UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROELETRÔNICA

ALAN CARLOS JUNIOR ROSSETTO

Modeling and Simulation of Self-heating Effects in *p*-type MOS Transistors

Thesis presented in partial fulfillment
of the requirements for the degree of
PhD in Microelectronics

Advisor: Prof. Dr.-Ing. Gilson Inacio Wirth

Porto Alegre
November 2018

CIP — CATALOGING-IN-PUBLICATION

Rossetto, Alan Carlos Junior

Modeling and Simulation of Self-heating Effects in *p*-type MOS Transistors / Alan Carlos Junior Rossetto. – Porto Alegre: PGMICRO da UFRGS, 2018.

149 f.: il.

Thesis (Ph.D.) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Microeletrônica, Porto Alegre, BR–RS, 2018. Advisor: Gilson Inacio Wirth.

1. CMOS. 2. Charge Traps. 3. Monte Carlo. 4. MOSFET. 5. Reliability. 6. Self-heating. I. Inacio Wirth, Gilson. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. Rui Vicente Oppermann
Vice-Reitora: Prof[a]. Jane Fraga Tutikian
Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves
Diretora do Instituto de Informática: Prof[a]. Carla Maria Dal Sasso Freitas
Coordenador do PGMICRO: Prof. Fernanda Gusmão de Lima Kastensmidt
Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

"(...) In our endeavor to understand reality we are somewhat like a man trying to understand the mechanism of a closed watch. He sees the face and the moving hands, even hears its ticking, but he has no way of opening the case. If he is ingenious he may form some picture of a mechanism which could be responsible for all the things he observes, but he may never be quite sure his picture is the only one which could explain his observations. (...)"

— ALBERT EINSTEIN

ACKNOWLEDGMENTS

At the completion of such a long journey, it may sometimes be difficult to identify each person who has contributed, either directly or indirectly, to the quality and content of my doctoral work. I will attempt to detail my gratitude herein, and apologize for any omissions.

First and foremost, I would like to thank my two thesis advisors and friends, Prof. Dr.-Ing. Gilson I. Wirth and Prof. Dr. Dragica Vasileska. As a CMOS reliability specialist, Prof. Gilson provided the necessary guidance and background I needed to go forward in a totally new research field to myself. Here, I also acknowledge him for supervising my research activities over the last 7 years, starting when I was an undergraduate student, through the master's degree and the completion of the current thesis.

The notorious expertise of Prof. Vasileska in the field of computational electronics also represented a remarkable improvement in the research reported here. Her guidance, patience, hospitality, friendship, and continuous support during the year I spent with her research group at Arizona State University are certainly something I will never ever forget. I also acknowledge the assistance and support of Dr. Gokul Kannan, M.Sc. Pradyumna Muralidharan, M.Sc. Robin Daugherty, and M.Sc. Suleman Qazi during my stay in Tempe, AZ.

I am also deeply in debt to my colleagues at Laprot/UFRGS. Dr. Vinicius Camargo provided the isothermal version of the Monte Carlo code for holes, as well as valuable discussions pertaining the simulation dynamic and flow. Dr. Fabio Vidor and Dr. Thiago Both substantially contributed with ideas and insights. The companionship of M.Sc. Pablo Vaz, M.Sc. Paulo Aguirre, Dr. Raphael Brum, M.Sc. Thales Becker, and M.Sc. Thiago Copetti is also acknowledged, as well as the fruitful discussions with B.Sc. Caroline Soares and M.Sc. Gabriela Furtado, for whom I am passing the torch.

At the conclusion of writing this Ph.D. thesis, I realize how much I owe to my family. I am very grateful to my parents, Nelci and Cleusa, and to my siblings, Jean and Janaina, for their unconditional support and love. I am also exceedingly thankful to my beloved wife Monique, my safe harbor and companion for both good and *not that good* times.

Finally, I would like to thank the Universidade Federal do Rio Grande do Sul for providing public and high-quality education, as well as the CAPES Funding Agency for the financial support that made this work possible.

ABSTRACT

The complementary metal-oxide-semiconductor (CMOS) scaling process of the recent decades, coupled with new device structures and materials, has aggravated thermal problems and turned them into major reliability issues for deeply-scaled devices. As a consequence, the thermal transport dynamic and its impact on the device performance at submicron dimensions is established as a contemporary theme. In this context, a new self-consistent electro-thermal particle-based device simulator for the study of self-heating effects in *p*-type metal-oxide-semiconductor field-effect transistors (MOSFETs) based in silicon is developed and presented. The electrical module of the tool utilizes the Ensemble Monte Carlo method to perform the charge transport, whereas the thermal module evaluates the non-isothermal temperature profiles by solving the phonon energy balance equations for both acoustic and optical phonon baths. These temperature profiles are fed back into the electrical module, which adjusts the carriers' scattering rate accordingly, thus, properly accounting for the device current capability degradation. The developed tool proved to be suitable for sub-100 nm device simulations, and it was used to perform relevant case study simulations of 24-nm channel length bulk and fully-depleted silicon-on-insulator (FD-SOI) MOSFETs. General device parameters extracted from the simulations are qualitatively in agreement with the expected behavior, as well as data from the literature, ensuring the proper operation of the tool. Electro-thermal simulations of bulk and FD-SOI devices provided both acoustic and optical phonon temperature profiles across the transistor structure, as well as the heat generation map and the device power dissipation. Some results were also extracted via Joule heating thermal model, and they are presented for comparison. The current degradation due to self-heating was found to be significant for FD-SOI devices, but very modest for bulk ones. At a fixed bias point of $V_D = V_G = -1.5$ V, for instance, bulk devices presented a current variation of as much as -0.75%, whereas for FD-SOI devices it reached up to -8.82% for $T_{gate} = 400$ K. Hot spot acoustic (lattice) and optical phonon temperatures were extracted as a function of the applied bias for both topologies. The lattice temperature rise, for instance, exceeded 10 K and 150 K over the heat sink temperature for bulk and FD-SOI transistors, respectively, observing the same bias point and gate temperature presented earlier. The particle-based nature of the tool is also suitable for the study of the impact of trap activity in MOSFETs and its interplay with self-heating effects. Simulations of charge traps were used to analyze the statistical distribution of the current deviations in 25-nm bulk MOSFETs

due to traps. The simulations showed that these deviations are exponentially-distributed, as experimentally observed and reported in the literature. Electro-thermal simulations of charge traps in bulk and FD-SOI transistors revealed that the largest degradation on the device current occurs when the effects of self-heating and trap activity take place simultaneously. At lower biases, the impact of charge traps dominates the current degradation, whereas the self-heating component prevails for larger biases.

Keywords: CMOS. Charge Traps. Monte Carlo. MOSFET. Reliability. Self-heating.

Modelagem e Simulação dos Efeitos de Auto Aquecimento em Transistores MOS do Tipo P

RESUMO

Nas últimas décadas, o processo de constante redução das dimensões de dispositivos semicondutores — aliado a novas estruturas de dispositivos e novos materiais —, tem agravado os problemas térmicos em tais estruturas, tornando-os importantes limitadores na confiabilidade destes dispositivos. Como consequência, a dinâmica do transporte térmico, e o impacto desta no desempenho de dispositivos semicondutores de escala nanométrica, se estabelece como um tema bastante atual. Neste contexto, esta tese desenvolve e apresenta um novo simulador eletrotérmico de dispositivos do tipo *particle-based* voltado ao estudo dos efeitos de auto aquecimento (*self-heating*) em transistores de efeito de campo de metal-óxido-semicondutor (MOSFETs) do tipo p baseados em silício. A ferramenta possui duas partes principais: um módulo elétrico — que realiza o transporte de carga baseado no método de Monte Carlo —, e um módulo térmico — que determina os perfis não-isotérmicos de temperatura dentro do dispositivo através da resolução das equações do balanço energético entre fônons óticos e acústicos. Tais perfis de temperatura são fornecidos ao módulo elétrico que, por sua vez, ajusta a frequência de espalhamento dos portadores com base na temperatura na vizinhança destes, permitindo ao simulador capturar a degradação na capacidade de corrente dos transistores devido aos efeitos de auto aquecimento. A ferramenta se mostrou adequada para a simulação de dispositivos de até 100 nm de comprimento de canal, e foi utilizada para realizar simulações estudo de caso de transistores MOS nas tecnologias *bulk* e *fully-depleted silicon-on-insulator* (FD-SOI) com 24 nm de comprimento de canal. Parâmetros elétricos de tais dispositivos extraídos via simulação se mostraram coerentes com o comportamento esperado e com dados da literatura, o que assegura o correto funcionamento do simulador. A partir de simulações eletrotérmicas, foram extraídos os perfis de temperatura ao longo da estrutura dos transistores, além do mapeamento da geração de calor dentro do dispositivo e a potência por este dissipada. Para comparação, alguns resultados também foram extraídos a partir do modelo de aquecimento Joule. A degradação na corrente do dispositivo devido ao efeito de auto aquecimento se mostrou muito mais significativa para transistores de tecnologia FD-SOI do que para transistores de tecnologia *bulk*. Para uma polarização de $V_\mathrm{D} = V_\mathrm{G} = -1.5$ V, por exemplo, transistores *bulk* apresentaram uma variação média na

corrente de -0.75%, enquanto que a corrente dos transistores FD-SOI variou, em média, -8.82%, considerando $T_{\text{gate}} = 400$ K. A temperatura de pico dos fônons acústicos (rede cristalina) e óticos foi extraída em função da polarização aplicada para ambas as topologias. Para a rede cristalina, por exemplo, o aumento da temperatura em relação à temperatura do dissipador foi de 10 K e 150 K para dispositivos *bulk* e FD-SOI, respectivamente, observando as mesmas condições apresentadas anteriormente. A característica *particle-based* da ferramenta também se mostrou adequada para o estudo da atividade de armadilhas (*traps*) em MOSFETs, bem como a interação desta com os efeitos de auto aquecimento. A simulação de armadilhas foi utilizada para analisar a distribuição estatística de impacto destas na corrente de dispositivos *bulk* com 25 nm de comprimento de canal. Tais simulações mostraram que as variações na corrente dos transistores devido à ação de armadilhas são exponencialmente distribuídas, em concordância com observações experimentais reportadas na literatura. Simulações eletrotérmicas de armadilhas de carga em transistores *bulk* e FD-SOI revelaram que a maior degradação na corrente do dispositivo ocorre quando os efeitos de auto aquecimento e atividade de armadilhas ocorrem simultaneamente. Para tensões de polarização mais baixas, o impacto das armadilhas de carga domina a degradação de corrente, enquanto que a degradação devido ao auto aquecimento prevalece para tensões mais elevadas.

Palavras-chave: Auto-aquecimento, Armadilhas de carga, CMOS, Confiabilidade, Monte Carlo, MOSFET.

LIST OF ABBREVIATIONS AND ACRONYMS

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
ASU	Arizona State University
BTE	Boltzmann Transport Equation
BTI	Bias Temperature Instability
CMOS	Complementary Metal-oxide-semiconductor
EMC	Ensemble Monte Carlo
FinFET	Fin Field-effect transistor
FD-SOI	Fully-depleted Silicon-on-insulator
LHS	Left-hand Side
LU	Lower-upper
MOS	Metal-oxide-semiconductor
MOSFET	Metal-oxide-semiconductor Field-effect Transistor
PEB	Phonon Energy Balance
RDF	Random Dopant Fluctuations
RHS	Right-hand Side
SHE	Self-heating Effects
SIP	Strong Implicit Procedure
SOI	Silicon-on-insulator
SOR	Successive-over-relaxation

LIST OF SYMBOLS

Symbol	Description
A	Inverse valence-band parameter
$a_{1,2}$	Energy dependent parameters
C_{A}	Volumetric heat capacity of acoustic phonons
C_{OP}	Volumetric heat capacity of optical phonons
dt	Differential time
d_{t}	Trap depth within the gate oxide
D_{J}	Junction depth
D_{op}	Optical phonon deformation potential
D_{T}	Total silicon depth
$D_{\mathrm{t}}K$	Deformation potential constant for optical and intervalley-phonon scattering with carriers
E	Electric field
	Energy
E'	Carrier final energy
E_{ac}	Acoustic deformation potential
E_{avg}	Average carrier energy
$E_{\mathrm{avg}}(x)$	Average carrier energy along the transistor length
$E_{\mathrm{avg}}(x,y,z)$	Average carrier energy for each position within the transistor
E_{c}	Energy at the edge of the conduction band
E_{g}	Band gap energy
E_{i}	Particle initial energy
$E_{\mathbf{k}}$	Particle energy at the state $\mathbf{k}$
$E_{\mathbf{k}'}$	Particle energy at the state $\mathbf{k}'$
E_{v}	Energy at the edge of the valence band

$f(\mathrm{Y})$	Attenuation of the trap impact as a function of its depth into the gate oxide
$f(\mathrm{Y_i})$	Attenuation of the impact of a trap at the position $\mathrm{Y_i}$
$g(\theta,\phi)$	Warping of the bands as a function of the angles θ and ϕ
$\hbar$	Reduced Planck constant
H	Volumetric heat generation rate
$H(x)$	Heat generation along the transistor length
$H(x,y,z)$	Heat generation for each position within the transistor
$H_{\mathrm{J\cdot E}}$	Volumetric heat generation evaluated via Joule heating model
H_{PEBs}	Volumetric heat generation evaluated via phonon energy balance equations model
i_{max}	Maximum number of mesh points along the x-axis
	Largest mesh point index along the x-axis
i_{min}	Lowest mesh point index along the x-axis
I_{D}	Distribution of device drain currents
$\overline{I_{\mathrm{D}}}$	Average drain current for an ensemble of devices
I_{D0}	Reference distribution of device drain currents
$\overline{I_{\mathrm{D0}}}$	Average reference drain current for an ensemble of devices
$I_{\mathrm{D0,i}}$	Reference drain current for a given transistor i
$I_{\mathrm{D}}(x_{\mathrm{t}},d_{\mathrm{t}},z_{\mathrm{t}})$	Drain current distribution for an ensemble of devices with a trap in the position $p_{\mathrm{t}}(x_{\mathrm{t}},d_{\mathrm{t}},z_{\mathrm{t}})$
$I_{\mathrm{D,i}}(x_{\mathrm{t}},d_{\mathrm{t}},z_{\mathrm{t}})$	Drain current for a transistor i with a trap in the position $p_{\mathrm{t}}(x_{\mathrm{t}},d_{\mathrm{t}},z_{\mathrm{t}})$
J	Electric current density
J_{n}	Electric current density due to electrons
j_{max}	Maximum number of mesh points along the y-axis
	Largest mesh point index along the y-axis
j_{min}	Lowest mesh point index along the y-axis

j_{OX}	Mesh points in the gate oxide
J_{p}	Electric current density due to holes
$\mathbf{k}$	wave vector of a carrier
$\mathbf{k}'$	wave vector of a carrier after scattering
k_{B}	Boltzmann constant
$k_{\max}$	Maximum number of mesh points along the z-axis
	Largest mesh point index along the z-axis
$k_{\min}$	Lowest mesh point index along the z-axis
L	Transistor channel length
L_{C}	Contact length
L_{D}	Drain region length
L_{G}	Gate length
L_{S}	Source region length
L_{T}	Total device length
L_{tr}	Square root of the trap impact area L_{tr}^2
L_{tr}^2	Trap impact area
m	Final band
m^*	Carrier effective mass
m_0	Electron rest mass
M	Number of thermal Gummel Cycles
n	Initial band
	Carrier concentration (in cm^{-2})
N	Number of Monte Carlo iterations
	Number of particles in the system
	Number of transistors in the ensemble
N_{A}	Source/drain doping concentration

N_{car}	Maximum number of carriers
N_{D}	Substrate doping concentration
N_{i}	Intrinsic doping concentration
N_{OP}	Thermal equilibrium number of optical phonons
N_{q}	Thermal equilibrium number of phonon with wave vector q
N_{sub}	Substrate doping concentration
$p_{\mathrm{t}}(x_{\mathrm{t}}, d_{\mathrm{t}}, z_{\mathrm{t}})$	Trap position within the device
P	Power dissipated as heat
$P(t)$	Probability that a carrier will suffer its next scattering event during a time dt around t
$P_{\mathrm{h,ac}}(E)$	Acoustic phonon scattering probability per unit time as a function of the energy
$P_{\mathrm{h,op}}(E)$	Non-polar optical phonon scattering probability per unit time as a function of the energy
$P_{\mathrm{J \cdot E}}$	Power dissipated as heat evaluated via Joule heating model
P_{PEBs}	Power dissipated as heat evaluated via phonon energy balance equations model
q	Elementary charge (with its sign)
$\mathbf{r}$	Position in real space
r	Uniformly distributed random number
R	Carrier generation/recombination rate
$SE_{\overline{\upsilon}_{\mathrm{D}}}$	Standard error of the drift velocity
t	Time
t_{BOX}	Buried oxide thickness
t_{dep}	Depletion region thickness (or depth)
t_{max}	Maximum simulation time
t_{OX}	Gate oxide thickness

t_r	Random free flight time
t_sat	Amount of time required to a carrier to cross the device channel traveling at the velocity saturation limit
t_Si	Silicon layer thickness
t_tran	Simulation transient time
t_w	Length of the simulation time window
T	Temperature
T_A	Acoustic phonon temperature
$T_\mathrm{A}(i,j,k)$	Acoustic phonon temperature at a given position $p(i,j,k)$ within the mesh
T_c	Carrier temperature
T_gate	Gate electrode temperature
T_{i_max}	Temperature at the surface with $i = i_\mathrm{max}$
T_{i_min}	Temperature at the surface with $i = i_\mathrm{min}$
T_ini	Initial temperature
T_{j_max}	Temperature at the surface with $j = j_\mathrm{max}$
T_{j_min}	Temperature at the surface with $j = j_\mathrm{min}$
$T_\mathrm{J\cdot E}$	Acoustic phonon temperature (lattice) evaluated via Joule heating model
T_L	Lattice temperature
T_OP	Optical phonon temperature
$T_\mathrm{OP}(i,j,k)$	Optical phonon temperature at a given position $p(i,j,k)$ within the mesh
T_PEB	Acoustic phonon temperature evaluated via PEB model
T_ph	Generic phonon temperature
T_sink	Heat sink temperature
V_D	Voltage applied at the drain electrode
V_DS	Voltage applied between the drain and source electrodes
V_G	Voltage applied at the gate electrode

V_{GS}	Voltage applied between the gate and source electrodes
V_{j}	Scattering potential of a given process j
V_{TH}	Transistor's threshold voltage
W	Device width
W_{A}	Energy density for the acoustic phonon bath
W_{c}	Carrier energy density
W_{OP}	Energy density for the optical phonon bath
x_{D}	Position along the transistor length of the metallurgical junction between the channel and the drain region
x_{max}	Maximum extension of the transistor along the transistor length
x_{S}	Position along the transistor length of the metallurgical junction between the source and the channel region
x_{t}	Trap position along the transistor channel length
$\mathrm{X_i}$	Random trap position along the transistor channel length
y_{max}	Maximum extension of the transistor along the transistor depth
y_{OX}	Maximum extension of the gate oxide above the Si/SiO_2 interface
y_{SD}	Position along the transistor depth of the metallurgical junction between the source/drain and the substrate region
$\mathrm{Y_i}$	Random trap depth into the gate oxide
z_{t}	Trap position along the transistor width
$\beta(E')$	Non-parabolicity factor as a function of the energy E'
Γ	Total scattering rate including self-scattering
Γ_{j}	Scattering rate of a given process j
Γ_{self}	Self-scattering rate
δI_{D}	Current deviation distribution for an ensemble of devices
$\delta I_{\mathrm{D,calc}}$	Current deviation calculated via uniform channel theory
$\delta I_{\mathrm{D,i}}$	Extrapolated current deviation due to a trap for a given transistor i

$\delta I_{\mathrm{D}}(\mathrm{X})$	Current deviation distribution for an ensemble of devices as a function of the trap position along the transistor length
$\delta I_{\mathrm{D}}(\mathrm{X_i})$	Drain current deviation due to a trap at the position $\mathrm{X_i}$
$\delta I_{\mathrm{D}}(x_{\mathrm{t}}, d_{\mathrm{t}}, z_{\mathrm{t}})$	Current deviation distribution for an ensemble of devices with a trap at the position $p_{\mathrm{t}}(x_{\mathrm{t}}, d_{\mathrm{t}}, z_{\mathrm{t}})$
$\delta I_{\mathrm{D,i}}(x_{\mathrm{t}}, d_{\mathrm{t}}, z_{\mathrm{t}})$	Current deviation due to a trap at $p_{\mathrm{t}}(x_{\mathrm{t}}, d_{\mathrm{t}}, z_{\mathrm{t}})$ for a given transistor i
Δ	Mesh spacing
ΔE_{ij}	Energy gap between band i and band j
Δn	Carrier concentration variation (in cm^{-2})
Δt	Time step
$\Delta \overline{I_{\mathrm{D}}}$	Average current deviation for an ensemble of devices
ε	Material electrical permittivity
$\varepsilon_{\mathrm{Si}}$	Silicon electrical permittivity
θ	Azimuthal angle between the states $\mathbf{k}$ and $\mathbf{k}'$
κ	Lattice thermal conductivity for silicon
κ_{A}	Acoustic phonon thermal conductivity for silicon
κ_{Si}	Thermal conductivity of silicon
λ_{D}	Debye length
μ	Carrier mobility
ρ	Carrier concentration
ρ_{m}	Material mass density
σ	Transistor channel conductance
$\sigma_{\overline{\upsilon}_{\mathrm{D}}}$	Standard deviation of the drift velocity
τ	Energy transfer time scale
$\tau_{\mathrm{c-A}}$	Carrier-acoustic phonon relaxation time
$\tau_{\mathrm{c-L}}$	Carrier-lattice scattering time constant

$\tau_{\mathrm{c-OP}}$	Carrier-optical phonon relaxation time
$\tau_{\mathrm{c-ph}}$	Carrier-phonon relaxation time
$\tau_{\mathrm{OP-A}}$	Optical phonon-acoustic phonon relaxation time
υ_{c}	Single carrier drift velocity
υ_{D}	Carrier drift velocity
$\overline{\upsilon}_{\mathrm{D}}$	Average drift velocity
$\upsilon_{\mathrm{D}}^{n}$	Drift velocity for the nth particle
$\upsilon_{\mathrm{D}}(x)$	Carrier drift velocity along the transistor length
$\overline{\upsilon}_{\mathrm{D}}(m\Delta t)$	Average drift velocity at the mth time step
$\upsilon_{\mathrm{D}}^{n}(m\Delta t)$	Drift velocity for the nth particle at the mth time step
υ_{Dsat}	Carrier saturation drift velocity
$\upsilon_{\mathrm{D}t}$	Absolute (ot total) carrier drift velocity
$\upsilon_{\mathrm{D}x}$	Carrier drift velocity along the x-axis
$\upsilon_{\mathrm{D}y}$	Carrier drift velocity along the y-axis
$\upsilon_{\mathrm{D}z}$	Carrier drift velocity along the z-axis
υ_{s}	Sound velocity in the material
υ_{T}	Carrier thermal velocity
ϕ	Polar angle between the states $\mathbf{k}$ and $\mathbf{k}'$
ϕ_{f}	Fermi potential
ϕ_{n}	Quasi-fermi level for electrons
ϕ_{p}	Quasi-fermi level for holes
ω	Phonon angular frequency
Ω	Crystal volume
ω_{0}	Initial phonon angular frequency
ω_{op}	Optical phonon angular frequency
ω_{p}	Plasma frequency

LIST OF FIGURES

Figure 2.1 Flowchart of a typical ensemble Monte Carlo device simulation. 34
Figure 2.2 Selection of the type of scattering terminating a free flight in the Monte Carlo algorithm..40
Figure 2.3 Ensemble Monte Carlo simulation in which a time step Δt is introduced over which the motion of particles is synchronized. 41
Figure 2.4 Flowchart for the electric portion of the ensemble Monte Carlo simulator developed in the work of Camargo (2016). 43
Figure 2.5 Hole drift velocity as a function of the electric field and carrier mobility as a function of the acceptor density for silicon. 45

Figure 3.1 Diagram and characteristics time scales of the energy transfer processes in silicon...50
Figure 3.2 Self-consistent loop scheme: exchange of variables between the electric and thermal kernels and choice of the proper scattering table. 60

Figure 4.1 Acoustic phonon scattering rates for holes scattering from the heavy hole, light hole, and split-off bands for temperatures of 300 K, 400 K, and 500 K..65
Figure 4.2 Non-polar optical phonon scattering rate for holes scattering from the heavy hole band for temperatures of 300 K, 400 K, and 500 K. 66
Figure 4.3 Histogram of the energy of the carriers during a simulation as a function of the applied bias for a bulk MOSFET..66
Figure 4.4 Silicon thermal conductivity as a function of the temperature. 71

Figure 5.1 Flowchart of the electro-thermal simulator...75
Figure 5.2 Thermal boundary conditions for bulk and FD-SOI case study transistors. 77
Figure 5.3 Simulation structure for bulk and FD-SOI case study transitors. 78
Figure 5.4 Hot spot temperature and number of iterations needed to solve the heat equation as a function of the solver tolerance. 82
Figure 5.5 Average carrier energy as a function of the time step for different biases...83
Figure 5.6 Average carrier energy as a function of the mesh spacing. 85
Figure 5.7 Number of particles (carriers) being used in the simulation along time.86
Figure 5.8 $I_D \times V_G$ characteristics of the bulk and FD-SOI case study MOSFETs.87
Figure 5.9 $I_D \times V_D$ characteristics of the bulk and FD-SOI case study MOSFETs.88

Figure 6.1 Average carrier energy as a function of the position along the transistor length for 100 bulk MOSFETs..91
Figure 6.2 Average carrier drift velocity as a function of the position along the transistor length for 100 bulk MOSFETs..91
Figure 6.3 Average electric field profile as a function of the position along the transistor length and depth for 100 bulk MOSFETs. 92
Figure 6.4 Distribution of the carrier travel time from source to drain as a function of the bias for 100 case study bulk MOSFETs. 92
Figure 6.5 Average carrier temperature along the transistor structure for 100 case study bulk MOSFETs...93
Figure 6.6 Average carrier concentration along the transistor structure for 100 case study bulk MOSFETs...93

Figure 6.7 Average acoustic phonon (lattice) temperature profile evaluated via the PEB model and Joule heating model for 100 case study bulk MOSFETs..95
Figure 6.8 Average optical phonon temperature profile for 100 case study bulk MOSFETs. 96
Figure 6.9 Average acoustic phonon peak temperature and optical phonon peak temperature as a function of the applied bias for 100 case study bulk MOSFETs. 97
Figure 6.10 Average volumetric heat generation within the transistor evaluated via phonon energy balance model and Joule heating model for 100 case study bulk MOSFETs. 99
Figure 6.11 Average heat generation along the transistor length according to the PEB model and Joule heating model for 100 case study bulk MOSFETs.100
Figure 6.12 Average power supplied to the transistor and average power dissipated as heat for each thermal model as a function of the applied bias for 100 case study bulk MOSFETs..101
Figure 6.13 Drain current as a function of the temperature Gummel cycle for isothermal and non-isothermal simulations. 103
Figure 6.14 Average drain current deviation for isothermal and non-isothermal simulations as a function of the applied bias. 104
Figure 6.15 Percentage of acoustic and optical scattering events as a function of the applied bias. 104
Figure 6.16 Average carrier energy along the transistor length for 100 case study FD-SOI MOSFETs..106
Figure 6.17 Average carrier drift velocity along the transistor length for 100 case study FD-SOI MOSFETs...106
Figure 6.18 Average electric field profile along the transistor length and depth for 100 case study FD-SOI MOSFETs. 107
Figure 6.19 Distribution of the carrier travel time from source to drain as a function of the applied bias for 100 case study FD-SOI MOSFETs. 107
Figure 6.20 Average carrier concentration and average carrier temperature along the transistor silicon layer for 100 case study FD-SOI MOSFETs................108
Figure 6.21 Average acoustic phonon temperature profile evaluated via PEB model and Joule heating model, and optical phonon temperature profile for 100 case study FD-SOI MOSFETs. 110
Figure 6.22 Average acoustic phonon and optical phonon temperature profiles for 100 case study FD-SOI MOSFETs with silicon thermal conductivity depending only on the temperature...111
Figure 6.23 Average optical phonon, acoustic phonon, and lattice peak temperature for 100 case study FD-SOI MOSFETs as a function of the applied bias and gate temperature...112
Figure 6.24 Volumetric heat generation profile evaluated via Joule heating model and PEB model for 100 case study FD-SOI MOSFETs. 114
Figure 6.25 Average heat generation along the transistor length for 100 case study FD-SOI MOSFETs...115
Figure 6.26 Electric power supplied to the device and power dissipated as heat for each thermal model according to the applied bias for 100 case study FD-SOI MOSFETs. 116
Figure 6.27 Average drain current deviation for isothermal and non-isothermal simulations as a function of the applied bias. 117

Figure 6.28 Drain current as a function of the temperature Gummel cycle for isothermal and non-isothermal simulations for multiple heat sink and gate temperatures. 118
Figure 6.29 Average total carrier drift velocity and total carrier drift velocity variation along the transistor length as a function of the gate temperature for non-isothermal simulations of 100 case study FD-SOI MOSFETs. 119
Figure 6.30 Current deviation distribution from $N = 400$ transistors with a single charge trap located in the middle of the conduction channel. 122
Figure 6.31 Normalized average drain current deviation caused by a single trap as a function of its position along the channel length. 123
Figure 6.32 Normalized average current deviation caused by a single trap as a function of the its depth into the oxide. 124
Figure 6.33 Distribution of drain current deviations estimated considering 10^6 charge traps randomly positioned within the transistor gate oxide. 126
Figure 6.34 Normalized current degradation due to self-heating and trap activity for bulk and FD-SOI MOSFETs as a function of the applied bias. 127
Figure 6.35 Normalized current deviation due to a single trap at the Si/SiO_2 interface as a function of the transistor channel area and applied bias. 128

Figure A.1 Graphical representation of the test structure. 145
Figure A.2 Simulated temperature profile for Case I and Case II. 146
Figure A.3 Thermal conductivity profile and temperature profile for Case III. 146
Figure A.4 Temperature profile for Case IV. 147
Figure A.5 Thermal conductivity profile and temperature profile for Case V. 147

LIST OF TABLES

Table 3.1 Comparison among the phonon models presented in Subsection 3.2.3. ... 56

Table 4.1 Description of the physical parameters and constants utilized for scattering rates calculation. ... 65

Table 5.1 Physical dimensions of the bulk and FD-SOI case study MOSFETs. ... 79

Table 5.2 Main simulation parameters for bulk and FD-SOI case study MOSFETs. 80

Table 6.1 Average drain current for isothermal and non-isothermal simulations as a function of the applied bias for 1000 case study bulk MOSFETs. The current variation for each case is calculated over the current extracted from the isothermal simulation with $T_{\text{sink}} = 300$ K. ... 102

Table 6.2 Average drain current for isothermal and non-isothermal simulations as a function of the gate temperature and applied bias. ... 117

CONTENTS

1 INTRODUCTION......29
2 ENSEMBLE MONTE CARLO METHOD......33
2.1 Scattering Rates......35
2.2 Free-Flight Time Generation......37
2.3 Scattering Events......39
2.4 Ensemble Monte Carlo Simulation......40
2.5 Device Simulation Utilizing the Ensemble Monte Carlo Method......42
3 THE LATTICE HEATING PROBLEM......47
3.1 Physical Mechanism......48
3.2 Modeling of Self-Heating......50
3.2.1 Joule Heating Model......51
3.2.2 Carrier-Lattice Scattering......52
3.2.3 Phonon Model......52
3.2.4 Phonon Monte Carlo Simulator......55
3.3 ASU Model Description......55
4 DEVELOPMENT OF THE THERMAL MODULE......63
4.1 Temperature Dependent Scattering Tables......63
4.2 Solving the Phonon Energy Balance Equations in 3D Coordinates......67
4.3 Thermal Conductivity of Silicon Dependence on the Temperature......69
4.4 Thermal Conductivity Dependence on the Silicon Doping and Thickness......71
5 ELECTRO-THERMAL SIMULATION PROCEDURE......73
5.1 Simulation Flow......73
5.2 Thermal Boundary Conditions......74
5.3 Case Study Devices......76
5.4 Optimization of Simulation Parameters......80
5.4.1 Thermal Solver Tolerance......81
5.4.2 Time Step......81
5.4.3 Mesh Size......84
5.4.4 Number of Particles in the Simulation......85
5.5 Device Characterization......86

6 RESULTS AND DISCUSSION **89**
6.1 Electro-thermal Simulations of Bulk MOSFETs **89**
6.1.1 Carrier Profiles 89
6.1.2 Temperature Profiles 94
6.1.3 Heat Generation Profile 98
6.1.4 Impact of Self-heating on the bulk MOSFET On-Current 102
6.2 Electro-thermal Simulations in FD-SOI MOSFETs **105**
6.2.1 Carrier Profiles 105
6.2.2 Temperature Profiles 109
6.2.3 Heat Generation Profile 113
6.2.4 Impact of Self-heating on the FD-SOI MOSFET On-Current 116
6.3 Particle-based Simulations of Charge Traps in MOSFETs **120**
6.3.1 Statistical Simulations of Charge Traps in Bulk MOSFETs 121
6.3.2 Electro-thermal Simulations of Charge Traps in Bulk and FD-SOI MOSFETs 125
7 CONCLUSION **129**
REFERENCES **133**
APPENDIX A — VALIDATION OF THE THERMAL SOLVER **145**
APPENDIX B — PUBLICATIONS **149**

1 INTRODUCTION

Over the years, CMOS technology scaling has provided continuous improvements to integrated circuits performance and density, and the constant trend to reduce the transistor size has led to the emergence of nanoscale devices. On the other hand, not only advantages for the semiconductor structures have arisen from this process. Several issues that played minor roles in older technologies became major or even critical reliability limiters for modern technology nodes. Particularly noteworthy are the effects of random dopant fluctuations (RDF) (BHAVNAGARWALA; TANG; MEINDL, 2001; MAHMOODI; MUKHOPADHYAY; ROY, 2005; GERRER et al., 2012) and charge trap activity (KACZER et al., 2011; THEODOROU et al., 2015).

The shrink of device dimensions has also lead to increased power density levels generated in the channel, which can be directly translated into a temperature rise in the device active region (MAKOVEJEV; OLSEN; RASKIN, 2011). Although thermal phenomena are intrinsically related to the operation of electronic devices, they are not directly responsible for electrical functionality and performance, but adversely affect transistor reliability and yield. It is well known, for instance, that the increase in the overall device temperature may degrade its current drive capability (TAKACS; TRAGER, 1987; CAVIGLIA; ILIADIS, 1993; JENKINS; RIM, 2002). Additionally, temperature variation across a single chip may induce mismatch issues for analog applications (TENBROEK et al., 1998), while device overheating is reported to be one of the most common causes of circuit failure (MITTAL; MAZUMDER, 2010).

The heat source causing such temperature rise of a semiconductor device (or circuit) can be either external to the structure — for instance, in a situation that the circuit is operating in a high-temperature environment —, or internal — when the heat is generated by the device itself. In the latter situation, the effect is appropriately called *self-heating*. The occurrence of self-heating in semiconductor devices is known since the 70's (POPESCU, 1970) and it was firstly reported for bulk MOS (SESNIC; CRAIG, 1972) and bipolar (GAUR; NAVON, 1976) technologies.

Nevertheless, self-heating effects started to capture significant attention with the emergence of silicon-on-insulator (SOI) devices (NEEL; HAOND, 1990; BERGER; CHAI, 1991; DALLMANN; SHENAI, 1995; TENBROEK et al., 1997). In these structures, the heat flow through the substrate is restricted, since the buried oxide layer has poor thermal conductivity. In addition, the thermal conductivity of silicon significantly decreases

as the silicon layer thickness is reduced to the nanometric scale (ASHEGHI et al., 1998; ASHEGHI et al., 2002), retaining the heat even more concentrated. As a consequence, the heat generated at the active area remains almost entirely confined in the silicon layer, and much higher temperature hot spots are expected for SOI in comparison with bulk devices (RALEVA et al., 2012). Self-heating has been also reported to occur in tri-gate (PRASAD et al., 2013) as well as bulk and SOI FinFET devices (LIU et al., 2014; JIANG et al., 2015; JEON; JHON; KANG, 2017). In these non-planar and also highly confined topologies, besides the typical performance degradation, self-heating acts as a transistor lifetime limiter (LEE et al., 2017; PRASAD; RAMEY; JIANG, 2017). Moreover, its occurrence is no longer restricted to extreme scenarios, being pervasive even at nominal usage conditions (PRASAD; RAMEY; JIANG, 2017).

In the literature, self-heating effects are widely covered for *n*-type MOS devices. On the other hand, there is a lack of studies and tools addressing these effects in *p*-type structures. From the CMOS design perspective, however, it is desirable that self-heating effects are well characterized for both device types. In addition, some reliability issues, such as the bias temperature instability (BTI), mainly affects *p*-type transistors (MAKABE; KUBOTA; KITANO, 2000; REDDY et al., 2005; STATHIS; ZAFAR, 2006), and the interplay with self-heating might stimulate further degradation in such devices. In this context, this thesis proposes a novel self-consistent electro-thermal particle-based device simulator for the study of self-heating effects in *p*-type MOS transistors based in silicon. The version of the simulator presented here covers planar technology, but it is easily scalable to any 3D structure, such as tri-gate and FinFET devices.

The electro-thermal tool itself is divided into two main modules. At the electrical portion of the simulator, the Ensemble Monte Carlo method is utilized to perform the charge transport by resolving the Boltzmann Transport Equation (BTE) self-consistently coupled with a Poisson equation solver. At the thermal portion of the simulator, the phonon energy balance equations are self-consistently solved for both acoustic and optical phonon baths utilizing data arising from the electrical module, providing the temperature profile for both acoustic and optical phonons at each point of the device. These temperatures are fed back into the electrical module, which adjusts the carriers' scattering rate accordingly, thus, properly accounting for the device current capability degradation. The developed tool proved to be suitable for sub-100 nm device simulations, and it was used to perform relevant case study simulations of 24-nm channel length bulk and fully-depleted silicon-on-insulator (FD-SOI) transistors.

The particle-based character of the tool is also suitable for the simulation of the impact of trap activity on the performance of *p*-type MOSFETs. The real-space treatment of the Coulomb interactions provided by the simulator allows one to account for the particle-like nature of carriers and dopants. In addition, the 3D-placement of the impurity atoms properly reproduces the RDF effect and allows the generation of unique random dopant profiles. The aforementioned features are key for the analysis of the impact of charge traps on the performance of MOS devices via computational simulations.

This work is organized as follows: first, a general and introductory overview of particle-based simulation and charge transport via the Ensemble Monte Carlo method is presented in Chapter 2. Still in the context of device simulation, Chapter 3 addresses the lattice heating problem in details; its physical mechanism and impact on the device performance, as well as the most commonly used modeling techniques. The development of the simulator thermal module is the subject of Chapter 4, which also addresses its coupling with the electrical module, turning the tool into an electro-thermal simulator. Next, the simulation procedure and the typical flowchart of an electro-thermal simulation are presented in Chapter 5, along with the optimization of some simulation parameters and the case study devices used in the scope of this work. Finally, Chapter 6 presents the results extracted for the case study structures via isothermal, electro-thermal, and charge trap simulations; and Chapter 7 gives the final remarks.

2 ENSEMBLE MONTE CARLO METHOD

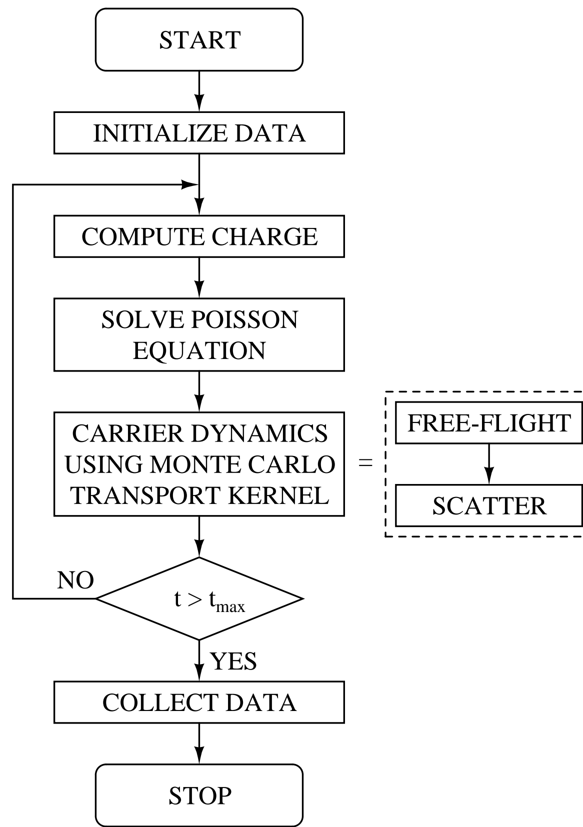
The Monte Carlo[1] method can be considered as a very general and established mathematical tool for the solution of a large variety of problems, being essentially based on the selection of random numbers (JACOBONI; LUGLI, 1989). The technique has been subject of numerous books and reviews (CASHWELL; EVERETT, 1959; RUBINSTEIN, 1981; JACOBONI; REGGIANI, 1983; JACOBONI; LUGLI, 1989; HESS, 1991; KALOS; WHITLOCK, 2008), and it is widely employed in many scientific fields, such as chemistry (DOLL; FREEMAN, 1994), physics (BINDER, 1995), biology (MANLY, 2007), finances (GLASSERMAN, 2003), etc.

With respect to the modeling and simulation of electronic devices, the application of the Monte Carlo method was firstly introduced by Kurosawa (1966), aiming to study high-field charge transport in semiconductors. Alike typical transport applications, the idea behind the simulation of charge transport problems through the Monte Carlo method is to keep track of a particle for a certain amount of time and emulate its free motion within a medium. The particle random walk is then subject to instantaneous random scattering events, which may or may not change either its trajectory or properties. In this way, the algorithm basically consists of generating random free flight times for each particle, choosing the type of scattering occurring at the end of the free flight, changing the final energy and momentum of the particle after scattering, and then repeating the procedure for the next free flight (VASILESKA; GOODNICK, 2002).

By simulating a one-particle system for a long enough time, one can gather statistics about the physical quantities of interest, such as the single particle distribution function, particle average energy, and the particle average drift velocity. This approach is known as *single-particle* Monte Carlo and it is appropriated for the calculation of steady-state carrier transport under uniform electric fields, such as the characterization of the charge transport in bulk materials. On the other hand, if the simulation is extended to an ensemble of particles — and the ensemble is representative of the entire physical system —, non-stationary time-dependent evolution of the carrier distribution under the influence of a time-dependent driving force can be simulated (VASILESKA; GOODNICK, 2002). The latter approach is the so-called *ensemble* Monte Carlo method and it is more useful for device simulation, where non-uniform electric fields typically occur.

[1]The origin of the term comes from the gambling casinos at the ward of Monte Carlo, in the city-state of Monaco. The method demands many repetitions of a random process, similar to what occurs in a casino, which wins against players on average (ROBERT, 2014).

Figure 2.1: Flowchart of a typical ensemble Monte Carlo device simulation.



Source: Adapted from Vasileska, Raleva and Goodnick (2010b).

The flowchart of a typical ensemble Monte Carlo simulation is depicted in Figure 2.1. As one can note, a Monte Carlo device simulation involves supplementary steps rather than just generating free-flight times and tracking particles. Since the Monte Carlo method is presented here in a conceptual manner, an extensive analysis of these additional features is out of the scope of this thesis. Nevertheless, an alternative literature, which covers the subject, is suggested to the reader. The work of Gross (1999) presents a full description of a Monte Carlo simulator for *n*-channel MOS devices, addressing specific aspects of each stage of the simulation. Besides the free flight-scatter kernel, the author describes in details the routines for carrier and parameters' initialization, charge assignment, Poisson equation solving, as well as particle-mesh coupling. The changes in the semiconductor band structure needed to adapt the tool for the simulation of *p*-channel MOS devices is covered in the work of Camargo (2016).

In this context, the following sections aim to present the essential characteristics that compose a typical Monte Carlo simulator. Section 2.1 introduces the scattering theory, also detailing the most relevant scattering processes which affect carrier transport in silicon. Section 2.2 addresses the free-flight time generation algorithm and the role of the

self-scattering. Section 2.3, in turn, presents the dynamics of the scattering events, as well as the methodology for choosing the final state after their occurrence. Next, Section 2.4 provides an overview about the simulation flow utilizing a synchronous scheme to coordinate the particle motion over a fixed time step, which allows the Monte Carlo method to be applied for multiple particles in a system. Finally, Section 2.5 details the simulation of semiconductor devices by employing the ensemble Monte Carlo method. The simulation flow is discussed in details and some results are presented.

2.1 Scattering Rates

In a semiconductor, free carriers interact with the crystal and with each other through a variety of scattering processes that modify the energy and momentum of the particles (VASILESKA; GOODNICK; KLIMECK, 2010). The probability of a carrier undergo a scattering event and experience a transition from some initial state $\mathbf{k}$ to some final state $\mathbf{k}'$ is quantified by the scattering rate. Basically, this probability depends on the type of carrier (electron or hole), the carrier energy, and the type of scattering mechanism, also varying from material to material.

In the context of a Monte Carlo simulation, scattering rates are calculated quantum mechanically via the solution of the time-dependent Schrödinger equation, evaluated using first-order perturbation theory (HE, 2000). Thus, the transition rate per unit time from an initial state $\mathbf{k}$ in band n to a final state $\mathbf{k}'$ in band m for the jth scattering mechanism is given by the expression (VASILESKA; GOODNICK; KLIMECK, 2010)

$$\Gamma_{\text{j}}(n,\mathbf{k};m,\mathbf{k}') = \frac{2\pi}{\hbar}\left|\left\langle m,\mathbf{k}'\middle|V_{\text{j}}(\mathbf{r})\middle|n,\mathbf{k}\right\rangle\right|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}} \pm \hbar\omega), \tag{2.1}$$

which is known as Fermi's Golden Rule. In Equation 2.1, $V_{\text{j}}(\mathbf{r})$ is the scattering potential of the process, and $E_{\mathbf{k}}$ and $E_{\mathbf{k}'}$ are the initial and final state energies of the particle, respectively (VASILESKA; GOODNICK; KLIMECK, 2010). The matrix element $|\langle m,\mathbf{k}'|V_{\text{j}}(\mathbf{r})|n,\mathbf{k}\rangle|^2$ incorporates the momentum conservation, while the delta function δ accounts for the energy conservation during the scattering process (HE, 2000). The total scattering rate out of state $\mathbf{k}$ is found by summing over all the final states $\mathbf{k}'$ for a given

band, which results in (VASILESKA; GOODNICK; KLIMECK, 2010)

$$\Gamma_{\mathrm{j}}(n,\mathbf{k}) = \frac{2\pi}{\hbar}\sum_{m,\mathbf{k}'}\left|\left\langle m,\mathbf{k}'\middle|V_{\mathrm{j}}(\mathbf{r})\middle|n,\mathbf{k}\right\rangle\right|^{2}\delta(E_{\mathbf{k}'}-E_{\mathbf{k}}\pm\hbar\omega), \tag{2.2}$$

or in the integral form (HATHWAR, 2011),

$$\Gamma_{\mathrm{j}}(n,\mathbf{k}) = \frac{\Omega}{(2\pi)^3}\int_0^{2\pi}d\phi\int_0^{\pi}\theta d\theta\int_0^{\infty}\Gamma_{\mathrm{j}}(n,\mathbf{k};m,\mathbf{k}')\mathbf{k}'^2 d\mathbf{k}'. \tag{2.3}$$

Here, Ω is the crystal volume, ϕ is the polar angle, and θ is the azimuthal angle between $\mathbf{k}$ and $\mathbf{k}'$.

The result obtained from Equation 2.3 is utilized to calculate the carrier scattering rate due to transitions induced by the different scattering sources present in the crystal. In silicon, the most important ones are phonons, impurities, and other carriers (JACOBONI; REGGIANI, 1983), where the former dominates the charge transport at room temperature and high electric fields (HE, 2000).

The interaction between carriers and phonons arises from lattice vibrations, which introduce small shifts in the semiconductor energy bands. Deviations of the bands from the static lattice position due to these small shifts lead to an additional potential that is responsible for the scattering process (FERRY, 2000). Such interactions can be divided into two types: acoustic phonon scattering and optical phonon scattering. Since the scattering potential $V_{\mathrm{j}}(\mathbf{r})$ is characteristic of each scattering mechanism, the matrix element for acoustic and optical phonon scattering is distinct from each other.

For acoustic phonon scattering, the matrix element is given by (FERRY, 2000)

$$\left|\left\langle m,\mathbf{k}'\middle|V_{\mathrm{j}}(\mathbf{r})\middle|n,\mathbf{k}\right\rangle\right|^{2} = \frac{E_{\mathrm{ac}}k_{\mathrm{B}}T}{\rho_{\mathrm{m}}\Omega\upsilon_{\mathrm{s}}^2}, \tag{2.4}$$

where E_{ac} is the acoustic deformation potential, k_{B} is the Boltzmann constant, T is the temperature, ρ_{m} is the mass density, and υ_{s} is the sound velocity.

For the non-polar[2] optical phonon scattering, in turn, the matrix element can be written as (FERRY, 1976)

$$\left|\left\langle m,\mathbf{k}'\middle|V_{\mathrm{j}}(\mathbf{r})\middle|n,\mathbf{k}\right\rangle\right|^{2} = \frac{\hbar(D_{\mathrm{t}}K)^2}{2\rho_{\mathrm{m}}\Omega\omega_0}\left(N_{\mathrm{q}}+\frac{1}{2}\pm\frac{1}{2}\right), \tag{2.5}$$

[2]Optical phonon scattering can be either polar or non-polar, depending on the material characteristic. In polar materials (e.g. gallium arsenide, indium phosphide, zinc sulfide, etc.), this mechanism is called *polar* optical phonon scattering.

where $(D_\text{t}K)^2$ is a optical coupling constant, which depends on the optical deformation potential and the phonon wave vector (JACOBONI; REGGIANI, 1983), ω_0 is the phonon frequency, and N_q is the phonon occupation factor, represented by the Bose-Einstein distribution as

$$N_\text{q} = \left[\exp\left(\frac{\hbar\omega_\text{q}}{k_\text{B}T}\right) - 1\right]^{-1}. \tag{2.6}$$

In Equation 2.5, the upper sign is for the emission of phonons and the lower sign is for the absorption of phonons by the carriers.

The total scattering rate for both acoustic and non-polar optical phonon scattering mechanisms can be obtained by substituting Equation 2.4 and Equation 2.5, respectively, into Equation 2.1 and then evaluate the integral. This step will not be discussed here but it is presented in details in the work of Hathwar (2011) and Camargo (2016) for electrons and holes, respectively. In addition, an extensive description of the scattering mechanisms which affect charge transport in semiconductor materials can be found in Jacoboni and Lugli (1989, chapter 2).

2.2 Free-Flight Time Generation

As already presented, within the Monte Carlo technique the dynamics of the carrier motion is assumed to consist of free flights terminated by instantaneous scattering events, which may change the momentum and energy of the particle (VASILESKA; GOODNICK, 2002). In order to determine the free flight times and simulate this process, however, the scattering probability density for the carrier at hand must be determined. The relationship between the free flight time and the scattering probability is developed in the following.

Let $\Gamma[\mathbf{k}(t)]dt$ be the probability that a carrier in the state $\mathbf{k}$ undergoes a scattering event during the time dt. Thus, the probability that a carrier which experienced a scattering event at the time $t = 0$ has not undergone another event after a time t is given by (JACOBONI; REGGIANI, 1983)

$$\exp\left\{-\int_0^t \Gamma[\mathbf{k}(t')]dt'\right\}, \tag{2.7}$$

which, basically, gives the probability of the interval [0, t] does not contain a scattering. Consequently, the probability $P(t)$ that a carrier will suffer its next scattering event during

dt around t is given by the joint probability (JACOBONI; REGGIANI, 1983)

$$P(t)dt = \Gamma[\mathbf{k}(t)]\exp\left\{-\int_0^t \Gamma[\mathbf{k}(t')]dt'\right\}dt. \tag{2.8}$$

In this way, random free flight times t_r can be generated from the probability $P(t)$ utilizing a direct method according to the expression (VASILESKA; GOODNICK, 2002)

$$r = \int_0^{t_\mathrm{r}} P(t)dt, \tag{2.9}$$

where r is a uniformly distributed random number in the range [0, 1]. By integrating Equation 2.9 with $P(t)$ defined by Equation 2.8, one gets (VASILESKA; GOODNICK, 2002)

$$r = 1 - \exp\left[-\int_0^{t_\mathrm{r}} \Gamma[\mathbf{k}(t')]dt'\right]. \tag{2.10}$$

In addition, since $1-r$ is statistically equivalent to r, Equation 2.10 can be rewritten to its final form

$$-\ln r = \int_0^{t_\mathrm{r}} \Gamma[\mathbf{k}(t')]dt'. \tag{2.11}$$

Equation 2.11 is the fundamental equation used to generate random free flight times within the Monte Carlo framework, resulting in a random walk process related to the underlying particle distribution function (VASILESKA et al., 2008).

Nevertheless, because of the complexity of the integral at hand, it is impractical to generate stochastic free flight times with the distribution of Equation 2.11 (JACOBONI; REGGIANI, 1983). Therefore, it is worthwhile the application of the so-called *self-scattering* method (REES, 1969). Self-scattering consists of a fictitious scattering mechanism whose scattering probability always adjusts itself in such a way the total scattering probability, i.e., real scattering plus self-scattering, is constant over time (VASILESKA; GOODNICK, 2002). In addition, if a carrier undergoes a self-scattering event, its state after the collision is taken to be equal to its state before the collision, causing the carrier trajectory to continue unperturbed as if no scattering had occurred (JACOBONI; REGGIANI, 1983). In this case, the total scattering probability Γ can be written as

$$\Gamma = \Gamma[\mathbf{k}(t)] + \Gamma_\mathrm{self}[\mathbf{k}(t)]. \tag{2.12}$$

where $\Gamma[\mathbf{k}(t)]$ is the real scattering probability, and $\Gamma_\mathrm{self}[\mathbf{k}(t)]$ is the self-scattering probability. Since now the dependency on time vanishes, Equation 2.11 can be rewritten as

$$-\ln r = \Gamma \int_0^{t_r} dt' \tag{2.13}$$

and the integral is trivially evaluated. Thus, free flight times can be obtained directly from the resulting expression

$$t_r = -\frac{1}{\Gamma}\ln r. \tag{2.14}$$

The self-scattering technique introduces significant simplification on the free flight time calculation, thus compensating the computer time spent taking care of such fictitious mechanism (JACOBONI; REGGIANI, 1983). One requirement to utilize the self-scattering approach is choosing the total scattering rate Γ so that it exceeds the maximum "real" scattering rate encountered during the simulation (VASILESKA; GOODNICK, 2002).

2.3 Scattering Events

Immediately upon the free flight time generated accordingly the algorithm presented in the previous section elapses, a scattering event takes place. At this point, the scattering mechanism that terminates the free flight must be known. The choice is performed based on the generation of a uniformly distributed random number r between 0 and the total scattering rate Γ, where Γ contains the cumulative scattering rates for the N mechanisms taken into account in the simulation plus the self-scattering, i.e.,

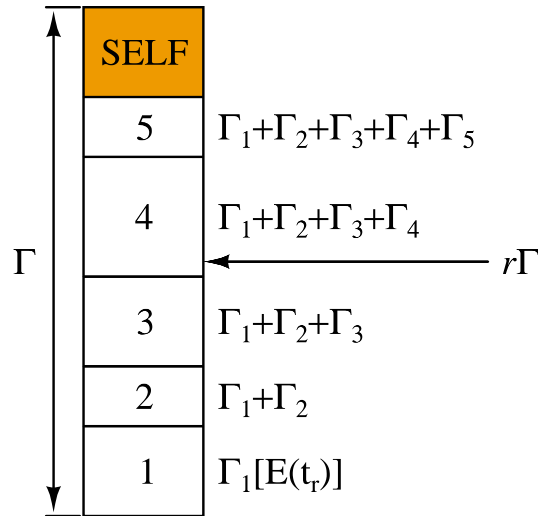
$$\Gamma = \Gamma_{\text{self}}[n,\mathbf{k}] + \Gamma_1[n,\mathbf{k}] + \Gamma_2[n,\mathbf{k}] + \cdots + \Gamma_{\text{N}}[n,\mathbf{k}]. \tag{2.15}$$

This process is graphically described in Figure 2.2.

Firstly, the product $r\Gamma$ is compared with the maximum scattering rate for a given energy due to mechanism 1. If the product is smaller than that, mechanism 1 is selected, otherwise $r\Gamma$ is compared again with the rate for mechanism 1 plus mechanism 2, and so on. In the case of the product $r\Gamma$ exceeds the maximum cumulative scattering rate for the real mechanisms, then the self-scattering is selected. Since the latter has a high occurrence probability — chiefly for low energy carriers —, it is usually checked first, for the sake of efficiency (CAMARGO, 2016).

After the scattering mechanism that terminates the particle free-flight being selected, the changes in the particle energy and momentum are treated accordingly. For in-

Figure 2.2: Selection of the type of scattering terminating a free flight in the Monte Carlo algorithm. In this particular case, mechanism 4 should be selected.



Source: Adapted from Vasileska, Raleva and Goodnick (2008).

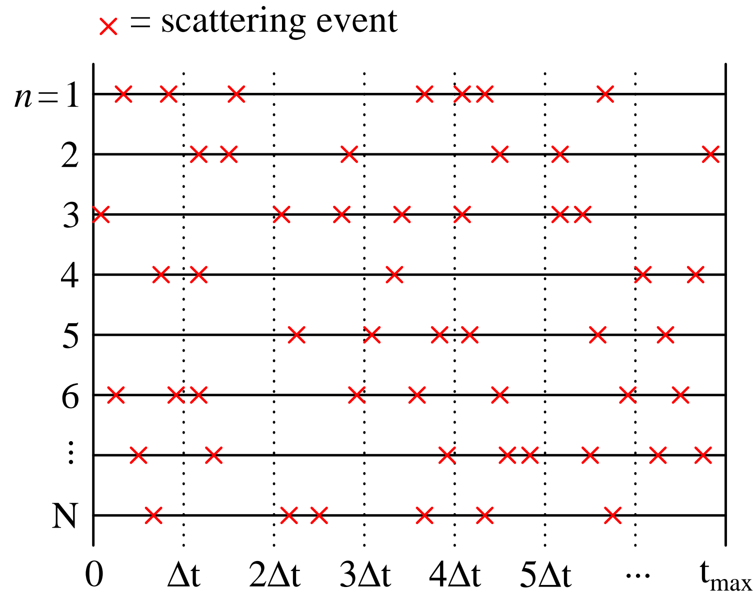
stance, if the scattering process at hand is elastic, which is the case for the acoustic phonon scattering, only the momentum is updated, since the particle energy remains unaltered. On the other hand, if the process is taken to be inelastic, both energy and momentum of the particle are modified.

2.4 Ensemble Monte Carlo Simulation

The techniques presented so far can be used for tracking either a single particle or multiple particles in a system. However, in order to get improved statistics — even with reduced simulation times —, as well as for transient simulations, the use of an ensemble of particles is preferred (VASILESKA; GOODNICK, 2002). In this case, the Monte Carlo algorithm is performed for each particle at once, and a fixed time step Δt is introduced in order to synchronize the motion of the particles. Basically, the time step represents the amount of time the algorithm spend keeping track of one particle before proceeding to the next one.

Figure 2.3 depicts the dynamic of a ensemble Monte Carlo simulation. In the beginning, i.e., time instant $t = 0$, the Poisson's equation is solved for the equilibrium condition, in order to provide initial values for the electric field and electric potential. After that, a random free flight time is generated for the first particle in the ensemble, and its motion is performed accordingly. Nevertheless, if the free flight time is larger than the

Figure 2.3: Ensemble Monte Carlo simulation in which a time step Δt is introduced over which the motion of particles is synchronized.



Source: Adapted from Vasileska et al. (2008).

time step, the particle is treated up to the end of the time step, and the remaining free flight time is stored to be utilized by the particle at hand when the next time step is performed.

Once the observation time elapses, the particle energy and momentum are updated, and the algorithm turns its attention to the second particle in the ensemble. This procedure is repeated until the Nth particle is treated. At that point, the Poisson's equation is solved — considering the change on the spatial particles distribution —, the electric field and electric potential are updated for each grid point, and the routine advances to the next time step. Here, the up to date value of the electric field is then used to govern the particles motion during the subsequent time window.

In this way, two main loops compose an ensemble Monte Carlo simulation: an inner loop, which goes over all the N particles of the ensemble; and an outer loop, which advances by increments of Δt until the total simulation time $t_{\max}$ is reached (VASILESKA et al., 2008). Finally, it is important to note that between two adjacent observation points, the system is considered to be "frozen", i.e., both the field and potential are constant. In addition, over each time step, the motion of each particle in the ensemble is performed independently of the other ones (VASILESKA et al., 2008).

From an ensemble of N particles, one can easily obtain the non-stationary particle distribution and its related quantities of interest, such as the drift velocity, energy, band or valley population, etc. These quantities are taken as averages over the ensemble at fixed

time steps (VASILESKA; GOODNICK; KLIMECK, 2010). For instance, the average drift velocity $\overline{\upsilon}_{\mathrm{D}}$ at a given time step $m\Delta t$ is estimated by the mean of the drift velocity υ_{D} for each particle individually, i.e.,

$$\overline{\upsilon}_{\mathrm{D}}(m\Delta t) \cong \frac{1}{N}\sum_{n=1}^{N}\upsilon_{\mathrm{D}}^{n}(m\Delta t) \qquad (2.16)$$

Since the result obtained in Equation 2.16 is an estimator of the "real" velocity, its standard error $SE_{\bar{\upsilon}_{\mathrm{D}}}$ given by (KALOS; WHITLOCK, 2008)

$$SE_{\overline{\upsilon}_{\mathrm{D}}} = \frac{\sigma_{\overline{\upsilon}_{\mathrm{D}}}}{\sqrt{N}}, \qquad (2.17)$$

where the standard deviation $\sigma_{\overline{\upsilon}_{\mathrm{D}}}$ can be expressed as

$$\sigma_{\overline{\upsilon}_{\mathrm{D}}} \cong \sqrt{\frac{N}{N-1}\left[\frac{1}{N}\sum_{n=1}^{N}(\upsilon_{\mathrm{D}}^{n})^{2} - \overline{\upsilon}_{\mathrm{D}}^{2}\right]}. \qquad (2.18)$$

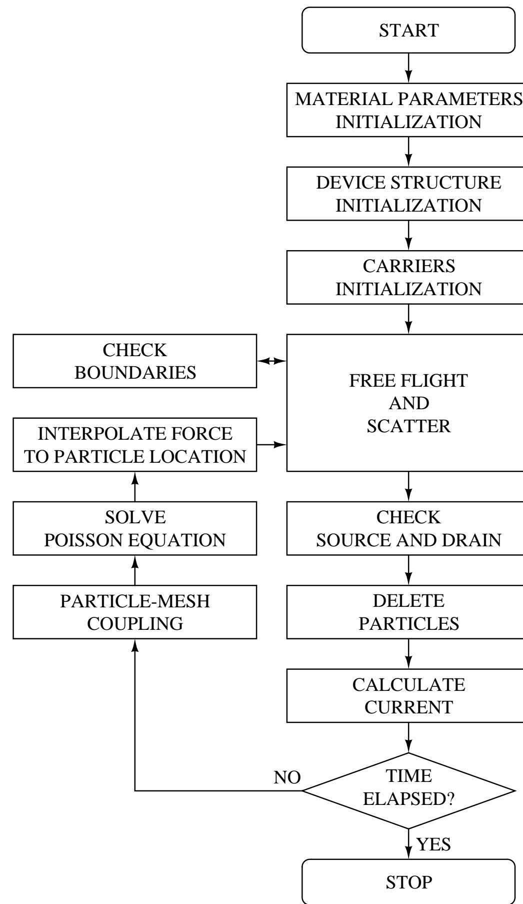
It is clear from Equation 2.17 that the estimation error reduces as the square root of the number of particles. In this way, Vasileska and Goodnick (2002) suggest that, for good statistics, typical ensemble ranges must be on the order of 10^4 to 10^5 particles.

2.5 Device Simulation Utilizing the Ensemble Monte Carlo Method

For the purpose of semiconductor device simulation, the ensemble Monte Carlo technique presented in the previous section — which is responsible for the charge transport dynamic —, must be self-consistently coupled to an appropriated field solver (VASILESKA et al., 2008). Within this scheme, the electric field, which arises from the solution of the Poisson's equation, is the driving force that accelerates the particles at the Monte Carlo phase. On the other hand, fixed (impurities) and mobile charges (carriers) compose the time-dependent charge profile of the device, which is fed back into the Poisson solver.

Furthermore, additional features are needed in order to facilitate the device simulation. In this context, Figure 2.4 depicts the complete flowchart for the electric portion of an ensemble Monte Carlo simulator, which was developed in the work of Camargo (2016) and it is employed in this thesis. In the following, each building block of the tool is briefly addressed, in order to provide to the reader an idea about the device simulation flow.

Figure 2.4: Flowchart for the electric portion of the ensemble Monte Carlo simulator developed in the work of Camargo (2016).



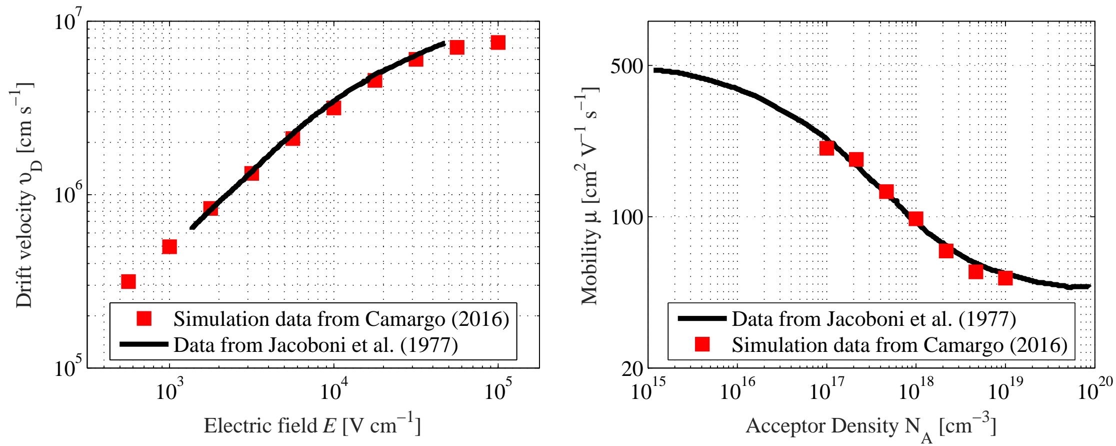
Source: Adapted from Camargo (2016).

- Step 1: *Material parameters initialization*: This routine loads the fundamental physical constants, specific features of the device to be simulated — e.g. width, length, oxide thickness, doping, etc. —, and additional parameters which will be used during the simulation, such as the temperature, the time step, the maximum simulation time, etc.
- Step 2: *Device structure initialization*: In this step, the continuous space and the discrete space (grid) are placed, the regions of the device are build up, and the boundaries, as well as the simulation domain, are defined.
- Step 3: *Carriers initialization*: This routine accounts for the initial number of carriers and their placement within the device, as well as their initial kinetic energy and initial momentum. Usually, particle number and position are assigned according to the doping profile, so that initially the system is charge neutral on the average (VASILESKA et al., 2008). Regarding the kinetic energy, the carriers are initialized according to a Maxwell-Boltzmann distribution (VASILESKA; GOODNICK; KLIMECK, 2010), i.e., the initial energy E_{i} for a given carrier is given by $E_{\mathrm{i}} = -3/2 k_{\mathrm{B}} T \ln(r)$, where r is an uniformly

distributed random number in the range]0, 1[. Owing to the initial energy for each carrier, the initial momentum is calculated by the energy-momentum relationship and the azimuthal and polar angles, which are randomly generated (CAMARGO, 2016).

- Step 4: *Free flight and scatter*: This routine accounts for the motion of the carriers according to the dynamic already discussed in the previous sections. Due to their movement, one or more particles might leave the simulation domain in this step. Therefore, the boundaries of the device must be continuously inspected, in order to assure no carrier exits the device unaccounted.
- Step 5: *Check source and drain*: here, the device contacts are checked in order to account for the number of carriers that have exited or entered the device through the contacts. Hence, the routine examines the grid cells that are adjacent to the contacts as well. If the cell at hand has extra positive charge, one or more carriers are swept out the simulation domain and, then, counted as if they exited the device. Similarly, if the cell has extra negative charge, one or more carriers are injected into the cell, counting as if they entered the device (CAMARGO, 2016).
- Step 6: *Delete particles*: after exiting the device — either via the contacts or not —, the particle at hand is no longer within the simulation domain and, consequently, it must be excluded from the analysis from that point on.
- Step 7: *Calculate current*: after deleting and injecting particles in the system, the charge flow through the contacts is used to calculate the terminal current. In the case where the substrate current is neglected, the current flowing through the source and drain contacts should be the same, ensuring charge conservation.
- Step 8: *Check simulation time*: at this point, the temporal evolution of the simulation is checked. If the current simulation time does not exceed the maximum simulation time t_{max}, another iteration is executed. Otherwise, the simulation is ended.
- Step 9: *Particle-mesh coupling*: the Poisson's equation is typically solved over a finite (discrete) grid, whereas the particles motion is performed over a continuous coordinate system. In this way, in order to map the particle charge from the continuous space to the discrete space, a charge redistribution scheme must be adopted. Basically, the particle charge is split into fractions of charge and then assigned to the adjacent grid points. Several charge redistribution techniques have been proposed, and the most utilized ones are the *nearest grid point*, *cloud-in-cell*, and *nearest element center* schemes (VASILESKA; GOODNICK; KLIMECK, 2010).
- Step 10: *Solve Poisson's Equation*: Owing to the charge profile arising from the

Figure 2.5: Hole drift velocity as a function of the electric field (*left panel*) and carrier mobility as a function of the acceptor density (*right panel*) for bulk silicon calculated with the ensemble Monte Carlo simulator developed in the work of Camargo (2016).



Source: Adapted from Camargo (2016).

particle-mesh coupling scheme, the Poisson's equation is solved and both potential and field are updated for each grid point.

- Step 11: *Interpolate force to particle location*: before proceeding to the free-flight-scatter phase, the force acting on the particle must be known. The force can be directly obtained from the electric field surrounding the particle. However, since the field is only known at the grid points, the same scheme applied in Step 9 is used here to interpolate the force from the discrete space to the carrier position in the real space.

In the approach presented here, there is no building block responsible for saving the output data, since this procedure is performed continuously during the simulation. Both snapshots and averaged values for the quantities of interest, such as the device current, the drift velocity, as well as the electric potential and electric field profile throughout the device, are stored. These results are then used to characterize the device under consideration.

Finally, in order to validate the aforementioned approach, Camargo extracted the hole drift velocity dependence on the electric field and the hole mobility dependence on the acceptor density for p-type bulk silicon from his simulator and compared the results with data from the literature. As depicted in Figure 2.5, the results are in agreement with those presented in Jacoboni et al. (1977).

3 THE LATTICE HEATING PROBLEM

Lattice heating is inherent to semiconductor device operation. It consists of a temperature rise in the device structure due to the dissipation of power as heat. Since the heat is generated by the device itself, this phenomenon is called *self-heating*. The temperature rise depends, basically, on the volumetric heat generation and the efficiency of the heat removal to the external medium. In regions with elevated temperature, some physical and electrical characteristics are degraded, affecting the overall device performance. In some critical cases, when the power dissipation exceeds a certain limit, an often-destructive phenomenon called *thermal runaway* may occur (POPESCU, 1970).

Self-heating effects have been reported and studied for a variety of semiconductor devices and topologies. Works in literature address this effect in bipolar structures (DENNISON; WALTER, 1989), JFETs (SESNIC; CRAIG, 1972), bulk and SOI LDMOS transistors (ARNOLD; PEIN; HERKO, 1994), regular MOSFETs (SHARMA; RAMANATHAN, 1984), strained-silicon MOSFETs (JENKINS; RIM, 2002) and power MOSFETs (BARLOW; DAVIS; LAZARUS, 1986), VLSI (TAKACS; TRAGER, 1987) and ULSI (MAUTRY; TRAGER, 1990) bulk MOSFETs, SOI MOSFETs (SU et al., 1994), multi-gate (MOLZER et al., 2006) and FinFET (POP; DUTTON; GOODSON, 2003) devices.

Although the occurrence of self-heating is known since the 70's (POPESCU, 1970), its effects started to receive major attention in recent years, due to the advent of nanoscale devices. The reduction on the transistor dimensions has intensified the electric field levels at the device junctions since the field increases with a decrease in the physical size (HATAKEYAMA; FUSHINOBU, 2008). In addition, the employment of a buried insulator layer in the transistor structure — characteristic of silicon-on-insulator devices — exacerbates the thermal dissipation problem, since the insulator has poor thermal conductivity and hinders the dissipation of heat (HATAKEYAMA; FUSHINOBU, 2008). As another consequence of the scaling process, more transistors are packed in a single die, increasing even more the power dissipation at the chip level (RALEVA et al., 2012). Consequently, the combination of all the aforementioned characteristics has turned self-heating into a major reliability issue for deeply-scaled CMOS technologies (HATAKEYAMA; FUSHINOBU, 2008; KAMAKURA et al., 2010).

Self-heating is known to change the charge transport dynamic and reduce the drain current of a device. In this context, the full understanding of the operation of a modern

transistor cannot be accomplished by assuming it as an isothermal problem (POP et al., 2001). Therefore, a rigorous device simulator must utilize a concurrent electrical and thermal model, which is capable of taking into account the lattice heating effect (LAI; MAJUMDAR, 1996). In addition, to properly address the temperature non-equilibrium phenomenon in silicon semiconductor devices, the modeling has to be performed in such a way that includes the three sub-systems involved in the process: carriers, optical phonons, and acoustic phonons (LAI; MAJUMDAR, 1996).

The following sections address the self-heating effect in details. Section 3.1 presents the physical mechanism which is responsible for the lattice heating phenomenon. Section 3.2 focus on the most commonly used modeling techniques and the state-of-the-art tools employed in the study of self-heating. Finally, Section 3.3 presents the approach developed at Arizona State University (ASU) for solving lattice heating problems in nanoscale devices, which is the method utilized in the context of this study.

3.1 Physical Mechanism

In the context of a MOSFET, whenever a gradient of electric potential exists between source and drain regions, a lateral electric field is formed across the device. Considering a typical bias configuration either for *n*- or *p*-channel transistors, this field peaks near to the device drain region (POP; SINHA; GOODSON, 2006). This field accelerates charge carriers from the source towards the drain. Once accelerated, the carriers gain energy and heat up.

The energy gained by the carriers can only be lost through scattering events. In silicon, the carriers can scatter with each other, with phonons, with material interfaces and imperfections, and with impurity atoms. Amongst these scattering mechanisms, carriers only lose net energy by scattering with phonons (POP; SINHA; GOODSON, 2006), since the others mainly affect the carrier momentum (LUNDSTROM, 2000). Basically, during scattering events, phonons absorb a fraction of the carriers energy and subsequently transfer it to the lattice, which warms up through a mechanism known as *Joule heating* (POP; SINHA; GOODSON, 2006).

Depending on its energy, a carrier can scatter with different phonon modes. At low electric fields, i.e. $|E| < 10^6$ V/m, the carriers mainly interact with acoustic phonons, since few carriers have enough energy to excite optical phonons (LAI; MAJUMDAR, 1996). Consequently, the carriers lose the excess energy directly to the acoustic phonon

bath. This mode has large group velocities, and quickly transport the heat out of the region of most intense scattering (POP et al., 2001). Therefore, both the carriers and the lattice exhibit similar levels of energy. Under these conditions, the ensemble carrier drift velocity υ_{D} linearly increases with the electric field E, and the carrier transport follows Ohm's Law (LAI; MAJUMDAR, 1996), i.e.,

$$\upsilon_{\mathrm{D}} = \mu E, \tag{3.1}$$

where μ is the carrier mobility.

Under high electric fields, i.e. $|E| \geq 10^6$ V/m, the carriers possess enough energy to interact with optical phonons, and the interaction rate can be high enough to saturate the ensemble carrier drift velocity (LAI; MAJUMDAR, 1996). At this point, the linear dependence of the carrier drift velocity with the electric field vanishes. In addition, some carriers might become highly energetic, being pushed far out from the equilibrium with the lattice. These carriers are called *hot carriers* and the concept of carrier temperature T_{c} is used to quantify this behavior (LAI; MAJUMDAR, 1996), i.e.,

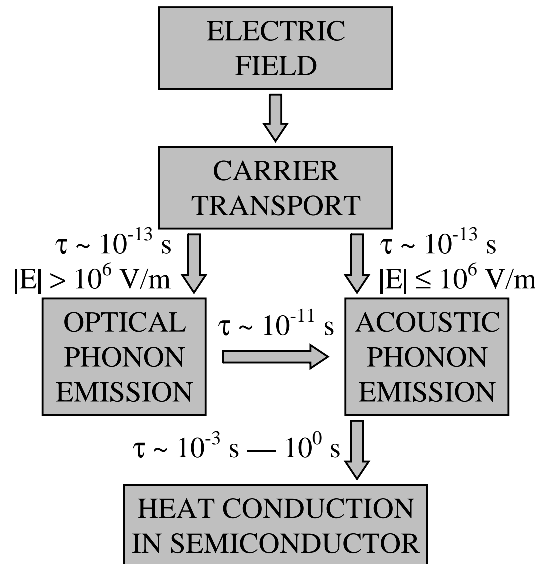
$$\frac{3}{2} k_{\mathrm{B}} T_{\mathrm{c}} = \frac{1}{2} m^{*} \upsilon_{\mathrm{T}}^{2}, \tag{3.2}$$

where υ_{T} stands for the thermal velocity and can be expressed in term of the carrier velocity υ_{c} and the ensemble carrier drift velocity υ_{D} as (LAI; MAJUMDAR, 1996)

$$\upsilon_{\mathrm{T}} = \upsilon_{\mathrm{c}} - \upsilon_{\mathrm{D}}. \tag{3.3}$$

Therefore, under high-field conditions, most of the carrier energy is transferred to the optical phonon bath. However, the optical mode has much smaller group velocity than the acoustic one and, thus, optical phonons remain relatively stationary (POP et al., 2001), not being able to transport the heat efficiently. On the other hand, optical phonons decay into acoustic phonons — which transport heat —, but in relatively long time scales ($\tau \sim 10^{-11}$ s) compared to the carrier-optical phonon scattering time ($\tau \sim 10^{-13}$ s). Consequently, an energy bottleneck may exist between the optical and acoustic phonon baths and it can cause the density of optical phonon modes to build up over time, leading to more scattering events and impeding carrier transport (POP; SINHA; GOODSON, 2006). This process is summarized in Figure 3.1.

Figure 3.1: Diagram and characteristics time scales of the energy transfer processes in silicon.



Source: Adapted from Pop et al. (2001).

3.2 Modeling of Self-Heating

Some of the earlier attempts to include self-heating in the simulation of semiconductor devices were performed by Gaur and Navon (1976), Adler (1978), and Chryssafis and Love (1979). These first models, however, were conceived mostly based on heuristic and empirical observations, thus, lacking physical foundation (LINDEFELT, 1994). Later, more refined modeling approaches emerged, deriving the heat conduction equation from the carrier energy continuity equations (WACHUTKA, 1990) and the carrier BTE (LINDEFELT, 1994). Coming after, Fushinobu, Majumdar and Hijikata (1995) and Lai and Majumdar (1996) derived for the first time the energy balance equations for the acoustic and optical phonons for germanium and silicon, respectively, allowing one to properly take into account the energy exchange between carriers and phonons. Their work represented a significant improvement for the approaches formulated thereafter.

In general, self-heating models differ from each other with respect to the computation of the heat source term. In this context, three different models are the most commonly used: (I) Joule heating, (II) carrier-lattice scattering, and (III) the phonon model. Although these three approaches yield identical results in equilibrium, under non-equilibrium conditions the results of the three models can vary significantly (RAMAN; WALKER; FISHER, 2003). Each individual model and its particularities are addressed in the following subsections.

3.2.1 Joule Heating Model

The Joule heating model consists of solving the heat diffusion equation using the Joule heating term as the source. In this approach, the heat source term H is computed directly from the dot product of the local electric field E and the electric current density J (GAUR; NAVON, 1976), i.e.,

$$H = J \cdot E \tag{3.4}$$

or hence

$$H = (J_\mathrm{n} + J_\mathrm{p}) \cdot E, \tag{3.5}$$

where J_n and J_p stand for the carrier current densities for electrons and holes, respectively. Therefore, the zone with the highest temperature will occur near the location where the product of the field and the current density is the largest.

In order to take into account the energy loss/gain to the lattice through recombination/generation processes, Adler (1978) extended the model to the form

$$H = \nabla \cdot \left(\frac{E_\mathrm{c}}{q} \cdot J_\mathrm{n} + \frac{E_\mathrm{v}}{q} \cdot J_\mathrm{p} \right), \tag{3.6}$$

where E_c and E_v are the band edges of conduction and valence bands, respectively. Expanding the divergent operator one gets

$$H = \frac{1}{q} \cdot J_\mathrm{n} \cdot \nabla E_\mathrm{c} + \frac{1}{q} \cdot J_\mathrm{p} \cdot \nabla E_\mathrm{v} + R \cdot (E_\mathrm{c} - E_\mathrm{v}), \tag{3.7}$$

where R is the generation/recombination rate, and the difference $E_\mathrm{c} - E_\mathrm{v}$ expresses the local band gap E_g. For non-degenerated materials, Equation 3.7 can be simplified as (SELBERHERR, 1984)

$$H = (J_\mathrm{n} + J_\mathrm{p}) \cdot E + R \cdot E_\mathrm{g}, \tag{3.8}$$

where the heat generated or consumed by the generation/recombination process becomes clearly apparent in Equation 3.8 when compared with Equation 3.5.

Chryssafis and Love (1979), on the other hand, proposed

$$H = \nabla \cdot (\phi_\mathrm{n} \cdot J_\mathrm{n} + \phi_\mathrm{p} \cdot J_\mathrm{p}), \tag{3.9}$$

where ϕ_n and ϕ_p are the quasi-fermi levels for electrons and holes, respectively. From the aforementioned models, however, Selberherr (1984, chapter 4) points out one should give

preference to Adler's approach. Addressing all the formulations proposed for the Joule heating generation modeling until that time, Wachutka (1990) proposed a more elaborated and sound theoretical model, taking into account for the heat generation in the stationary and transient regimes as well.

The Joule heating approach is widely used within the thermal package of commercial tools, such as SILVACO's Giga Simulator (SILVACO, 2018). Nevertheless, the model is unable to properly take into account the microscopic nature of the heat flow and the thermal non-equilibrium between the acoustic and optical phonon heat baths (VASILESKA; RALEVA; GOODNICK, 2008). In addition, this formulation also does not differentiate between carrier energy exchange with the various phonon modes and does not give any spectral information regarding the types of phonons emitted in the process (POP; SINHA; GOODSON, 2006).

3.2.2 Carrier-Lattice Scattering

Within the carrier-lattice scattering model, the heat generation in the semiconductor is taken to be due to temperature non-equilibrium between the electron/hole bath and the lattice. The thermal system is represented as a single lattice temperature and is considered to be in thermal equilibrium. The heat source is then taken as a scattering term obtained from the relaxation time approximation and moments of the BTE (RAMAN; WALKER; FISHER, 2003), having the form

$$H = \frac{3\rho k_{\mathrm{B}}}{2}\left(\frac{T_{\mathrm{c}} - T_{\mathrm{L}}}{\tau_{\mathrm{c-L}}}\right), \tag{3.10}$$

where T_{c} is the carrier temperature, T_{L} is the lattice temperature, and $\tau_{\mathrm{c-L}}$ is the carrier-lattice scattering time constant. However, since the heating term in this model is computed from the scattering of carriers with phonons, Pop, Sinha and Goodson (2006) state that only a simulation approach which deliberately incorporates all such scattering mechanisms will capture the complete microscopic and detailed picture of lattice heating.

3.2.3 Phonon Model

Phonon model is the most complete approach to study the thermal phenomena in nanoscale devices (LAI; MAJUMDAR, 1996). The expression, however, is not clearly

defined in the literature and, generally, the term is used to refer to a modeling approach in which a system of two different phonon modes is utilized to represent the thermal non-equilibrium condition (RAMAN; WALKER; FISHER, 2003; VASILESKA; RALEVA; GOODNICK, 2008). The first phonon model for silicon devices was derived by Lai and Majumdar (1996), and their formulation is presented in details in Section 3.3. After that, however, several other modeling approaches that take into account the energy transfer between carriers and phonons emerged. Some of them are addressed in the present subsection.

One of the best approaches for the modeling of thermal transport in submicron structures was proposed by Narumanchi, Murthy and Amon (2004) (VASILESKA; RALEVA; GOODNICK, 2008). They developed a tool that numerically solves the 2D phonon BTE utilizing the finite volume method (CHAI; LEE; PATNAKAR, 1994). Their model accounts for phonon dispersion and phonon polarization, as well as includes frequency-dependent phonon relaxation times. The tool, however, explores the problem from a physical point of view. Consequently, its primary focus is on the impact of lattice heating effects on material characteristics, rather than to a particular semiconductor structure. In this way, the authors do not present any data regarding self-heating effects on device reliability. On the other hand, they show their model can satisfactorily predict the specific heat and thermal conductivity of silicon over a range of temperatures. The predictions for bulk, as well as for doped and undoped silicon thin films, satisfactorily match experimental data.

Starting from a device-oriented perspective, E. Pop, S. Sinha, K. Goodson and coworkers proposed a model based on the Monte Carlo method for the study of thermal non-equilibrium in nanoscale silicon transistors (POP; DUTTON; GOODSON, 2004; POP; SINHA; GOODSON, 2006; POP, 2010). Their approach solves the phonon BTE taking into account the electron-phonon scattering term, as well as considering the phonon dispersion in the acoustic and optical branches. Unlike Lai and Majumdar's approach, which requires the solution of the phonon energy balance equations, Pop's model provides a detailed picture of the power dissipation within the device by computing it as the sum of all phonon emission minus all phonon absorption events. This model has already been applied to the analysis of heat generation and transport processes in n-type bulk and strained silicon structures (POP; DUTTON; GOODSON, 2004), 90 nm channel-length n-channel MOSFET (SINHA et al., 2006), and 1D $n+/n/n+$ silicon device (ROWLETTE; GOODSON, 2008).

Another model dedicated to analyzing the electro-thermal behavior of bulk silicon CMOS devices was formulated by Hatakeyama and Fushinobu (2008). As a case study structure, they utilized a structure composed of a pair of 90-nm channel length n-channel and p-channel MOSFETs. The formulation was derived from the hydrodynamic model and takes into account both carrier generation/recombination process and non-equilibrium between charge carriers and phonons. The tool provides good insights into the temperature cross-talk effect in a CMOS cell, but the impact of the temperature on the device current is not deeply investigated. In addition, since the hydrodynamic model is applied to the charge transport as well, this approach may not be suitable for the study of deeply scaled devices.

Kamakura et al. (2010) proposed a tool for analyzing the transient electro-thermal behavior of a 10-nm channel length $n{-}i{-}n$ device in silicon. The simulator solves the coupled transient BTE for carriers and phonons together, however, neglecting the contribution of optical phonons. The energy stored in the optical modes could be dissipated only through the conversion into acoustic modes. The authors present data for device current, lattice and phonon temperature distribution. Moreover, they validate the solution of the transient phonon BTE by calculating the silicon thermal conductivity for a wide range of temperatures, which matches experimental data. Nevertheless, despite its accuracy, this approach is highly resource consuming, despite being a 1D tool (NGHIEM; SAINT-MARTIN; DOLLFUS, 2014). Recently, their simulation approach was extended to n-channel double-gate MOSFET devices (KAMAKURA et al., 2014).

Ni et al. (2012) developed a tool which couples the anisotropic phonon BTE simulation to the computation of the phonon generation spectrum calculated using the Monte Carlo method. The simulator was utilized to study the hot spot temperature in bulk n-channel MOSFET devices. In addition, the tool also provides a good understanding on the heat generation and heat transport due to different phonon modes. They present, for instance, that the optical phonon mode can be responsible for as much as 25% of the heat transport in silicon, which in agreement with some works reported in the literature (BEECHEM et al., 2010; TIAN et al., 2011). The main drawback of their approach, however, is that the simulator is able to perform only one-way simulations. In other words, the temperature solution is not coupled back to the electric portion of the code and, thus, it does not affect the carrier-phonon scattering and the charge transport. In this way, the impact of self-heating on the device current cannot be captured by their tool.

More recently, Nghiem, Saint-Martin and Dollfus (2014) introduced a self-consist-

ent electron-phonon transport model, which was applied to investigate self-heating effects in a 20 nm-long double gate *n*-channel MOSFET. The developed phonon transport model includes both the phonon generation rates given by the electron Monte Carlo simulation and the decay of optical phonons into acoustic ones. The carrier transport was performed by a particle-based Monte Carlo code, self-consistently coupled to a 2D Poisson's equation solver, whereas the heat transport — based on the Boltzmann formalism — was carried out in 1D. The authors provide detailed insights into electro-thermal effects in the case study device, such as the impact on phonon distribution, lattice temperature, electric potential, carrier energy, and drain current. A comparison among the phonon models presented in this subsection is provided in Table 3.1.

3.2.4 Phonon Monte Carlo Simulator

Phonon Monte Carlo simulations are commonly used for modeling extreme nanoscale and hot-carrier devices. Basically, it utilizes the Ensemble Monte Carlo method to perform both charge and heat transport in the semiconductor. Similarly to carriers, which follow a gradient of electric field, phonons follow a gradient of temperature. Thus, the heat transport is carried out accordingly. Phonons are treated as quasi-particles that carry heat energy, obeying a corresponding Boltzmann transport equation, which can be used to study their transport (SHAIK, 2016).

In addition, the concept of scattering tables is also employed within the thermal phase of the tool, in order to generate the corresponding scattering rates for the phonon-phonon interactions. Hence, this approach is typically able to account for phonon creation and phonon annihilation processes as well. The main drawback of such type of tool, however, is its own complexity and the computational cost the simulations may require. A state-of-the-art tool developed for the study of heat transport by phonons which utilizes the Monte Carlo method is presented in the work of Shaik (2016).

3.3 ASU Model Description

As earlier presented, in order to properly treat the lattice heating phenomenon without making any approximations in the problem at hand, one has to solve the coupled Boltzmann transport equations for the carrier and phonon systems together (VASI-

Table 3.1: Comparison among the phonon models presented in Subsection 3.2.3.

Author	Device		Dimensionality		Does the model evaluate		
	Type	Channel length	Electrical module	Thermal module	Current degradation?	Phonon temp. non-equilibrium?	Heat generation?
Sinha et al. (2006)	*n*-channel MOSFET	90 nm	2D	2D	✓	✓	✓
Rowlette and Goodson (2008)	$n+/n/n+$ device	20 nm	1D	1D	✓	✓	✓
Hatakeyama and Fushinobu (2008)	*n*- and *p*-channel MOSFETs	90 nm	2D	2D	✓		✓
Kamakura et al. (2010)	*n-i-n* diode	10 nm	1D	1D	✓	✓	
Ni et al. (2012)	*n*-channel MOSFET	50 nm	3D	2D		✓	✓
Kamakura et al. (2014)	*n-i-n* DG-MOSFET	12 nm	2D	1D	✓	✓	
Nghiem et al. (2014)	*n*-channel DG-MOSFET	20 nm	2D	1D	✓	✓	✓
This work	*p*-channel MOSFET	24 nm	3D	3D	✓	✓	✓

LESKA; GOODNICK; KLIMECK, 2010). In other words, one has to solve the coupled carrier−optical phonons−acoustic phonons−heat bath problem, where each individual sub-process involves different time scales and has to be addressed in a somewhat peculiar manner and included in the global picture via a self-consistent loop (VASILESKA; RALEVA; GOODNICK, 2008).

In addition, for a full physical understanding of the lattice heating dynamic in semiconductors, phonon dispersion, and phonon polarization effects, as well as frequency-dependent relaxation times, must be considered in the analysis (NARUMANCHI; MURTHY; AMON, 2004). However, the solution of the carrier-phonon coupled set of equations including all the aforementioned characteristics becomes a formidable task even for today's high-performance computing systems (VASILESKA; RALEVA; GOODNICK, 2008). Consequently, it is clear that simplifications in the global problem are needed.

In this way, if one is primarily focused on calculating the impact of self-heating on the I-V characteristics of semiconductor devices, the heat generation and transport can be treated in a more approximate manner, but still more accurate than the local heat conduction model (VASILESKA; RALEVA; GOODNICK, 2010b). In this context, the foundation of the ASU approach is the analysis performed by Lai and Majumdar (1996). Starting from the phonon Boltzmann equations, they derived energy balance equations separately for the optical phonon and the acoustic phonon baths. The quantities involved in the process can be straightforwardly related to carrier transport parameters, but some assumptions are necessarily made. Optical phonons are considered to have zero group velocity, and constant relaxation times are used. Hence, it is assumed the heat transport is isotropic.

Their analysis was firstly presented for electrons in gallium arsenide (FUSHINOBU; MAJUMDAR; HIJIKATA, 1995) and in silicon (LAI; MAJUMDAR, 1996), however, the same formulation can also be applied to holes in both material. Starting from the principle of energy conservation, phonon energy balance equations in silicon can be developed as follows. Under high electric fields, carriers lose energy to optical phonons and optical phonons decay into acoustic phonons. Thus, the energy conservation equations for optical and acoustic phonons can be written as (LAI; MAJUMDAR, 1996)

$$\frac{\partial W_{\mathrm{OP}}}{\partial t} = \left(\frac{\partial W_{\mathrm{c}}}{\partial t}\right)_{\mathrm{coll}} + \left(\frac{\partial W_{\mathrm{OP}}}{\partial t}\right)_{\mathrm{coll}} \tag{3.11}$$

and

$$\frac{\partial W_{\mathrm{A}}}{\partial t} = \nabla \cdot (\kappa_{\mathrm{A}} \nabla T_{\mathrm{A}}) - \left(\frac{\partial W_{\mathrm{OP}}}{\partial t}\right)_{\mathrm{coll}}, \tag{3.12}$$

where W_{OP}, W_{A} and W_{c} are optical phonon, acoustic phonon, and carrier energy densities, respectively, κ_{A} is the acoustic phonon thermal conductivity, and T_{A} is the acoustic phonon temperature[1]. Due to their zero group velocity, there is no heat flux associated with the optical phonon mode and, thus, there is no divergence term on the right-hand side of Equation 3.11. Hence, the phonon energy densities can also be associated with the volumetric heat capacity of their respective mode, i.e.,

$$dW_{\mathrm{OP}} = C_{\mathrm{OP}} \cdot dT_{\mathrm{OP}} \tag{3.13}$$

and

$$dW_{\mathrm{A}} = C_{\mathrm{A}} \cdot dT_{\mathrm{A}}, \tag{3.14}$$

where C_{OP} and C_{A} represent the heat capacity of optical and acoustic phonons, respectively, and T_{OP} is the optical phonon temperature. The volumetric heat capacities C_{OP} and C_{A} can be estimated from the Einstein and Debye models (SZE, 2006), respectively.

The collision terms are expressed utilizing the relaxation time approximation, which results in (FUSHINOBU; MAJUMDAR; HIJIKATA, 1995)

$$\left(\frac{\partial W_{\mathrm{c}}}{\partial t}\right)_{\mathrm{coll}} = -\left[\frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}} - T_{\mathrm{ph}}}{\tau_{\mathrm{c-ph}}}\right) + \frac{1}{2}\frac{\rho m^* \upsilon_{\mathrm{D}}^2}{\tau_{\mathrm{c-ph}}}\right] \tag{3.15}$$

and

$$\left(\frac{\partial W_{\mathrm{OP}}}{\partial t}\right)_{\mathrm{coll}} = -C_{\mathrm{OP}}\frac{T_{\mathrm{OP}} - T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}, \tag{3.16}$$

where ρ is the carrier concentration, k_{B} is the Boltzmann constant, T_{c} is the carrier temperature, m^* is the effective mass, υ_{D} is the carrier drift velocity, and T_{ph} can be either the optical or acoustic phonon temperature, depending on which type of phonons the carriers interact with. In the same way, $\tau_{\mathrm{c-ph}}$ can be either carrier-optical phonon or carrier-acoustic phonon relaxation time, whereas $\tau_{\mathrm{OP-A}}$ stands for the optical phonon-acoustic phonon relaxation time. Finally, combining the Equations 3.11−3.16, one gets

$$C_{\mathrm{OP}}\frac{\partial T_{\mathrm{OP}}}{\partial t} = \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}} - T_{\mathrm{OP}}}{\tau_{\mathrm{c-OP}}}\right) + \frac{\rho m^* \upsilon_{\mathrm{D}}^2}{2\tau_{\mathrm{c-OP}}} - C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}} - T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right) \tag{3.17}$$

[1] The subscript *coll* means *due to collisions*.

and

$$C_{\mathrm{A}}\frac{\partial T_{\mathrm{A}}}{\partial t} = \nabla\cdot(\kappa_{\mathrm{A}}\nabla T_{\mathrm{A}}) + C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}}-T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right). \tag{3.18}$$

The first two terms in the right-hand side of Equation 3.17 represent the energy gain from the carriers, while the last term is the energy transferred to the acoustic phonons. The latter appears as a gain term on the right-hand side of Equation 3.18, and the first term on the right-hand side of this equation accounts for the heat diffusion.

Under low electric fields, however, the carriers do not possess enough energy to interact with the optical phonons and, thus, lose their energy directly to the acoustic ones. In this case, Equation 3.12 simplifies to

$$C_{\mathrm{A}}\frac{\partial T_{\mathrm{A}}}{\partial t} = \nabla\cdot(\kappa_{\mathrm{A}}\nabla T_{\mathrm{A}}) - \left(\frac{\partial W_{\mathrm{c}}}{\partial t}\right)_{\mathrm{coll}} \tag{3.19}$$

and it can be rewritten as

$$C_{\mathrm{A}}\frac{\partial T_{\mathrm{A}}}{\partial t} = \nabla\cdot(\kappa_{\mathrm{A}}\nabla T_{\mathrm{A}}) + \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}}-T_{\mathrm{A}}}{\tau_{\mathrm{c-A}}}\right) + \frac{1}{2}\frac{\rho m^{*}\upsilon_{\mathrm{D}}^{2}}{\tau_{\mathrm{c-A}}}, \tag{3.20}$$

where $\tau_{\mathrm{c-A}}$ stands for the carrier-acoustic phonon relaxation time. Carrier-phonon relaxation times are reported in the literature and they may range from 0.1 ps (FERRY, 2000) to 0.3 ps (TIEN; MAJUMDAR; GERNER, 1998) for carrier-optical phonon, and from 6 ps to 10 ps for carrier-acoustic phonon interactions (TIEN; MAJUMDAR; GERNER, 1998).

Finally, in order to cover the situation when carriers interact with both optical and acoustic phonons, Equation 3.18 must be expanded to account for the portion of carrier energy lost directly to the acoustic phonon bath (TIEN; MAJUMDAR; GERNER, 1998). In this way, the phonon energy balance equations can be expressed as

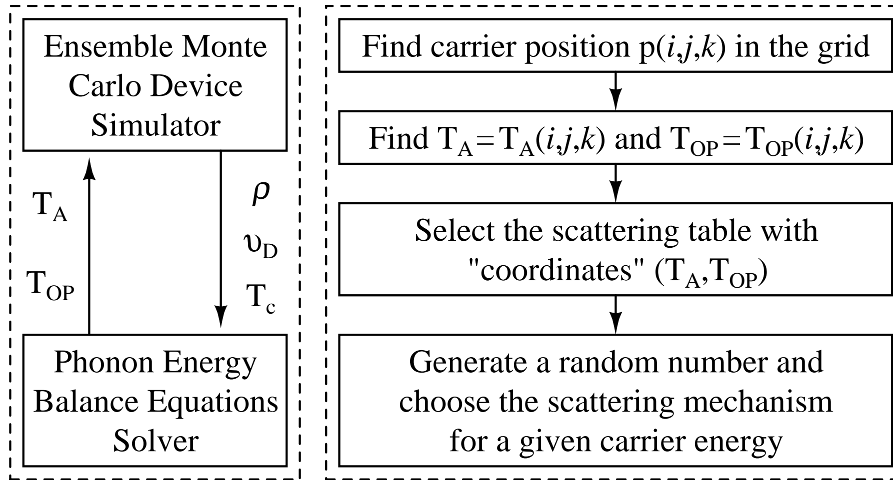
$$C_{\mathrm{OP}}\frac{\partial T_{\mathrm{OP}}}{\partial t} = \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}}-T_{\mathrm{OP}}}{\tau_{\mathrm{c-OP}}}\right) + \frac{\rho m^{*}\upsilon_{\mathrm{D}}^{2}}{2\tau_{\mathrm{c-OP}}} - C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}}-T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right) \tag{3.21}$$

and

$$C_{\mathrm{A}}\frac{\partial T_{\mathrm{A}}}{\partial t} = \nabla\cdot(\kappa_{\mathrm{A}}\nabla T_{\mathrm{A}}) + C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}}-T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right) + \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}}-T_{\mathrm{A}}}{\tau_{\mathrm{c-A}}}\right), \tag{3.22}$$

where the last term in the right-hand side of Equation 3.22 represents the energy loss to the acoustic phonons. Nevertheless, if the carrier-acoustic phonon scattering is treated as elastic, the computation of this term must be suppressed from the analysis (VASILESKA; RALEVA; GOODNICK, 2008). Note that the acoustic phonon thermal conductivity κ_{A}

Figure 3.2: Self-consistent loop scheme: exchange of variables between the electric and thermal kernels (*left panel*), and choice of the proper scattering table (*right panel*).



Source: Adapted from Vasileska, Raleva and Goodnick (2008).

and acoustic phonon temperature T_A can be approximated to the lattice thermal conductivity κ and temperature T_L, respectively.

Within the ASU approach, steady-state version of Equations 3.21 and 3.22 are solved self-consistently within a thermal module. Parameters such as the ensemble carrier drift velocity υ_D, carrier concentration ρ, and carrier temperature T_c are determined at the electric portion of the simulator and then transferred to the thermal module, which is responsible for solving the temperatures for acoustic and optical phonons. However, in order to properly couple the fluid-like characteristic of the equations within the thermal module with the particle-like characteristic of the variables stemming from the electric module, both time and space averaging must be applied to these quantities. Hence, the smoothing of such variables is necessary because most of the mesh points, especially at the interface and pinch-off region, are rarely populated with carriers (VASILESKA; RALEVA; GOODNICK, 2010b). The exchange of variables scheme between the two kernels is depicted in the left panel of Figure 3.2.

As the lattice temperature is expected to vary along the device, it is unpractical to calculate the scattering rate for each particle based on its temperature inside the Monte Carlo loop. Rather than that, *temperature dependent* scattering tables are created for a wide range of acoustic and optical phonon temperatures. Even though the calculation of the scattering rate is no longer needed, the process of addressing these scattering tables still involves additional steps to the Monte Carlo phase. Now, in order to choose randomly a scattering mechanism for a given carrier energy, it is also necessary to find the

corresponding scattering table for a given temperature.

To do that, every time a scattering event takes place, the position of the carrier at hand in the grid has to be found, in order to determine the acoustic and optical phonon temperatures in the vicinity of the particle. With this information, a scattering table with "temperatures coordinates" (T_A, T_{OP}) is selected and, according to the carrier energy, a random scattering rate for a certain mechanism is read from that table. For temperatures whose scattering table is not directly tabulated, an interpolation scheme is employed to provide a better estimative of the scattering rate. This process is illustrated in the right panel of Figure 3.2.

Regarding the simulation flow, the system starts at room temperature. Once the steady-state is reached, the electric parameters start to be sampled for each iteration. After a predefined time elapses, these variables are averaged and transferred to the thermal kernel, which solves the phonon energy balance equations and updates the acoustic and optical temperatures for each mesh point. By utilizing the temperatures provided by the thermal module, the scattering routine at the electric kernel is then able to include the temperature effect on the scattering rate as described above, which immediately affects the charge transport. This process is repeated for several times during the simulation until the desired convergence is achieved. In this way, both thermal and electric kernels are constantly exchanging up to date parameters with each other, providing the self-consistency needed for the study of self-heating effects on the I-V characteristics of semiconductor devices.

The model described in this section has already been applied to the study of self-heating effects in *n*-channel FD-SOI MOSFET devices (RALEVA et al., 2008; RALEVA et al., 2012), heat transport in 22-nm channel length planar *n*-channel MOSFET devices using multi-scale approach (SHAIK, 2016), and it is currently being implemented for *n*-channel strained-silicon SOI MOSFET devices. In the first study, self-heating effects were reported for SOI MOS transistors within a wide range of geometries. In a 25-nm channel length device, for instance, self-heating can lead the hot spot temperature to exceed 600 K, causing the device on-current to reduce up to 17% its nominal (isothermal) value.

4 DEVELOPMENT OF THE THERMAL MODULE

As discussed in Chapter 3, in order to make an isothermal Monte Carlo device simulator capable of accounting for self-heating effects, the software demands the incorporation of additional modules, which will be responsible to deal with the temperature-related phenomena. In addition, it is desirable that the electrical and the thermal kernel of the tool are coupled self-consistently, which means that the quantities within a module change with respect to the quantities in the other, and vice versa. A simulator that presents such features is typically named as *non-isothermal* or even *electro-thermal*.

The very first version of the software expanded in this thesis was the isothermal simulator developed by W. Gross (1999) during his Ph.D. studies within the Computational Electronics Research Group at Arizona State University, under the guidance of Professor D. Vasileska. At that time, the code was used for the modeling and simulation of deeply-scaled *n*-channel MOS transistors. Aiming the study of trap-related reliability issues, such as Bias Temperature Instability and Random Telegraph Noise — which chiefly affect *p*-channel devices —, V. Camargo (2016) modified W. Gross's code in order to make it capable of simulating *p*-channel MOS transistors. This latter version of the isothermal simulator is, thus, the one that is being extended in this work. As in its previous versions, the code was developed in Fortran language.

The construction of the modules that allows one to turn an electrical simulator into an electro-thermal simulator is the subject of the present chapter, which is organized as follows. Section 4.1 addresses the impact of the temperature on the scattering rates, as well as the generation of the temperature-dependent scattering tables. Next, Section 4.2 details the implementation of the numerical solver for solving the phonon energy balance equations at the steady state approximation. Finally, Section 4.3 focus on the silicon thermal conductivity dependence on the temperature, whereas Section 4.4 addresses its dependence on the doping and thickness of the silicon layer.

4.1 Temperature Dependent Scattering Tables

As presented in Section 2.1, the most relevant scattering mechanisms due to transitions for holes in silicon are the acoustic and non-polar optical phonon scattering. The

scattering probability per unit of time of these mechanisms are given by

$$P_{\mathrm{h,ac}}(E) = \frac{E_{\mathrm{ac}}^2 k_{\mathrm{B}} T_{\mathrm{A}}}{(2\pi)^2 \hbar^4 \rho_{\mathrm{m}} \upsilon_{\mathrm{s}}} \left(\frac{2m_0}{|A|\hbar^2}\right)^{3/2} \frac{\sqrt{E'}}{1-\beta(E')} \left(1 + \frac{2a_2 E' + a_1}{1-\beta(E')}\right) I(\theta,\phi) \tag{4.1}$$

and

$$P_{\mathrm{h,op}}(E) = \frac{D_{\mathrm{op}}^2 (2m_0)^{3/2}}{2(2\pi)^2 \rho_{\mathrm{m}} \omega_{\mathrm{op}} \hbar^3 A^{3/2}} \begin{bmatrix} N_{\mathrm{OP}} \\ N_{\mathrm{OP}}+1 \end{bmatrix} \frac{\sqrt{E'}}{[1-\beta(E')]^{3/2}} \left(1 + \frac{2a_2 E' + a_1}{1-\beta(E')}\right) I(\theta,\phi), \tag{4.2}$$

respectively (CAMARGO, 2016), where

$$I(\theta,\phi) = \int_0^{2\pi} \int_0^{\pi} \frac{sin(\theta)}{(1 \pm g(\theta,\phi))^{3/2}} d\theta d\phi. \tag{4.3}$$

Here, the function β represents the non-parabolicity of the bands, whereas g accounts for the warping of the bands. N_{OP} stands for the equilibrium optical phonon number and it is given by Bose-Einstein distribution as

$$N_{\mathrm{OP}} = \left[\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}} T_{\mathrm{OP}}}\right) - 1\right]^{-1}, \tag{4.4}$$

where ω is the phonon frequency and T_{OP} is the optical phonon mode temperature. E' is the final carrier energy and it is given by

$$E' = E - \hbar\omega - \Delta E_{\mathrm{ij}}, \tag{4.5}$$

where ΔE_{ij} represents energy gap between the initial band i and the final band j. The description of the additional physical parameters and constants utilized for scattering rates calculation is presented in Table 4.1.

From the set of equations presented earlier, it is clear that the scattering rates for both acoustic and optical modes depend on the temperature. Acoustic phonon rate depends linearly with the temperature T (see Equation 4.1), whereas non-polar optical phonon rate depends on the equilibrium phonon number N_{OP} which, in turn, depends on the temperature T_{OP}, as in Equation 4.4. The impact of the temperature on the phonon scattering rates is depicted in Figure 4.1 and 4.2.

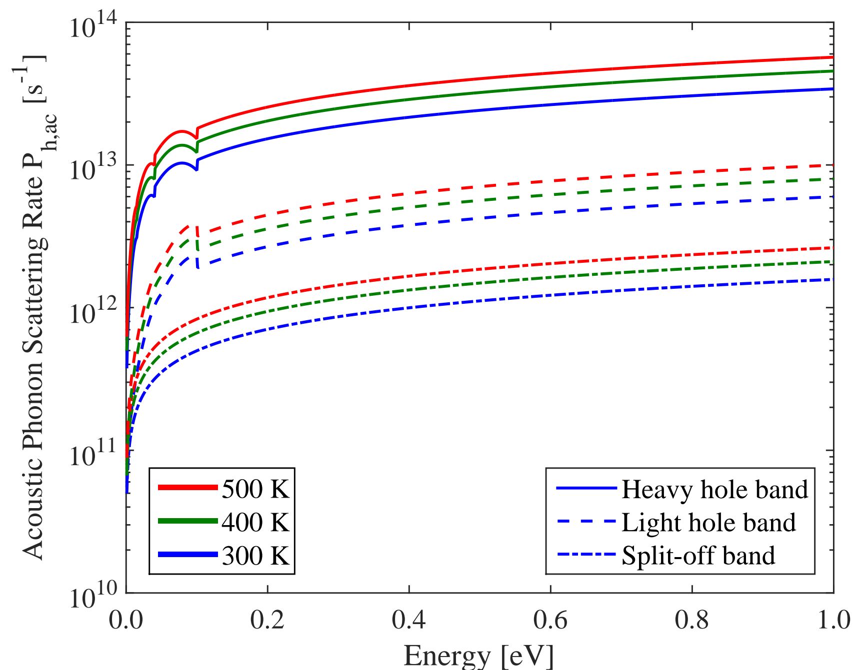
In order to evaluate the proper scattering rate for a given lattice or optical phonon bath temperature, the scattering rates for both modes are calculated for a wide range of temperatures. In this work, the temperature for the scattering rates calculation varies from

Table 4.1: Description of the physical parameters and constants utilized for scattering rates calculation.

Parameter	Description
E_{ac}	Acoustic phonon deformation potential
D_{op}	Optical phonon deformation potential
k_{B}	Boltzmann constant
T	Temperature
m_0	Electron rest mass
ρ_{m}	Mass density
υ_{s}	Sound velocity
A	Inverse valence-band parameter
a_1, a_2	Energy dependent parameters
θ	Carrier wave vector polar angle
ϕ	Carrier wave vector azimuthal angle
$\hbar$	Reduced Planck constant
ω_{op}	Optical phonon angular frequency

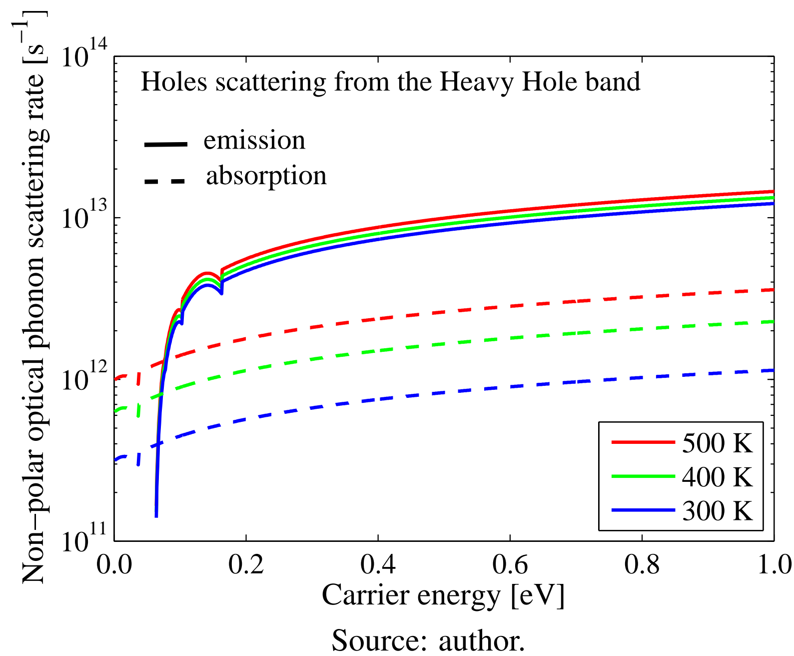
300 K to 500 K, in discrete steps of 5 K. These rates are then tabulated in temperature-dependent scattering tables, as already presented in Section 3.3.

Figure 4.1: Acoustic phonon scattering rates for holes scattering from the heavy hole, light hole, and split-off bands for temperatures of 300 K, 400 K, and 500 K.



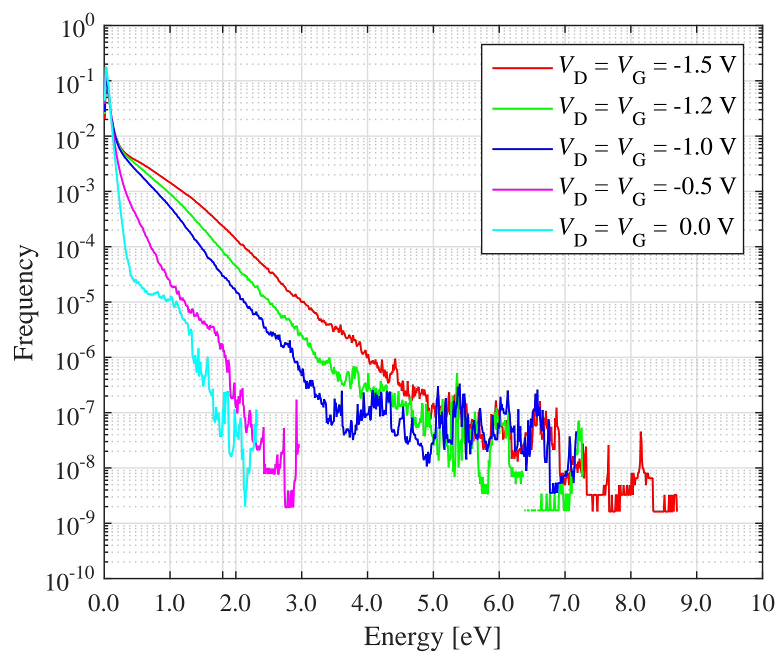
Source: author.

Figure 4.2: Non-polar optical phonon scattering rate for holes scattering from the heavy hole band for temperatures of 300 K, 400 K, and 500 K.



Source: author.

Figure 4.3: Histogram of the energy of the carriers during a simulation as a function of the applied bias for a bulk MOSFET. The cumulative probability of the carrier energy to exceed the ionization threshold energy is — according to the bias —, as much as $2.9 \times 10^{-4}\%$ for $V_D = V_G = 0.0$ V, $1.5 \times 10^{-3}\%$ for $V_D = V_G = -0.5$ V, 0.09% for $V_D = V_G = -1.0$ V, 0.27% for $V_D = V_G = -1.2$ V, and 0.89% for $V_D = V_G = -1.5$ V.



Source: author.

Besides the scattering due to transitions, at least two additional scattering mechanisms are important for charge transport in silicon: Coulomb scattering and impact ionization (JACOBONI; REGGIANI, 1983). In the present simulator, however, Coulomb scattering is performed directly in the real-space, via carrier-carrier and carrier-impurity interactions modeling, and thus, no thermal compensation is demanded. Scattering due to impact ionization, in turn, is not taken into account in the code. For ultra-small devices, the typical bias does not exceed 1.5 V and, thus, the carrier energy rarely overcomes the ionization threshold energy (see Figure 4.3), which for holes is on the order of 1.8 eV (ANDERSON; CROWELL, 1972).

4.2 Solving the Phonon Energy Balance Equations in 3D Coordinates

In order to obtain the acoustic and the optical phonon temperatures within the device, the phonon energy balance equations have to be solved. As presented in Section 3.3, the phonon energy balance equations for the optical and the acoustic phonon baths have the form

$$C_{\mathrm{OP}}\frac{\partial T_{\mathrm{OP}}}{\partial t} = \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}} - T_{\mathrm{OP}}}{\tau_{\mathrm{c-OP}}}\right) + \frac{\rho m^* \upsilon_{\mathrm{D}}^2}{2\tau_{\mathrm{c-OP}}} - C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}} - T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right) \tag{4.6}$$

and

$$C_{\mathrm{A}}\frac{\partial T_{\mathrm{A}}}{\partial t} = \nabla\cdot(\kappa_{\mathrm{A}}\nabla T_{\mathrm{A}}) + C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}} - T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right) + \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}} - T_{\mathrm{A}}}{\tau_{\mathrm{c-A}}}\right). \tag{4.7}$$

respectively. Since steady state regime is considered here, the dependence on time vanishes and the LHS of Equation 4.6 and Equation 4.7 can be expressed as

$$C_{\mathrm{OP}}\frac{\partial T_{\mathrm{OP}}}{\partial t} = 0 \tag{4.8}$$

and

$$C_{\mathrm{A}}\frac{\partial T_{\mathrm{A}}}{\partial t} = 0. \tag{4.9}$$

Thus, Equation 4.6 can be rewritten as

$$C_{\mathrm{OP}}\left(\frac{T_{\mathrm{OP}} - T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}}\right) = \frac{3}{2}\rho k_{\mathrm{B}}\left(\frac{T_{\mathrm{c}} - T_{\mathrm{OP}}}{\tau_{\mathrm{c-OP}}}\right) + \frac{\rho m^* \upsilon_{\mathrm{D}}^2}{2\tau_{\mathrm{c-OP}}}. \tag{4.10}$$

Finally, substituting Equation 4.9 and Equation 4.10 into Equation 4.7 yields

$$-\nabla \cdot (\kappa_{\mathrm{A}} \nabla T_{\mathrm{A}}) = \frac{3}{2}\rho k_{\mathrm{B}} \left(\frac{T_{\mathrm{c}} - T_{\mathrm{OP}}}{\tau_{\mathrm{c-OP}}} \right) + \frac{\rho m^* \upsilon_{\mathrm{D}}^2}{2\tau_{\mathrm{c-OP}}} + \frac{3}{2}\rho k_{\mathrm{B}} \left(\frac{T_{\mathrm{c}} - T_{\mathrm{A}}}{\tau_{\mathrm{c-A}}} \right) \tag{4.11}$$

and

$$T_{\mathrm{OP}} = \left[\frac{3}{2}\rho k_{\mathrm{B}} \frac{T_{\mathrm{c}}}{\tau_{\mathrm{c-OP}}} + \frac{\rho m^* \upsilon_{\mathrm{D}}^2}{2\tau_{\mathrm{c-OP}}} + C_{\mathrm{OP}} \frac{T_{\mathrm{A}}}{\tau_{\mathrm{OP-A}}} \right] \cdot \left[C_{\mathrm{OP}} \frac{1}{\tau_{\mathrm{OP-A}}} + \frac{3}{2}\rho k_{\mathrm{B}} \frac{1}{\tau_{\mathrm{c-OP}}} \right]^{-1}, \tag{4.12}$$

respectively. The right-hand side of Equation 4.11 is commonly named *forcing function* and determines the amount of heat that is being generated (or consumed) in the system. The left-hand side, in turn, accounts for the heat diffusion in the structure, which is modulated by the thermal conductivity κ of the material. Since in this work the acoustic phonon scattering mechanism is defined as elastic, the last term of the RHS of Equation 4.11 was properly neglected.

Nevertheless, solving analytically a second-order partial differential equation defined in the continuum domain is unpractical, since it may involve an enormous number of points to be evaluated. In this way, the continuous function has to be defined over a finite domain and represented by a finite amount of data. In other words, the equation must assume a discrete form in which its solution approximates the solution of the continuous one. This process is called *discretization* and it allows the differential equation to be solved numerically.

There are several numerical recipes that can be employed to solve a discretized equation, depending upon the complexity and size of the system which is represented. For low and medium-sized systems, direct elimination techniques, such as the Gauss elimination method and lower-upper (LU) decomposition method, are preferred (VASILESKA; GOODNICK; KLIMECK, 2010). For large systems, iterative schemes such as Successive Over-Relaxation (SOR) method (YOUNG, 1954), Conjugate Gradient method (HESTENES; STIEFEL, 1952) and Multigrid method (BRANDT, 1973), are almost always used (VASILESKA; GOODNICK; KLIMECK, 2010). For the approach presented in this thesis, however, the adopted numerical method was the Stone's Strongly Implicit Procedure (SIP) (STONE, 1968), since this method proved to be very efficient computationally for 3D systems which have to be repeatedly solved (GROSS, 1999).

In order to ensure the proper operation of the numeric solver, a variety of thermal situations were tested. First, temperature diffusion across the test structure was simulated considering uniform and non-uniform material thermal conductivity. Then, simulations

were performed for the homogeneous and inhomogeneous cases as well, i.e., in the absence and in the presence of a heat source within the structure, respectively. The process of validation of the thermal solver is presented in details in Appendix A.

4.3 Thermal Conductivity of Silicon Dependence on the Temperature

In order to preserve the self-consistency of the electro-thermal simulator, another problem that has to be addressed properly is the silicon thermal conductivity dependence upon the temperature. In a semiconductor, several types of carriers contribute to the heat transport — such as phonons, photons, electron-hole pairs, and the separate electron and holes (GLASSBRENNER; SLACK, 1964) — and it is known their individual contribution to the thermal conductivity varies depending on the temperature of the material (SHANKS et al., 1963).

In silicon, even in the presence of large concentrations of free charge carriers, the thermal conduction is dominated by phonon transport (WEBER; GMELIN, 1991; ASHEGHI et al., 2002). Consequently, the mechanisms which affect the thermal conductivity are mostly associated with the scattering of phonons, such as phonon-phonon scattering, phonon-boundary scattering, phonon-carrier scattering, and phonon-impurity scattering (ASHEGHI et al., 2002). Phonon-boundary scattering is especially important in thin films of silicon, whereas phonon-carrier and phonon-impurity scattering play an important role at low temperatures, as it is presented further in this section.

Phonon-phonon scattering, in turn, dominates the thermal transport at high temperatures, which in this context means 300 K and above. Basically, as the system becomes hotter, the phonon population is increased, which causes the collision frequency among phonons to be high. The increase of such collisions reduces the phonon mean free path and restricts them to move from hot to cold regions and vice versa (VASILESKA; RALEVA; GOODNICK, 2008). Since the energy transport decreases, the thermal conductivity also decreases.

First attempts to characterize the behavior of the silicon thermal conductivity according to the temperature were done in the 50's, after the experimental measurements of Rosenberg (1954), White and Woods (1956), Carruthers et al. (1957) and others. In the same decade, Callaway (1959) developed a model which was able to explain the thermal conductivity dependence on the temperature for semiconductors in general. In 1963, Holland applied Callaway's model to silicon (and germanium) and expanded Callaway's

approach to the high-temperature range as well. Holland's model was then validated through experimental measurements carried out by Glassbrenner and Slack, which were published later, in 1964.

The models such those presented above are typically dedicated to cover a wide range of temperatures for a certain material and they may depend upon several other additional physical parameters besides the temperature itself. For the purpose of device simulation, however, the range of temperatures in which the device is typically subject is restricted. In this case, the thermal conductivity dependence on the temperature can be appraised to a simple power law function, which in turn, depends solely on a few and well-characterized parameters. In this context, simplified expressions for the modeling of the silicon thermal conductivity along the temperature have been extensively proposed in the literature (GAUR; NAVON, 1976; ADLER, 1978; CHRYSSAFIS; LOVE, 1979; LETURCQ et al., 1987; PALANKOVSKI; SCHULTHEIS; SELBERHERR, 2001). In this particular work, however, major attention is given to the models developed by Selberherr (1984) and Lacroix, Joulain and Lemonnier (2005).

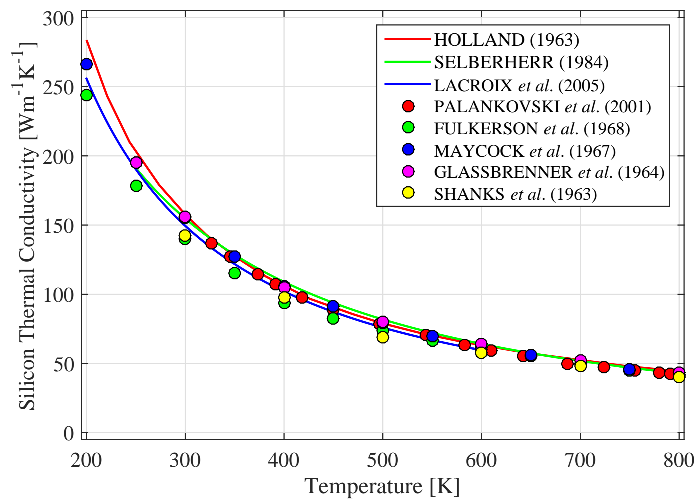
Selberherr's approach is derived from the study of Glassbrenner and Slack (1964) and its validity is limited for temperatures within the range from 250 K to 1000 K. Lacroix, Joulain and Lemonnier's approach, in turn, is derived from Monte Carlo calculations and theoretical values, and it is valid for temperatures within the range from 200 K to 600 K. The expressions for the silicon thermal conductivity κ_{Si} dependence on the temperature T are presented in Equation 4.13 and 4.14 for both Selberherr's and Lacroix's models, respectively. Regarding the units of measurement, the temperature is typically given in K while the thermal conductivity is given in W/mK.

$$\kappa_{\mathrm{Si}}(T) = 154.86 \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-4/3} \tag{4.13}$$

$$\kappa_{\mathrm{Si}}(T) = \frac{\exp(12.570)}{T^{1.326}} \tag{4.14}$$

A compilation of values for the silicon thermal conductivity along the temperature is shown in Figure 4.4, considering a temperature range from 200 K to 800 K. Solid lines represent model predictions (HOLLAND, 1963; SELBERHERR, 1984; LACROIX; JOULAIN; LEMONNIER, 2005), whereas the circles represent experimental data (SHANKS et al., 1963; GLASSBRENNER; SLACK, 1964; MAYCOCK, 1967; FULKERSON et al., 1968; PALANKOVSKI; SCHULTHEIS; SELBERHERR, 2001). Comparatively,

Figure 4.4: Silicon thermal conductivity as a function of the temperature. Solid *lines* represent model predictions and *circles* represent experimental data.



Source: author.

the results for all the three aforementioned models are agreement with experimental data. However, if one sets a temperature range from 300 K to 500 K — which is the interval adopted here —, Lacroix, Joulain and Lemonnier's formula is the one which results in the best fit for this particular region of interest. Therefore, the modeling performed in the tool presented here utilizes the relationship given by Equation 4.14.

4.4 Thermal Conductivity Dependence on the Silicon Doping and Thickness

Besides its dependence on the temperature, the thermal conductivity of silicon is also a function of the doping and thickness of the silicon film. At low temperatures, where the phonon-carrier and phonon-impurity scattering becomes important, the thermal conductivity of heavily doped materials is strongly reduced, as it is shown in the work of Asheghi et al. (2002). At room temperature and above, however, the thermal transport is dominated by phonon-phonon scattering and the thermal conductivity is virtually independent of the impurity concentration (WEBER; GMELIN, 1991).

The thickness of the film also plays a major role in silicon thermal conductivity. For thin layers — with the thickness in the order of the phonon mean free path[1] or less —, the phonon-boundary scattering poses as the dominant scattering mechanism, and it strongly reduces the thermal conductivity (ASHEGHI et al., 2002; JU, 2005). In addition,

[1] In silicon, the phonon mean free path is around 300 nm (ASHEGHI et al., 1998).

the thermal conductivity is still dependent on the temperature even for these very thin films, and taking this characteristic into account is particularly important for the study and modeling of heat transport in SOI and FinFET devices.

In this context, Vasileska, Raleva and Goodnick developed a theoretical model that accounts for the temperature and thickness dependence of the silicon thermal conductivity. They reported silicon thermal conductivity values for several temperatures and for a wide range of thicknesses, and their results are in good agreement with experimental data, as presented in Vasileska, Raleva and Goodnick (2010a).

For bulk technologies, however, the silicon layers are typically much thicker and, consequently, the phenomenon is not observed (ASHEGHI et al., 1997).

5 ELECTRO-THERMAL SIMULATION PROCEDURE

This chapter is dedicated to cover the electro-thermal simulation procedure. It starts addressing the simulation flow in Section 5.1, which also presents a complete flowchart of the tool. Next, the definition of the thermal boundary conditions is detailed in Section 5.2. Section 5.3, in turn, presents the structure and characteristics of both bulk and FD-SOI case study MOSFETs used in the scope of the present work. Section 5.4 addresses the optimization of some critical simulation parameters, such as the time step and mesh size. Finally, Section 5.5 presents the I–V curves for the case study transistors introduced in Section 5.3.

5.1 Simulation Flow

Figure 5.1 depicts the flowchart of the electro-thermal simulator, highlighting the electric, thermal, and mixed building blocks. Regarding the thermal part, at the beginning of the simulation, the scattering rates are tabulated in scattering tables for a wide range of acoustic and optical phonon temperatures. Next, both acoustic and optical phonon temperatures are initialized with a uniform initial temperature T_{ini} that, for convenience, is set to be equal to the heat sink temperature T_{sink}. This temperature is used in the following temperature-dependent free flight routine.

Resolving the temperature for each iteration i is very time consuming and not necessary. In this way, another loop (over j) is defined for this purpose in the simulation. This loop is called *temperature Gummel cycle*. Only a few iterations j are performed in a simulation. Basically, a simulation with total time t_{max} is divided into M time windows of length t_{w}. The temperature cycle is then executed only at the end of the window, using the data sampled during the current time interval. In addition, the temperature loop only starts to be performed after the simulation reaches the steady state, i.e., when the simulation time t is larger than the transient time t_{tran}.

Upon executed, the first task within the temperature Gummel cycle is averaging the quantities of interest stemming from the sampling routine. After that, they are transferred to the numerical solver, which solves the acoustic and optical phonon balance equations. Next, the acoustic and optical phonon temperatures are updated for each grid point, as well as the temperature dependent parameters. This temperature loop is performed until the Mth time window is treated, which coincides with the end of the simulation. One

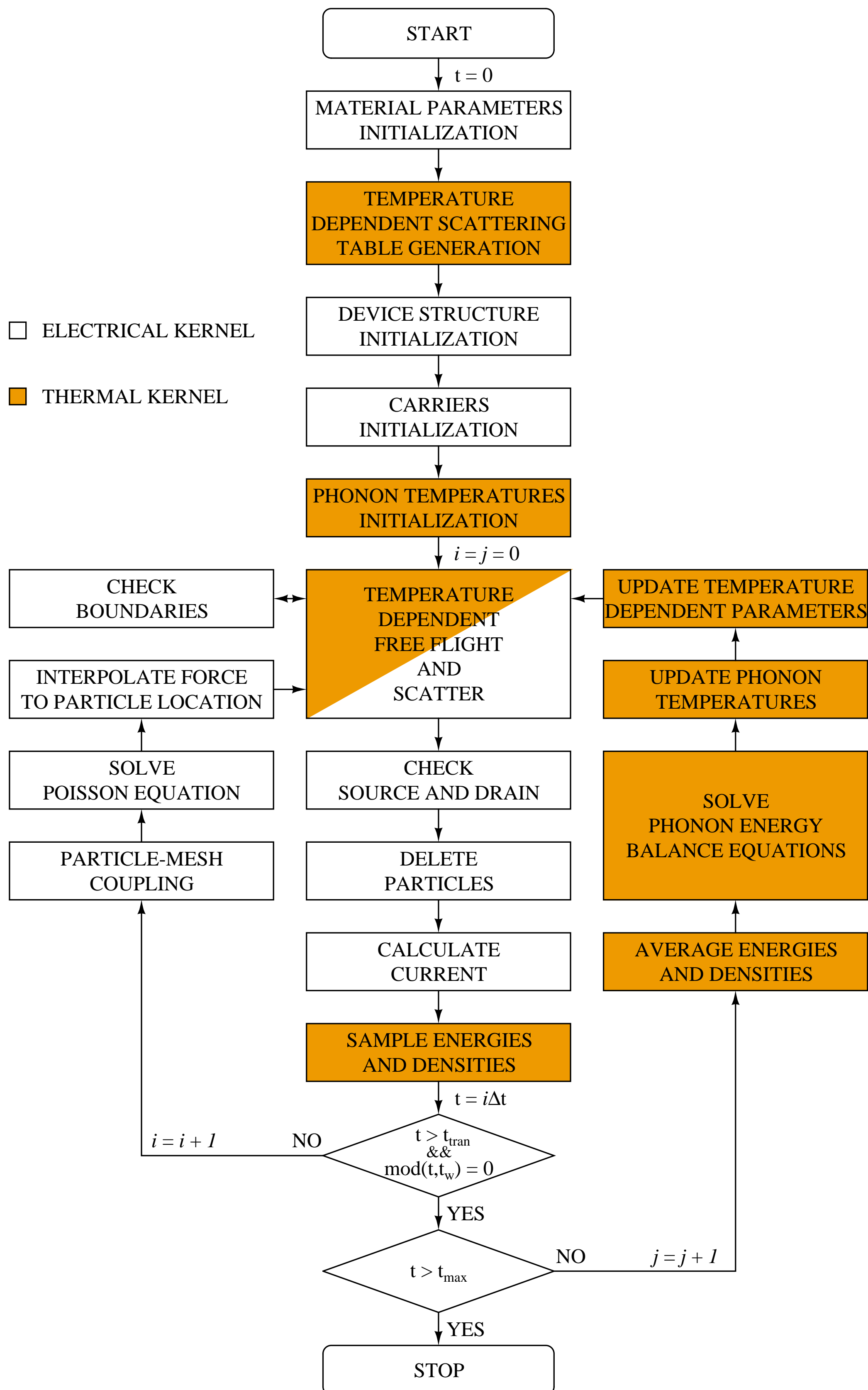
important observation, however, is that non-uniform phonon temperatures at the free flight stage are only possible after the first temperature Gummel cycle be performed (i.e., for the 2nd, 3rd,...). For this reason, the current calculated within the first time window is not subjected to lattice heating effects and, thus, it should be equal to the current evaluated for an isothermal simulation at a uniform temperature T_{sink}.

5.2 Thermal Boundary Conditions

Boundary conditions play an important role in thermal simulations since the temperature rise strongly depends upon them (LAI; MAJUMDAR, 1996). In addition, in order to obtain realistic temperature profiles, the thermal boundaries must be specified as close as possible to a typical operating condition. In this study, the thermal boundary conditions are given as follows:

- The bottom of the device is assumed to be attached to a good heat sink, which is kept at a constant temperature T_{sink}. Therefore, Dirichlet (CHENG; CHENG, 2005) — or fixed — boundary condition is applied to the bottom of the transistor. The temperature T_{sink} is assumed to be 300 K unless otherwise specified.
- On all lateral sides of the device, Neumann (CHENG; CHENG, 2005) — or adiabatic — boundary conditions are applied, i.e., it is assumed that no heat flow occurs through the laterals of the transistor. This condition can be justified if one assumes that the device under consideration does not operate alone. If neighboring devices operate with nearly the same thermal dissipation, an adiabatic boundary represent a good approximation of the real picture (RAMAN; WALKER; FISHER, 2003).
- On the top surface, the layers of oxide limit the temperature diffusion, being the adiabatic boundary a good approximation of this condition (RAMAN; WALKER; FISHER, 2003). Specifically in the simulation of SOI devices, however, the gate electrode is assumed to play a major role on the transistor thermal performance (RALEVA et al., 2008). In addition, if no heat is considered to be flowing through the top surface of the structure, the silicon layer temperature delivered from the simulation will be unphysically overestimated, due to the negligible heat flow through the buried oxide layer. In this way, the gate contact is treated as a fixed-temperature boundary condition, in agreement with the treatment performed in commercial tools (RALEVA et al., 2008). The heat dissipating through interconnects, in turn, is con-

Figure 5.1: Flowchart of the electro-thermal simulator.



Source: author.

sidered to be small (LEE; PALISOC; MIN, 1989) and, thus, it is neglected. In this way, both source and drain metal contacts are left floating.

In a real operating chip, however, the lattice temperature is expected to be higher than a single device lattice temperature, due to the power dissipation of all active devices (LAI; MAJUMDAR, 1996). This situation can be simulated by increasing the substrate boundary temperature T_{sink}. Hence, the efficacy of the gate as a heat sink and its impact on the device performance can be extracted by adopting several different gate temperatures in the simulations. The adopted boundary conditions are summarized in Figure 5.2, according to the transistor structure.

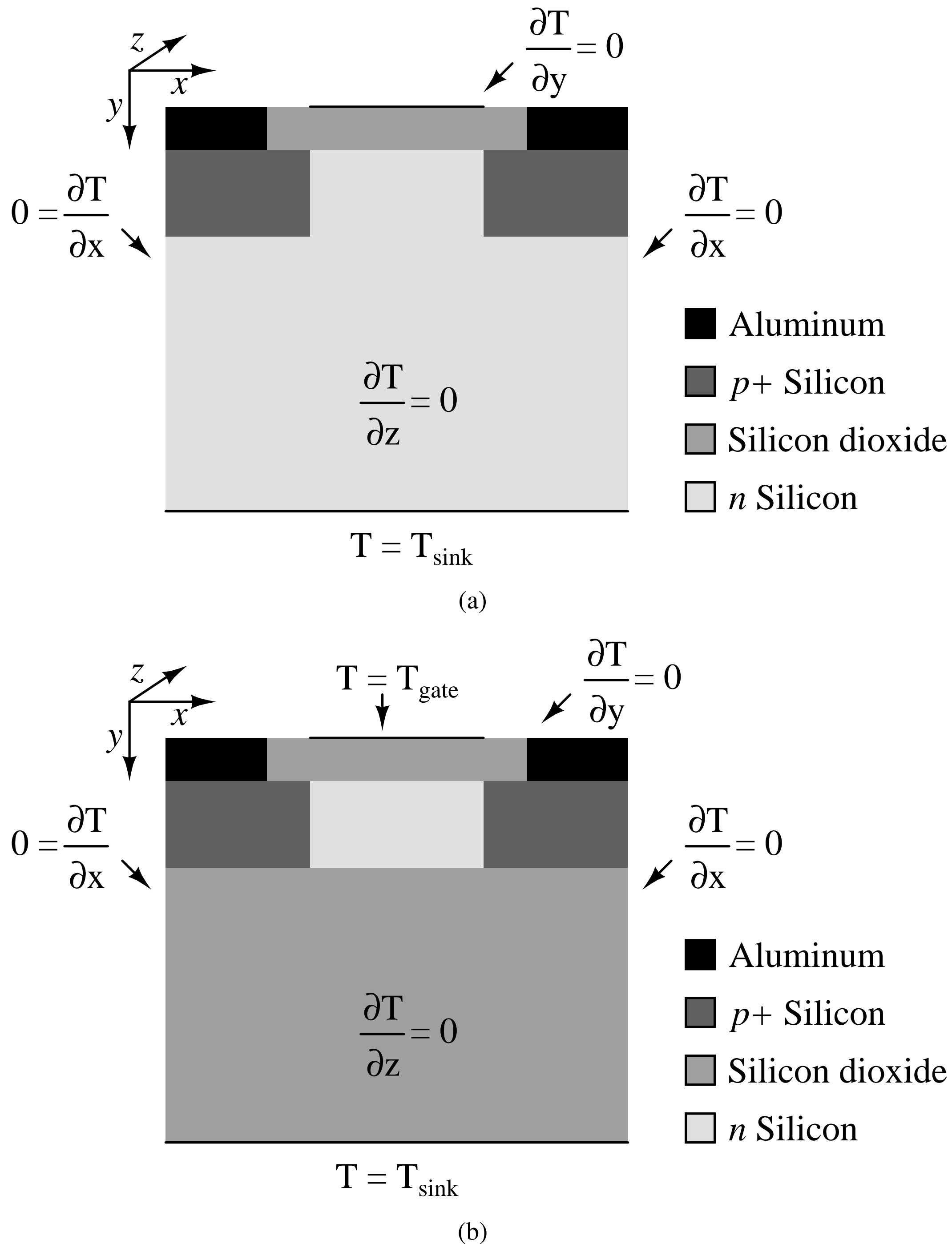
5.3 Case Study Devices

In order to standardize the simulations to be performed in this work, case study devices were specified for both bulk and FD-SOI technology. The structure of the simulated *p*-channel MOS transistors is depicted in Figure 5.3, which also highlights the different regions that compose the devices. The physical dimensions shown in the figure are summarized in Table 5.1.

Figure 5.3 also presents the most relevant points in the coordinate system along the length−depth plane (*xy*-axes), which will be used in Chapter 6 for presenting some of the results. Here, x_{S} and x_{D} represent the position along the length of the device where the metallurgical junctions between the substrate and the source and drain regions are formed, respectively. Likewise, the position where the junction between these regions and the substrate occurs along the depth is defined by y_{SD}. Although omitted in Figure 5.3, the width of the device is represented by W, which in the xyz coordinate system extends itself from 0 to z_{max}.

Besides the geometry of the device and the characteristic physical constants of the materials that constitute the transistor, specific simulation parameters for the device at hand must be defined. The most important ones are summarized in Table 5.2. For the tool presented here, a typical simulation of a *p*-channel MOS transistor requires between 2 and 3 ps to reach the steady state. Thus, the transient time t_{tran} is chosen to be 5 ps. During this part of the simulation, all sampled data are not used to compute the outputs. By utilizing a time step of 0.1 fs per iteration, a simulation of 55 ps requires 550,000 iterations to be completed. The choice of such a time step is detailed further in this chapter.

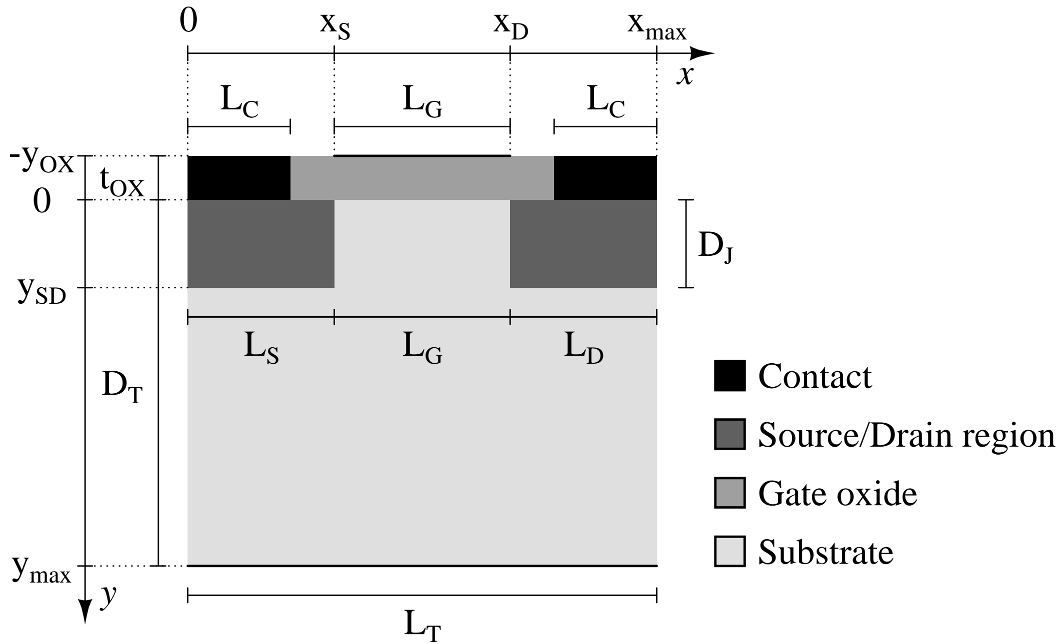
Figure 5.2: Thermal boundary conditions for bulk (a) and FD-SOI (b) transistor structure. For the sake of simplicity, the extension of the devices along the *z*-axis is omitted in the figures.



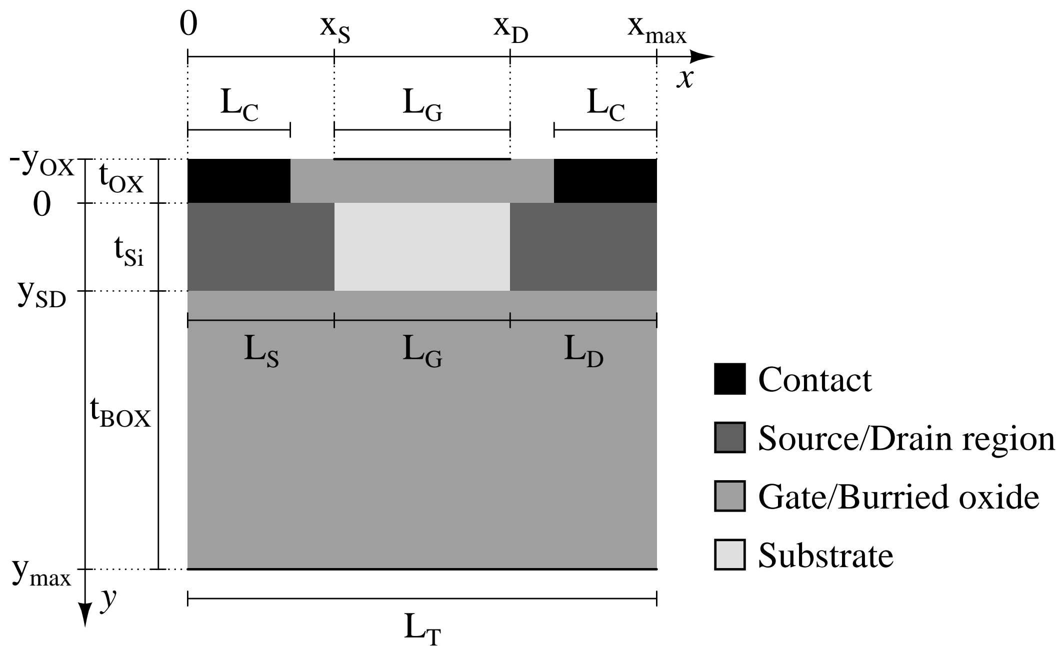
Source: author.

The computer time needed for completing a simulation is strongly dependent on the number of mesh points of the system to be evaluated. In this regard, the number of mesh points in the bulk of the transistor was defined according to its physical dimensions in such a way a uniform mesh spacing Δ of 4.0 nm was obtained for all directions. This

Figure 5.3: Bulk (a) and FD-SOI (b) case study device structure: transistor dimensions, regions, and coordinates in the length-depth (*xy*-) plane. For the sake of simplicity, the extension of the devices along the *z*-axis is omitted in the figures.



(a)



(b)

Source: author.

distance was chosen aiming to minimize self-forces in the mesh, and it also provides suitable smoothness and resolution for the quantities defined in the discrete domain. For the oxide region, however, the thickness t_{OX} was divided uniformly in 5 and 8 mesh points along the *y*-axis for bulk and FD-SOI, respectively, since in this region a more detailed

Table 5.1: Physical dimensions of the bulk and FD-SOI case study MOSFETs.

Parameter		Bulk	FD-SOI	Unit
Source region length	L_S	24.0	24.0	nm
Drain region length	L_D	24.0	24.0	nm
Gate length	L_G	24.0	24.0	nm
Contact length	L_C	20.0	20.0	nm
Total device length	L_T	72.0	72.0	nm
Junction depth	D_J	20.0	–	nm
Total silicon depth	D_T	100	–	nm
Silicon layer thickness	t_{Si}	–	12.0	nm
Buried oxide thickness	t_{BOX}	–	52.0	nm
Device width	W	100	100	nm
Gate oxide thickness	t_{OX}	1.2	2.0	nm

profile of the vertical field is desirable.

The simulation computational cost also grows with the number of particles in the system. In this way, net impurity concentrations larger than 10^{19} cm^{-3} for source and drain regions are prohibitive if one aims reasonable simulation times[1]. Since the devices have poor channel engineering (e.g., no halo implants), a substrate doping N_D of as much as 5.0×10^{18} cm^{-3} was used to prevent excessive leakage currents and short channel effects.

One may argue, however, that $N_D = 5.0 \times 10^{18}$ cm^{-3} is a reasonable substrate concentration for a bulk device, but it is a excessively high doping for an FD-SOI device, which may cause it to operate partially depleted. In practice, an SOI device will operate fully depleted when the depletion depth t_{dep} overcomes the physical silicon thickness t_{Si}, so a neutral region no longer exists between the source and drain (EBINA et al., 2000). The depth of the depletion region can be related to the substrate doping N_{sub} as (PELLOIE, 1997)

$$t_{dep} = \sqrt{\frac{4\varepsilon_{Si}\phi_f}{qN_{sub}}}, \tag{5.1}$$

where ε_{Si} is the silicon permittivity, and ϕ_f is the Fermi potential and it is given by

$$\phi_f = \frac{1}{q}k_B T \ln\left(\frac{N_{sub}}{N_i}\right). \tag{5.2}$$

[1]Considering the transistor width W is in the order or larger than 100 nm.

Table 5.2: Main simulation parameters for bulk and FD-SOI case study MOSFETs.

Parameter		Bulk	FD-SOI	Unit
Transient time	t_{tran}	5.0	5.0	ps
Simulation time	t_{max}	55	55	ps
Time step	Δt	0.1	0.1	fs
Monte Carlo iterations	N	5.5×10^5	5.5×10^5	–
Thermal Gummel Cycles	M	10	10	–
Mesh points along x-axis	i_{max}	18	18	–
Mesh points along y-axis	j_{max}	25	16	–
Mesh points along z-axis	k_{max}	25	25	–
Mesh points in the gate oxide	j_{OX}	5	8	–
Maximum number of carriers	N_{car}	5.0×10^4	5.0×10^4	–
Mesh spacing	Δ	4.0	4.0	nm
Source/Drain doping	N_{A}	1.5×10^{19}	1.5×10^{19}	cm^{-3}
Substrate doping	N_{D}	5.0×10^{18}	5.0×10^{18}	cm^{-3}
Intrinsic doping	N_{i}	1.5×10^{10}	1.0×10^{10}	cm^{-3}
Initial temperature	T_{ini}	300	300	K
Heat sink temperature	T_{sink}	300	300	K

For a substrate doping of 5.0×10^{18} cm^{-3}, t_{dep} is approximately 16 nm, so a silicon layer with thickness $t_{\text{Si}} = 12$ nm was adopted. Additionally, preliminary SOI device simulations showed that the concentration of majority carriers is larger than the concentration of minority carriers for all over the extension of the device, characterizing the fully depleted operation.

Finally, the buried oxide thickness t_{BOX} was chosen with regard on the case study FD-SOI device proposed by Raleva et al. (2008), and the gate oxide thickness t_{OX} was extended to 2.0 nm aiming the adjustment of the threshold voltage. Hence, both the initial T_{ini} and heat sink T_{sink} temperatures were set at room temperature, i.e., 300 K.

5.4 Optimization of Simulation Parameters

This subsection addresses the optimization of some key simulation parameters, such as the thermal solver tolerance, the time step used in the simulations, and the mesh spacing for the real space discretization. These parameters directly impact the simulation output and numerical stability, so they must be properly tuned. By the end of the

subsection, the evolution of the number of particles in the simulation with time is also presented.

5.4.1 Thermal Solver Tolerance

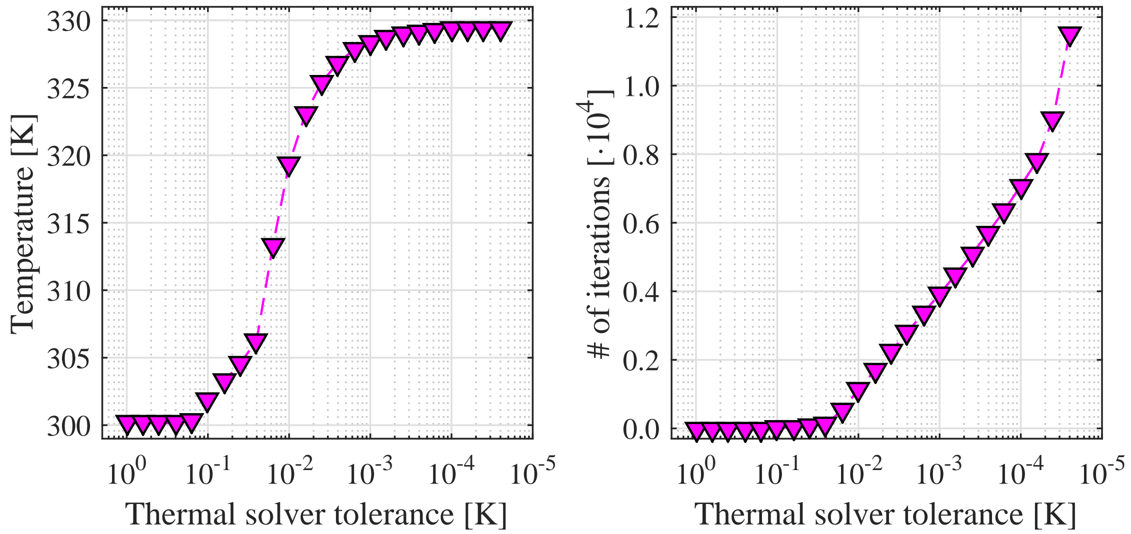
When dealing with numerical solvers, one must establish a good compromise between the computational cost required to solve the system and the desired accuracy of the solution. For this reason, the tolerance of a solver has to be chosen wisely, in order to deliver a reliable value for the solution and avoiding the excessive expense of computational resources at the same time. Attempting to obtain such an optimal tolerance for the thermal solver, the Equation 4.11 was solved for an arbitrary forcing function and for different values of the tolerance. The maximum temperature of the system and the number of iterations need to achieve the solution were extracted as a function of the solver tolerance. The results are depicted in Figure 5.4.

The solution of the system is strongly dependent on the solver tolerance for values in the range of 10^{-1} K to 10^{-3} K. As the tolerance approaches to 10^{-4} K, such dependence is significantly reduced, and for values beyond that point, the dependence of the solution on the solver tolerance is negligible. Similarly, the number of iterations required to evaluate the solution of the system also depends upon the tolerance, as expected. The system is solved very quickly for tolerances up to 5.0×10^{-1} K. For lower values, the number of iterations starts to increase logarithmically as the tolerance decreases, up to around 1.6×10^{-4} K. If the tolerance is even more reduced, the number of iterations continues to increase in a logarithmic fashion, but with a more aggressive slope. Consequently, there is no practical reason to use a thermal solver tolerance lower than 10^{-4} K, since the gain on the accuracy of the solution is negligible, and the number of iterations needed to resolve the system is largely increased.

5.4.2 Time Step

The magnitude of the time step plays a major role for Monte Carlo simulations. Ideally, it should be as small as possible, but in practice, it just needs to be small enough to prevent oscillations that could arise from the particles traveling too far without updating the field (GROSS, 1999). In other words, if the time step is too large, the carriers are

Figure 5.4: Hot spot temperature (*left panel*) and number of iterations needed to solve the heat equation (*right panel*) as a function of the solver tolerance.



Source: author.

driven by erroneous values of the electric field by the most part of the time, causing the simulation to deliver unrealistic outputs.

For the sake of numerical stability, the time step is generally correlated to the mesh size, which is, in turn, associated with the Debye length (VASILESKA et al., 2008). In addition, the time step must be much smaller than twice the inverse plasma frequency ω_{p} of the system (HOCKNEY; EASTWOOD, 1988, chapter 7), i.e.,

$$\Delta t \ll \frac{2}{\omega_{\mathrm{p}}}, \tag{5.3}$$

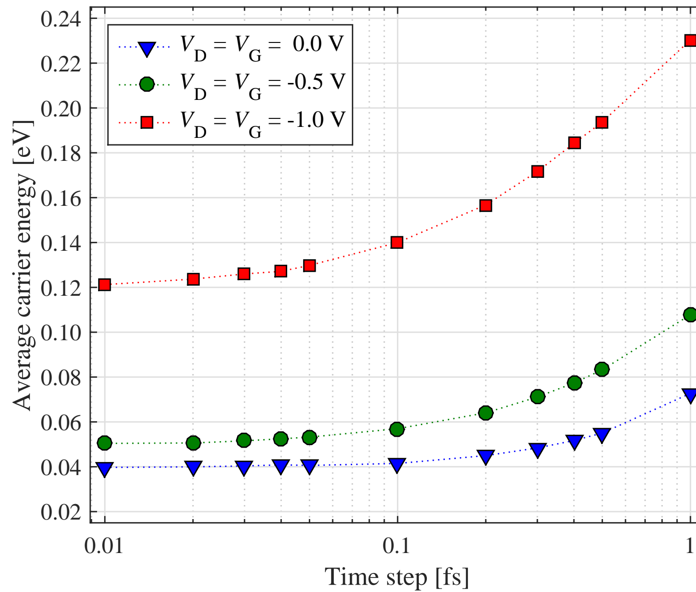
where the plasma frequency can be defined as (FISCHETTI; LAUX, 1988)

$$\omega_{\mathrm{p}} = \sqrt{\frac{q^2 \rho}{\varepsilon_{\mathrm{Si}} m^*}}. \tag{5.4}$$

Here, q is the elementary charge, ρ is the carrier concentration, $\varepsilon_{\mathrm{Si}}$ is the silicon permittivity, and m^* is the carrier effective mass. In determining the appropriated time step of a simulation, ρ must be the highest carrier concentration found in the simulated device. Similarly, for multi-valley or multi-band semiconductors, the value of m^* to be used corresponds to the smallest effective mass encountered by the carriers (VASILESKA et al., 2008).

The main drawback of very small time steps is that the simulation becomes more computer time consuming, i.e., a larger number of iterations N is needed to reach a given

Figure 5.5: Average carrier energy as a function of the time step for different biases. Error bars are not shown since they are smaller than the marker size.



Source: author.

simulation time $t_{\max}$, since $t_{\max} = N\Delta t$. For a device with peak carrier concentration on the order of 10^{19} cm^{-3}, the typical time step should be on the order of femtoseconds or less (GROSS, 1999). In this case, ω_{p} can be as high as 130 THz[2] and, thus, the time step should be much smaller than 15 fs. In the literature, reliable simulation results are presented for Δt = 0.1 fs (GROSS; VASILESKA; FERRY, 2000) and Δt = 0.2 fs (FISCHETTI; LAUX, 1988) for the same doping level.

In order to determine a proper value for the time step to be used in the simulations, a set of 10 bulk case study devices were simulated at three different biases with the simulation time step varying from $\Delta t = 0.01$ fs to $\Delta t = 1$ fs. The output quantity monitored in these simulations was the average carrier energy since it impacts directly the scattering probability of the carriers and the magnitude of the heat generation within the device. As presented in Figure 5.5, the average carrier energy changes with respect to the time step and bias. By adopting a time step of 0.01 fs, the accuracy of the outputs is maximized for all biases, but it would result in extremely time-consuming simulations. In this way, aiming to keep a good compromise among numerical stability, accuracy and simulation time, a time step of as much as 0.1 fs was chosen.

[2]For *p*-type semiconductor and considering the light hole effective mass, i.e., $m^* = 0.16m_0$.

5.4.3 Mesh Size

Another parameter related to the simulation stability and the accuracy of the outputs is the mesh spacing used for the discretization of the simulation domain. Monte Carlo simulators typically obey a general rule that establishes the largest mesh spacing Δ_{max} within the simulation domain should not exceed the Debye length λ_{D} (VASILESKA et al., 2008), which is given by (ZEGHBROECK, 2011)

$$\lambda_{\mathrm{D}} = \sqrt{\frac{\varepsilon_{\mathrm{Si}} k_{\mathrm{B}} T}{q^2 N}}, \tag{5.5}$$

where N can be either N_{D} for n-type or N_{A} for p-type devices. For the sake of numerical stability, the Debye length can be related to the plasma frequency ω_{p} as (RAMBO; DENAVIT, 1993)

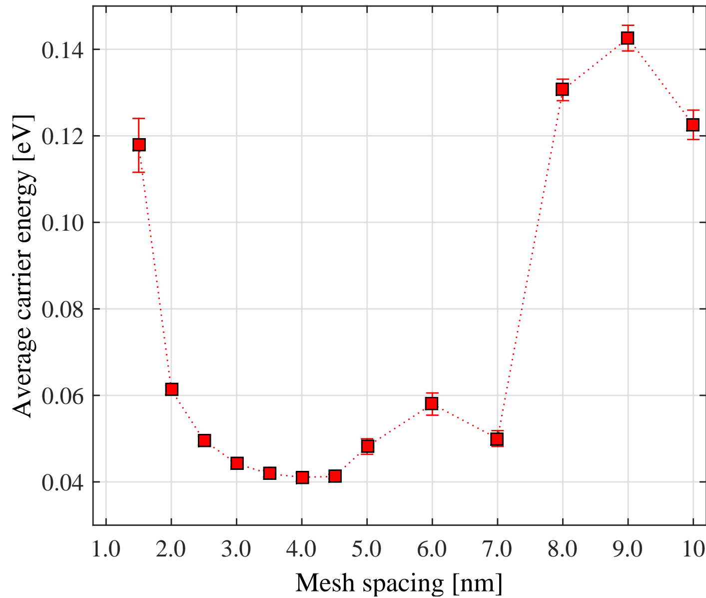
$$\lambda_{\mathrm{D}} = \frac{1}{\omega_{\mathrm{p}}} \sqrt{\frac{k_{\mathrm{B}} T}{m^*}}. \tag{5.6}$$

For those simulation tools in which the charge assignment is performed via the P3M algorithm — which is the case here —, the aforementioned criteria are not a restriction (WORDELMAN; RAVAIOLI, 2000): the Debye length is still used in the simulation, but as a normalization (scaling) factor only.

The mesh spacing used in the simulations was chosen aiming the minimization of the *self-force*. In the context of device simulation, self-force refers to an artificial force a charge may exert upon itself when it is represented in a discretized domain (FISCHETTI; LAUX, 1988). This artificial force may lead to artificial energy in the system, and it is known to increase as the mesh size decreases (HOCKNEY; EASTWOOD, 1988).

In this way, a set of 10 bulk case study devices were simulated for several values of mesh spacing, from $\Delta = 1.5$ nm to $\Delta = 10$ nm, considering $\Delta t = 0.1$ fs and $V_{\mathrm{D}} = V_{\mathrm{G}} = 0.0$ V. The average carrier energy as a function of the mesh spacing is presented in Figure 5.6. Note that self-energy (which arises from self-force) is minimized by adopting a mesh spacing between 3.5 nm to 4.5 nm. For lower values, the artificial energy largely increases as the mesh spacing decreases. For larger mesh sizes, in turn, the self-energy also increases, but with an indefinite behavior, indicating that the stability of the simulation is possibly compromised. In this way, a mesh size $\Delta = 4.0$ nm was adopted. Although the average carrier energy dependence either on the mesh size or the time step are aside from the main focus of this thesis, it will be subject of future studies.

Figure 5.6: Average carrier energy as a function of the mesh spacing for $V_{\text{D}} = 0.0$ V and $V_{\text{G}} = 0.0$ V.



Source: author.

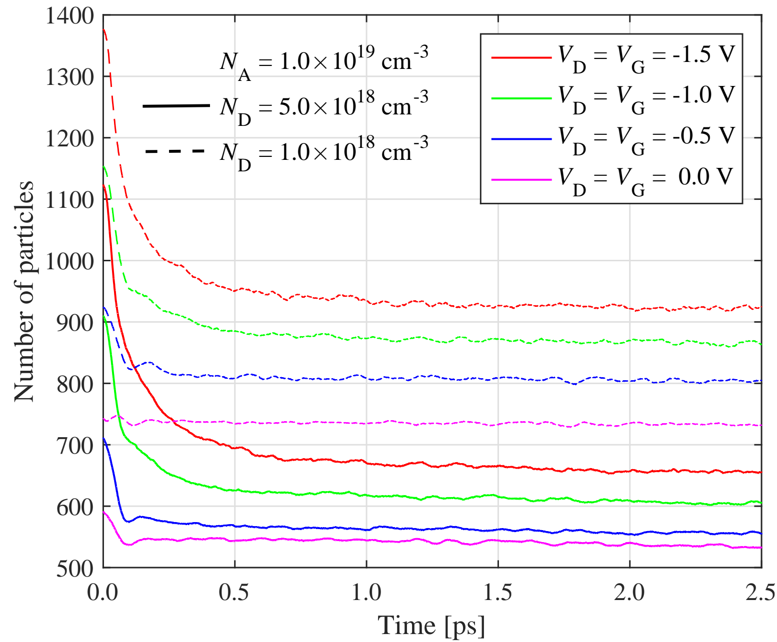
5.4.4 Number of Particles in the Simulation

In the scope of the approach presented in this thesis, the number of particles (i.e., carriers) in the simulation cannot be specified to be constant. Basically, this parameter is determined by the physics itself, and so, the tool uses as many carriers as it needs. However, a maximum number of carrier N_{car} is defined, in order to avoid unnecessary computer time expense in those routines which go over the entire stack of particles (used and not used) of the system.

At the beginning of the simulation, the number of particles in the system is determined according to the doping, i.e., each impurity atom (acceptor for holes and donor for electrons) is assumed to "generate" a free carrier in average. This process, however, might create extra net charge within the device. As the simulation evolves, on the other hand, the algorithm starts to delete (or inject) additional particles, aiming to reach charge neutrality. After some transient time, the number of carriers being used in the simulation converges to a nearly constant value. The evolution of the number of carriers in time is depicted in Figure 5.7.

The amount of particles in a simulation also changes with respect to the Fermi level of the semiconductor. In this way, by changing either the impurity concentration or the

Figure 5.7: Number of particles (carriers) being used in the simulation along time.



Source: author.

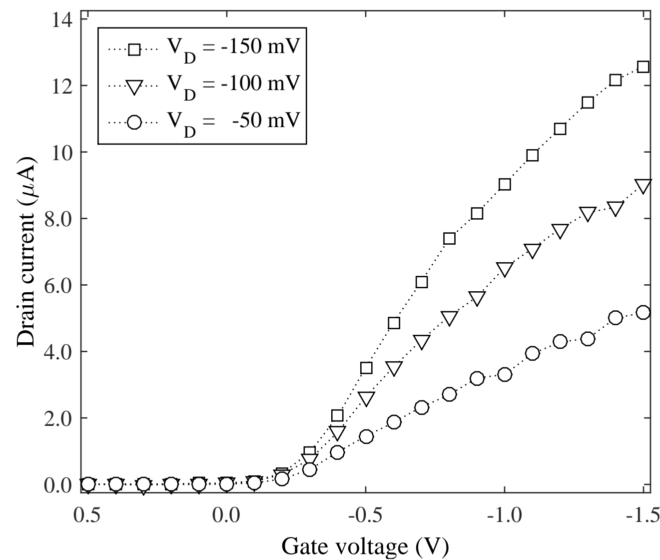
bias, the number of carriers changes accordingly. This behavior is clearly seen in Figure 5.7. In the tool presented here, only majority carriers are treated as particles. Minority carriers, which are also important within the simulation, are accounted and represented as a distributed quantity (i.e., a distributed variable and not a discrete particle), and evaluated via Fermi-Dirac statistics.

5.5 Device Characterization

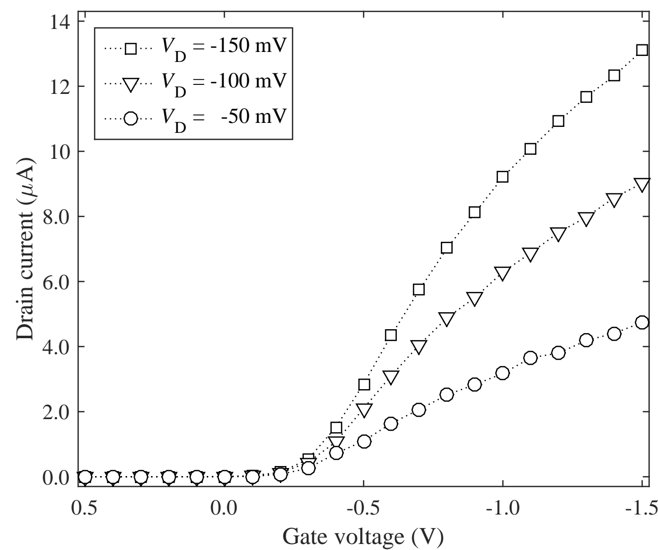
Before proceeding to the electro-thermal simulations, the I−V characteristics of the case study devices presented in Section 5.3 were extracted, in order to ensure the simulated devices were properly defined. To do that, two different sets of simulations were performed for each structure: one set dedicated to exploring the drain current dependence on the gate voltage, i.e., $I_D \times V_G$; and the other one for exploring the drain current dependence on the drain voltage, i.e., $I_D \times V_D$.

For the $I_D \times V_G$ simulations, three different drain biases were used: $V_D = -50$ mV, $V_D = -100$ mV, and $V_D = -150$ mV. The gate voltage V_G was swept from 0.5 V to -1.5 V in steps of 0.1 V, covering thus all the device operation regions, i.e., from accumulation to strong inversion. The extracted $I_D \times V_G$ characteristics of the case study

Figure 5.8: $I_D \times V_G$ characteristics of the case study transistors: bulk (a) and FD-SOI (b) MOSFETs.



(a)



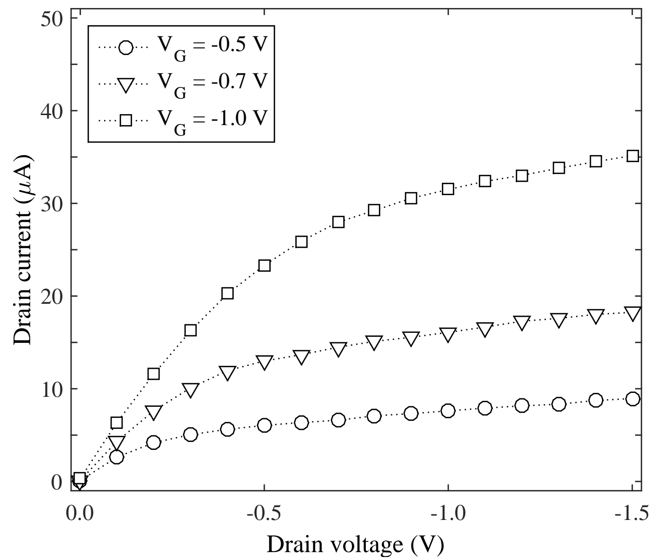
(b)

Source: author.

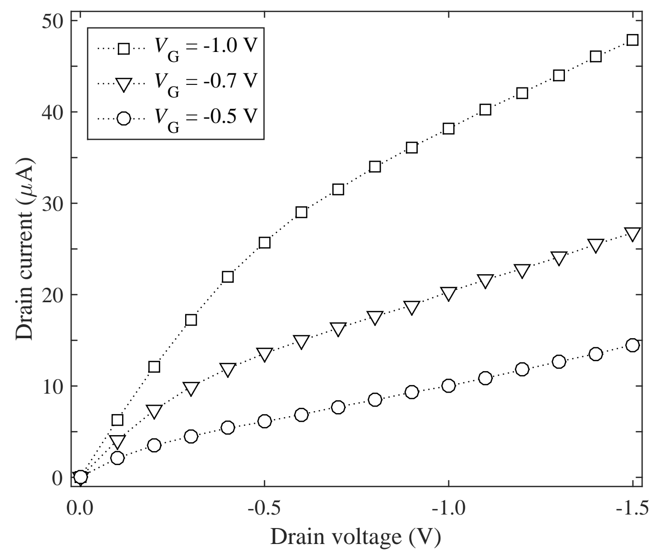
devices are depicted in Figure 5.8, where each point in the curve represents the average obtained from 10 samples. Applying the linear extrapolation method (ORTIZ-CONDE et al., 2002), it can be inferred from the plot that the device threshold voltage V_{TH} lays around -0.2 V for both bulk and FD-SOI transistors.

For the $I_D \times V_D$ simulations, in turn, the characteristics were extracted for three different gate voltages, namely $V_G = -0.5$ V, $V_G = -0.7$ V, and $V_G = -1.0$ V, all of them biasing the transistors in the strong inversion region. The drain voltage V_D was swept

Figure 5.9: $I_D \times V_D$ characteristics of the case study transistors: bulk (a) and FD-SOI (b) MOSFETs.



(a)



(b)

Source: author.

from 0.0 V to -1.5 V, also in steps of 0.1 V. The extracted $I_D \times V_D$ curves are depicted in Figure 5.9, for both bulk and FD-SOI transistors. As one can observe in the plots, the case study transistors displayed proper characteristic curves, qualitatively in agreement with those expected for a MOSFET device. Note, however, that the slope of the $I_D \times V_D$ curves at the saturation region is not negligible, as one may desire. This behavior is due to the intrinsic and relatively high contact resistance.

6 RESULTS AND DISCUSSION

Aiming to analyze the proper operation of the simulator and to validate the simulation model proposed in this thesis, both bulk and FD-SOI MOSFETs were simulated for a variety of situations, i.e., considering several doping configurations, device geometry, and biases. In this chapter, however, the results are presented only for the case study structures introduced in Section 5.3. Most part of the results were extracted adopting a bias point of $V_{\mathrm{G}} = -1.0$ V and $V_{\mathrm{D}} = -1.0$ V, biasing the transistor in the saturation region. These applied voltages represent a typical biasing condition for the characteristic lengths of the device (TAUR; WANN; FRANK, 1998; SAHA, 2001). A portion of the results, however, was extracted as a function of the bias point.

First, the impact of self-heating on the case study devices was simulated, and the results are presented in Section 6.1 and Section 6.2 for bulk and FD-SOI transistors, respectively. Each one of these sections starts presenting a few general extracted parameters pertaining to the respective case study device operation. Then, temperature profiles for both acoustic and optical phonons, as well as the heat generation distribution within the structure are presented. The sections end addressing the impact of the self-heating on the device current capability. Besides the thermal analysis performed utilizing the phonon energy balance (PEB) model, some results evaluated via the Joule heating model are also presented for comparison. Section 6.3, in turn, presents results pertaining particle-based simulations of charge traps, as well as the interplay between self-heating effects and charge trap activity. The results presented herein show that the tool is suitable for performing relevant case study simulations.

6.1 Electro-thermal Simulations of Bulk MOSFETs

6.1.1 Carrier Profiles

Before proceeding to the analysis of the thermal results, some general parameters of the case study device were extracted for the isothermal condition, such as the carrier energy, carrier drift velocity, carrier travel time, electric field, carrier temperature and carrier density along the transistor structure. These data are presented and discussed in the following. Note that these data correspond to averages of an ensemble of devices, so eventual fluctuations can be seen. These fluctuations can be related, for instance, to the

random nature of the Monte Carlo algorithm or even the RDF effect.

Figure 6.1 and Figure 6.2 depict respectively the average carrier energy E_{avg} and carrier drift velocity υ_{D} as a function of the position along the transistor length. For a given position along the length x, these quantities are calculated as the summation of the quantity along the depth y and width z, and then, normalizing the sum by the carrier concentration ρ, i.e.,

$$E_{\text{avg}}(x) = \frac{\sum_{y=0}^{y_{max}} \sum_{z=0}^{z_{max}} E_{\text{avg}}(x,y,z)}{\sum_{y=0}^{y_{max}} \sum_{z=0}^{z_{max}} \rho(x,y,z)} \tag{6.1}$$

and

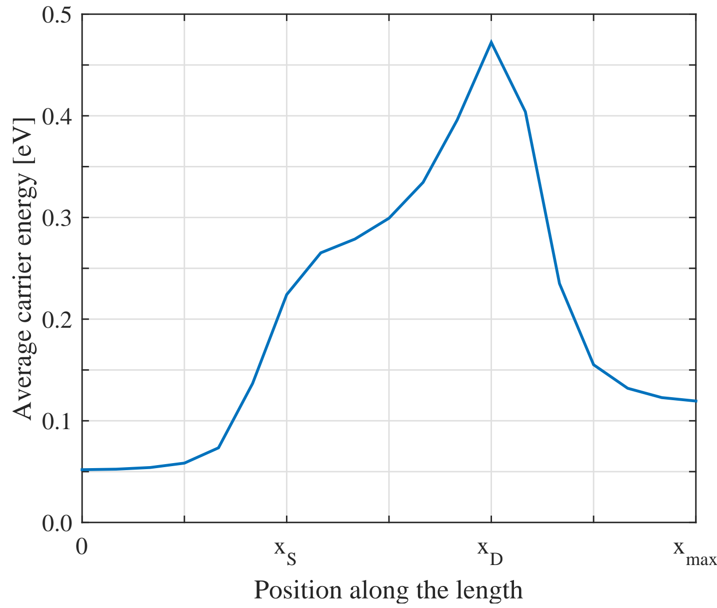
$$\upsilon_{\text{D}}(x) = \frac{\sum_{y=0}^{y_{max}} \sum_{z=0}^{z_{max}} \upsilon_{\text{D}}(x,y,z)}{\sum_{y=0}^{y_{max}} \sum_{z=0}^{z_{max}} \rho(x,y,z)}. \tag{6.2}$$

Unlike other plots in this section, these characteristics are better visualized in 1D rather than in 3D. Regarding the carrier energy, it is possible to note that carriers at the source region exhibit an energy level which is close to $3/2k_{\text{B}}T$. As they move towards the junction and enter into the conduction channel — driven by the electric field —, their energy rapidly increases, and it peaks at the junction between the channel and the drain region. As carriers enter into the drain region, they promptly start to lose energy due to the increased number of collisions. At a certain point within the drain region, however, the carrier energy starts to stabilize at a certain level, which is slightly higher than the energy level at the source, due to the applied voltage. Additionally, the carrier energy can be directly translated into carrier temperature, which is depicted in Figure 6.5.

Regarding the total carrier drift velocity $\upsilon_{\text{D}t}$, presented in Figure 6.2, its dominant component is the velocity along the x-axis (length) $\upsilon_{\text{D}x}$ due to the intense lateral electric field. Along the y-axis (depth), carriers also exhibit significant velocity at some regions, due to the vertical electric field. Negative velocity along such direction means that the carriers are moving from the substrate towards the Si/SiO_2 interface, and positive values mean the opposite. At the contacts, for instance, the velocity along the y-axis is typically negative since the carriers are leaving the device bulk and entering the metal contact. The velocity along the z-axis (width) $\upsilon_{\text{D}z}$, in turn, is virtually negligible, since there is almost no electric field gradient in that direction. An average snapshot of the electric field profile across the transistor structure is depicted in Figure 6.3.

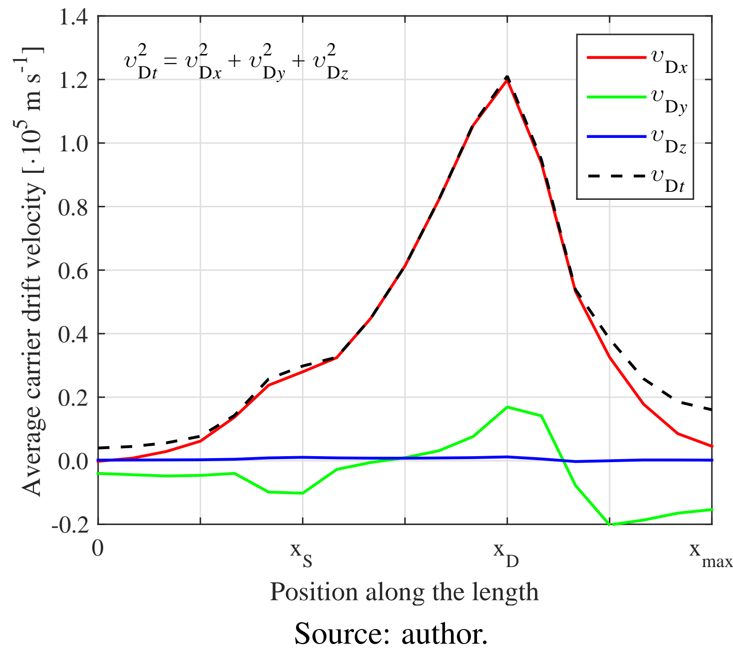
Another parameter extracted from the simulations was the time a carrier takes to go from the source region to the drain region. The distribution of the carrier travel time is presented in Figure 6.4, which also highlights the travel time for a particle traveling

Figure 6.1: Average carrier energy as a function of the position along the transistor length for 100 case study bulk MOSFETs at $V_D = V_G = -1.0$ V.



Source: author.

Figure 6.2: Average carrier drift velocity as a function of the position along the transistor length for 100 case study bulk MOSFETs at $V_D = V_G = -1.0$ V.

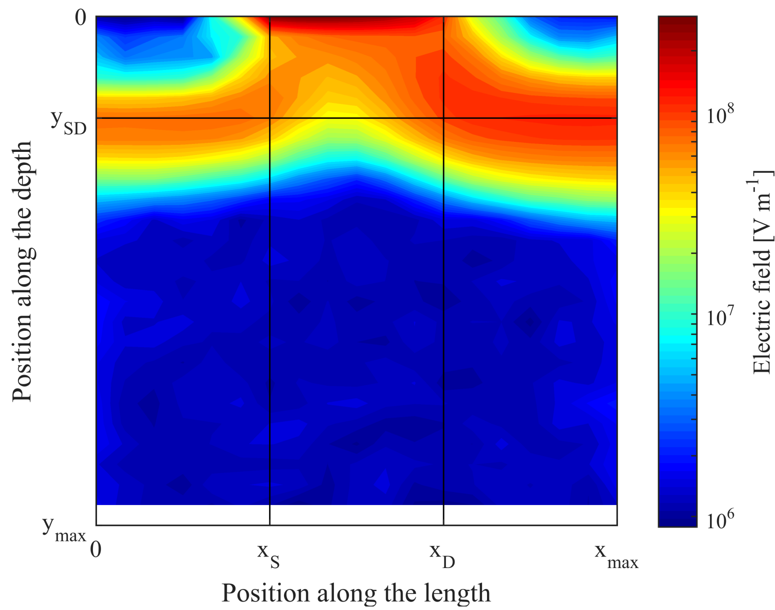


Source: author.

at the velocity saturation and at the thermal velocity limits[1]. It is possible to note that a

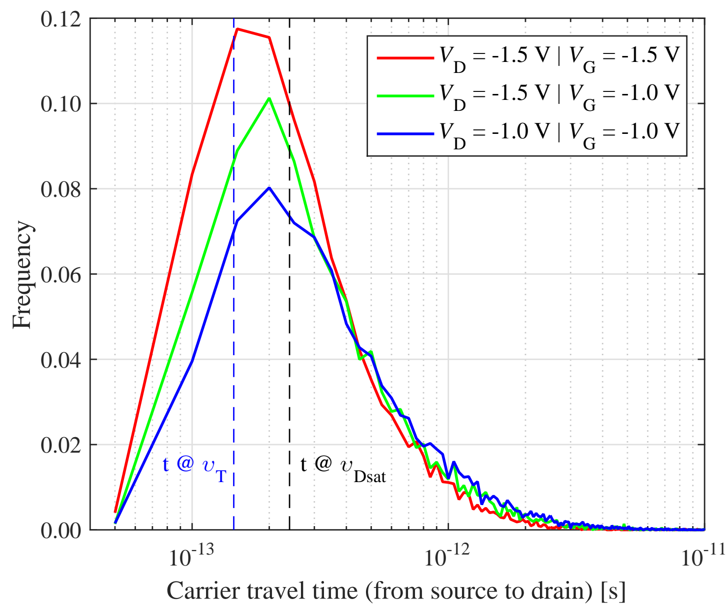
[1] At 300 K, the hole saturation velocity υ_{Dsat} in silicon is in the order of 0.7×10^5 ms^{-1} for electric fields up to 5×10^6 Vm^{-1}(CANALI; OTTAVIANI; QUARANTA, 1971), but it can be as high as $0.96 \sim 1.06 \times 10^5$ ms^{-1} for electric fields larger than 2×10^7 Vm^{-1} (SEIDEL; SCHARFETTER, 1967), which is the case here. Considering a channel with length $L_G = 24$ nm, the time spent by a certain particle to cross it, traveling at υ_{Dsat}, is $t_{\text{sat}} \approx 2.4 \times 10^{-13}$ s. Hole thermal velocity was assumed to be 1.65×10^5 ms^{-1}(IOFFE, 2018).

Figure 6.3: Average electric field profile as a function of the position along the transistor length and depth for 100 case study bulk MOSFETs at $V_{\mathrm{D}} = V_{\mathrm{G}} = -1.0$ V.



Source: author.

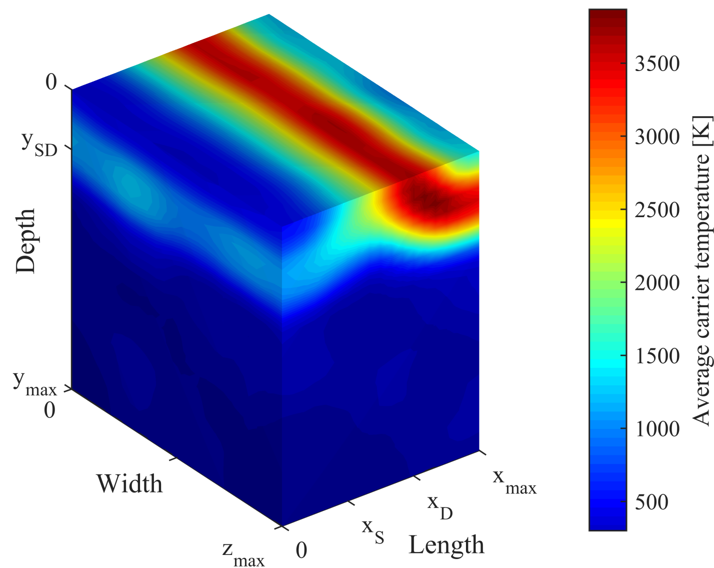
Figure 6.4: Distribution of the carrier travel time from source to drain as a function of the bias for 100 case study bulk MOSFETs.



Source: author.

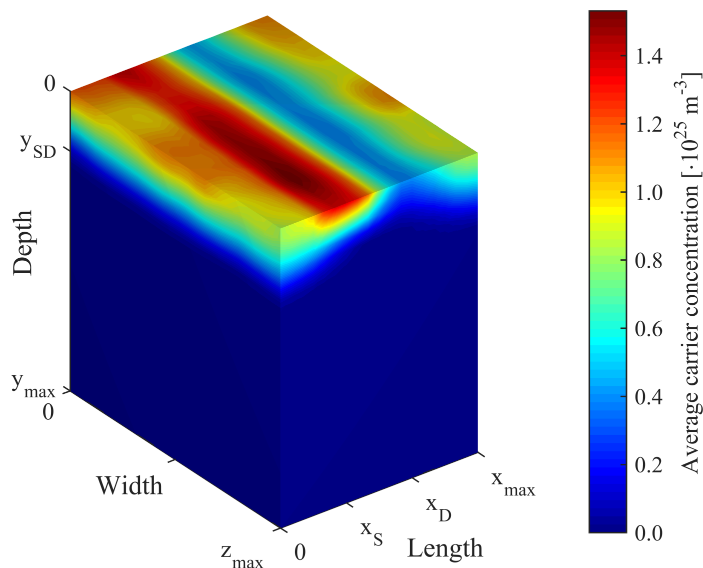
portion of the carriers crosses the channel ballistically, at very short times ($t \leq 10^{-13}$ s). A significant portion of the carriers crosses the channel with velocities slightly higher than the saturation velocity (i.e., $t < t_{\mathrm{sat}}$), indicating the occurrence of velocity overshoot effect

Figure 6.5: Average carrier temperature along the transistor structure for 100 case study bulk MOSFETs at $V_D = V_G = -1.0$ V.



Source: author.

Figure 6.6: Average carrier concentration along the transistor structure for 100 case study bulk MOSFETs at $V_D = V_G = -1.0$ V.



Source: author.

(LAUX; FISCHETTI, 1997; SINITSKY et al., 1997). On the other hand, there is also very sluggish carriers, which take one order of magnitude or higher times to travel the same distance. By increasing the bias, the travel time does not reduce significantly, i.e., the

distribution practically does not shift horizontally in Figure 6.4; the current is increased since the number of particles traveling at higher velocities increases (the distribution shifts vertically).

Finally, Figure 6.6 depicts the carrier density along the case study transistor. At the source and drain regions, the concentration remains close to the net doping concentration $|N_A - N_D|$, as expected. Hence, the inversion layer at the channel region is noticeable, and the pinch-off region, characteristic of devices operating in strong inversion condition, is clearly visible.

It shall be reinforced, however, that the data regarding the device electrical characteristics presented in this subsection are not a novelty. Instead, they emphasize and reaffirm the proper operation of the case study bulk device. Additionally, these data are helpful to explain some of the results presented further in the present section.

6.1.2 Temperature Profiles

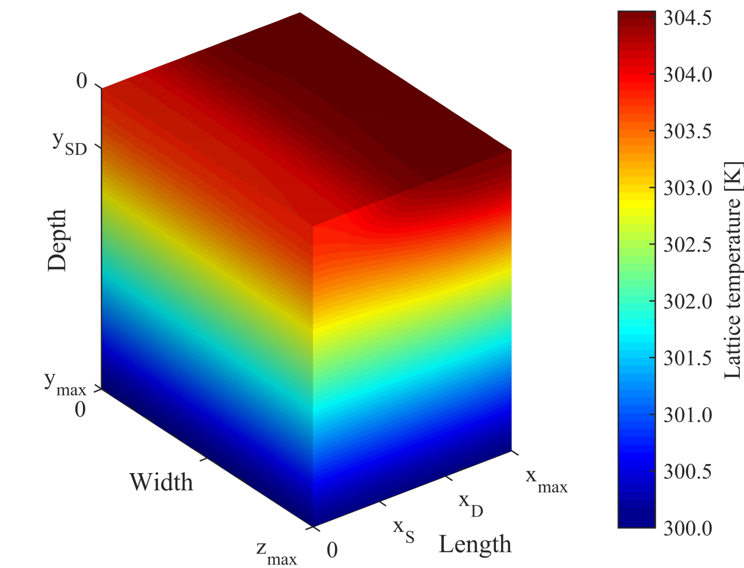
One of the most important results that can be extracted from an electro-thermal simulation is the temperature profile. It provides, basically, the temperature distribution across the device; the hottest and coldest spots in the structure. This information is helpful, for instance, to determine where in the transistor the charge transport is more or less affected by the temperature.

In this regards, Figure 6.7 depicts the acoustic phonon (lattice) temperature profile for the case study transistor evaluated via the PEB model (T_{PEB}) and Joule heating model ($T_{J\cdot E}$), respectively[2]. For both cases, the temperature rise peaks on the drain region of the device, as expected (LAI; MAJUMDAR, 1996; RAMAN; WALKER; FISHER, 2003; POP et al., 2001; RALEVA et al., 2012). The characteristics of the temperature hot spot, however, differs from model to model. While for the PEB model the hot spot is found to be spread out along the drain region, for the Joule heating model the hot spot is located exactly at the junction between the conduction channel and the drain region, where both electric field and carrier drift velocity are maximum. Hence, for this particular bias condition, the temperature rise calculated from the PEB model is found to be slightly higher than the temperature rise extracted via the PEB model.

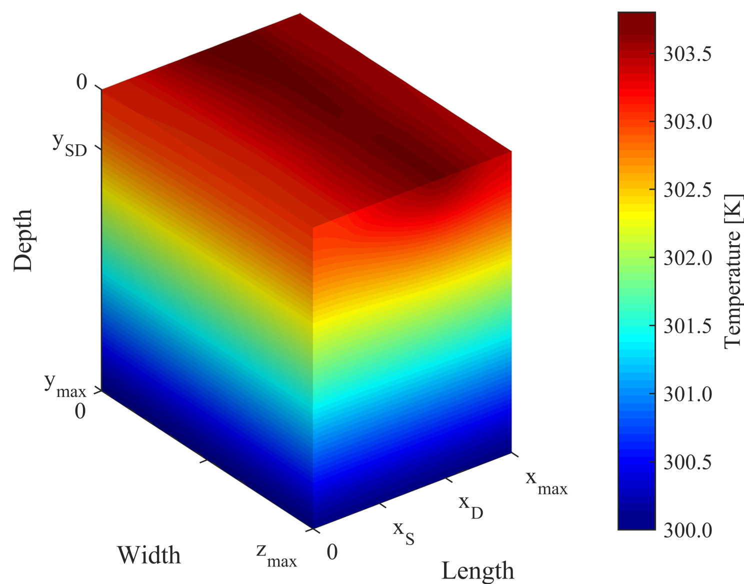
Figure 6.8 presents the optical phonon temperature profile within the transistor.

[2]Since the thermal transport is treated in steady-state, the plots present the temperature at the thermal equilibrium.

Figure 6.7: Average acoustic phonon (lattice) temperature profile evaluated via the PEB model (a) and Joule heating model (b) for 100 case study bulk MOSFETs at $V_{\mathrm{D}} = V_{\mathrm{G}} = -1.0$ V.



(a)

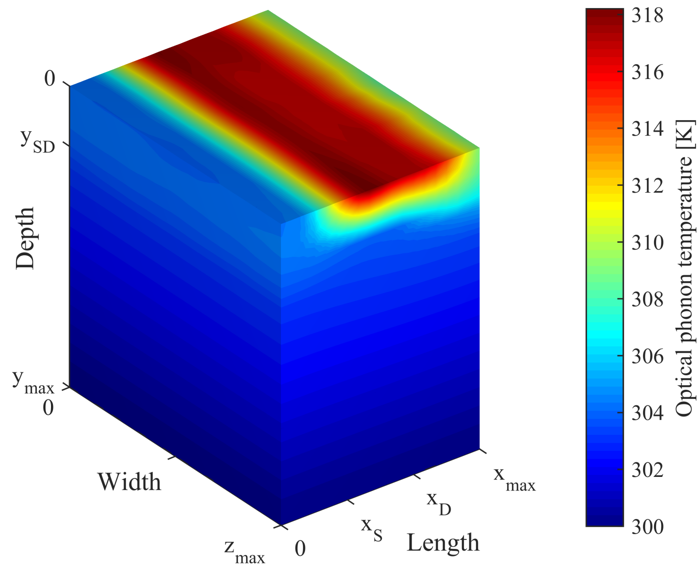


(b)

Source: author.

Unlike the acoustic phonon temperature, optical phonon temperature hot spot appears to be more distributed across the channel of the device. Analyzing the expression for the optical phonon temperature, given in Equation 4.12, one may infer that, assuming constant carrier concentration ρ, the optical phonon temperature is expected to follow

Figure 6.8: Average optical phonon temperature profile for 100 case study bulk MOSFETs at $V_D = V_G = -1.0$ V.

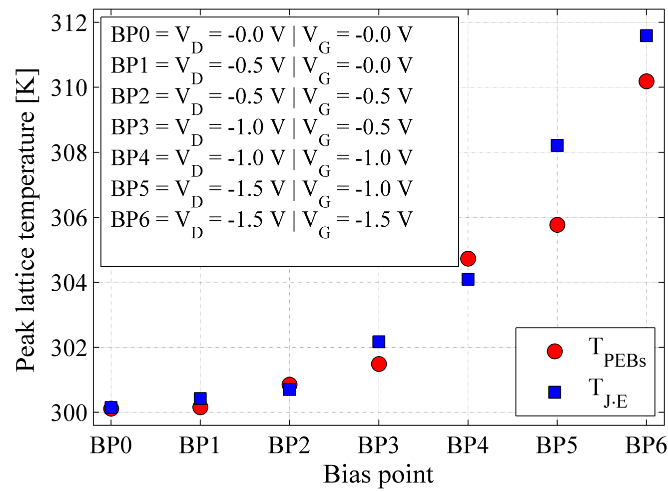


Source: author.

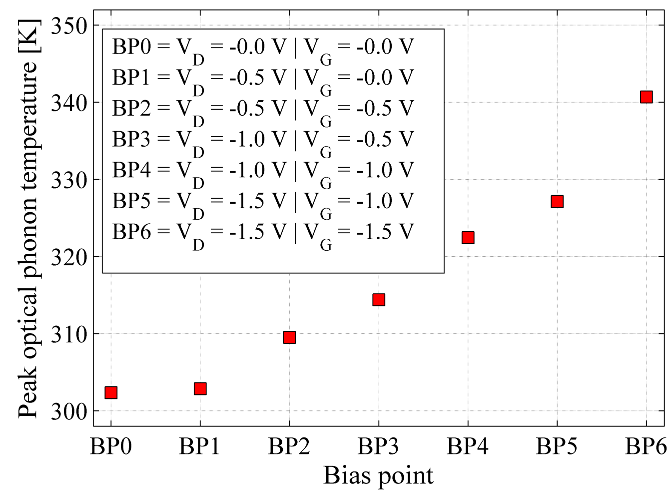
the carrier temperature T_c. On the other hand, assuming uniform carrier temperature, the optical phonon temperature is expected to follow the carrier concentration. However, in a practical situation, the transistor typically presents non-uniform concentration and non-uniform carrier temperature. In this case, the optical phonon temperature will follow the product of ρ and T_c. Therefore, as the carrier concentration peaks at the source end of the channel and the carrier temperature peaks at the drain end of the channel, optical phonon temperature is more pronounced between those regions. In the regions where either the carrier concentration is small or the carrier temperature approaches 300 K, the optical phonon temperature approaches the acoustic phonon temperature, as expected.

One may argue, however, that the phonon temperature rises extracted for the case study device are quite modest, mainly the acoustic phonon one, which increases roughly 4 K. This occurs due to the fact that the temperature rise is strongly dependent on the biasing conditions. Figure 6.9 depicts the peak temperature for both acoustic and optical phonons as a function of the applied bias. Note that for higher biases, for instance $V_D = V_G = -1.5$ V, acoustic phonon temperature may exceed 310 K, whereas the optical phonon temperature can reach 340 K. For the highest drain biases, the peak temperatures T_{PEBs} and $T_{J\cdot E}$ slightly deviate from each other. This occurs because, for the Joule heating model, the heat dissipation is concentrated at the channel-drain junction and, thus, it induces larger temperature rise in that location. The heat dissipation evaluated via the phonon energy

Figure 6.9: Average acoustic phonon (lattice) peak temperature (a) and optical phonon peak temperature (b) as a function of the applied bias for 100 case study bulk MOSFETs. Error bars are in the order of the marker size and are not shown.



(a)



(b)

Source: author.

balance model, on the other hand, is more spread out along the drain region, so a smaller temperature increment is expected.

6.1.3 Heat Generation Profile

The heat generation[3] within the transistor can also be extracted and it provides insights about the temperature hot spot location and power dissipation, for example. It is expected that the temperature rise peaks at or nearly the region with the most intense heat dissipation. Hence, by integrating the heat generation over the volume, one may obtain the total power that is being dissipated as heat.

In the regard of the phonon energy balance model, the amount of heat generated H_{PEBs} is given by the RHS of Equation 4.11, which is rewritten here for the sake of convenience.

$$H_{\text{PEBs}} = \frac{3}{2}\rho k_{\text{B}}\left(\frac{T_{\text{c}} - T_{\text{OP}}}{\tau_{\text{c}-\text{OP}}}\right) + \frac{\rho m^* \upsilon_{\text{D}}^2}{2\tau_{\text{c}-\text{OP}}}. \quad (6.3)$$

Note that in Equation 6.3 the acoustic phonon contribution on the heat generation is properly neglected. Hence, the carrier-optical phonon relaxation time $\tau_{\text{c}-\text{OP}}$ used in the simulations for the bulk case study device was 0.2 ps.

Similarly, for the Joule heating model the heat generation can assume any form between Equation 3.4 and Equation 3.9. Nevertheless, Monte Carlo simulators are typically unable to properly deal with generation/recombination processes. Thus, the heat generated through the Joule heating model $H_{\text{J}\cdot\text{E}}$ in the scope of this work was computed via the most straightforward approach, i.e.,

$$H_{\text{J}\cdot\text{E}} = J \cdot E. \quad (6.4)$$

Note that the current density J is given by (ZEGHBROECK, 2011)

$$J = q\rho\upsilon_{\text{D}}, \quad (6.5)$$

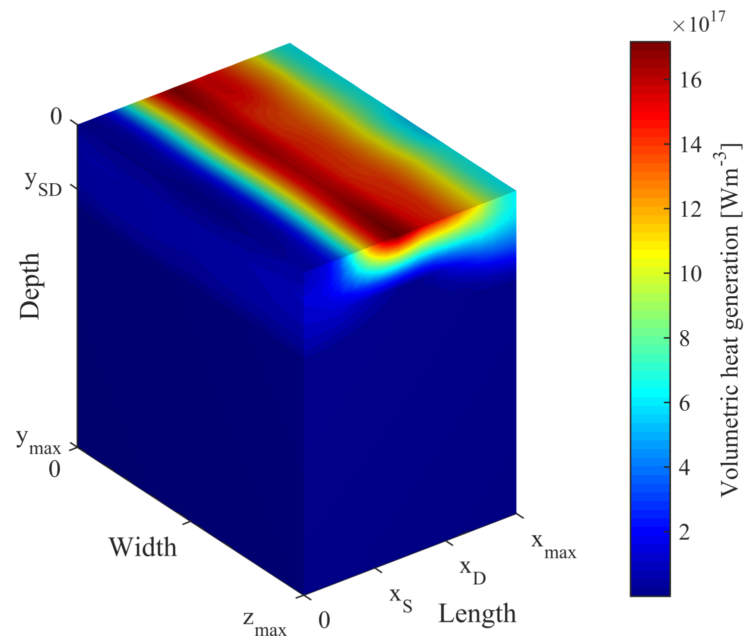
and, thus, the heat generation $H_{\text{J}\cdot\text{E}}$ can be rewritten as

$$H_{\text{J}\cdot\text{E}} = q\rho(\upsilon_{\text{D}} \cdot E). \quad (6.6)$$

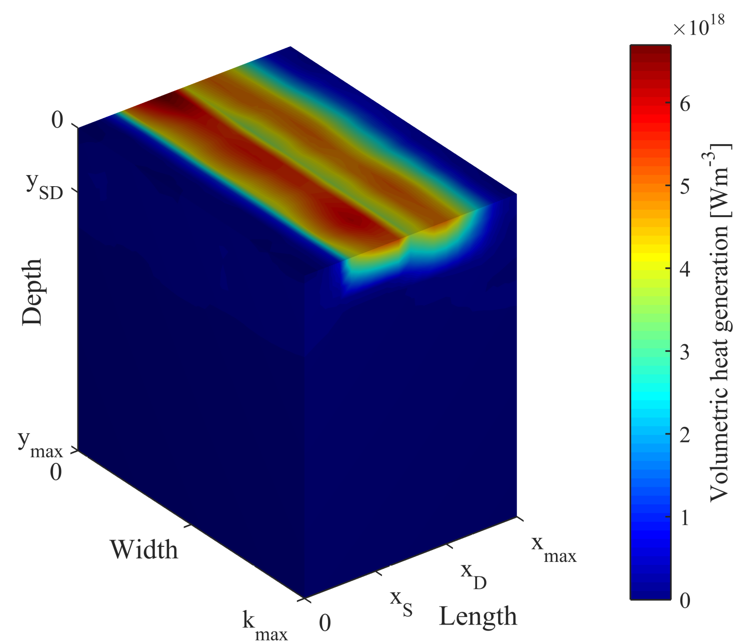
The volumetric heat generation H within the transistor evaluated through the two models presented above is depicted in Figure 6.10. From the PEB model, the heat generation across the device is found to be maximum at the interface and it is spread out through the device, extending itself from the source-channel junction, where it peaks due to the

[3] In a 3D domain, the heat generation is often called *volumetric* heat generation. Its usual unit is Wm^{-3}.

Figure 6.10: Average volumetric heat generation within the transistor evaluated via phonon energy balance model (a) and Joule heating model (b) for 100 case study bulk MOSFETs at $V_D = V_G = -1.0$ V.



(a)

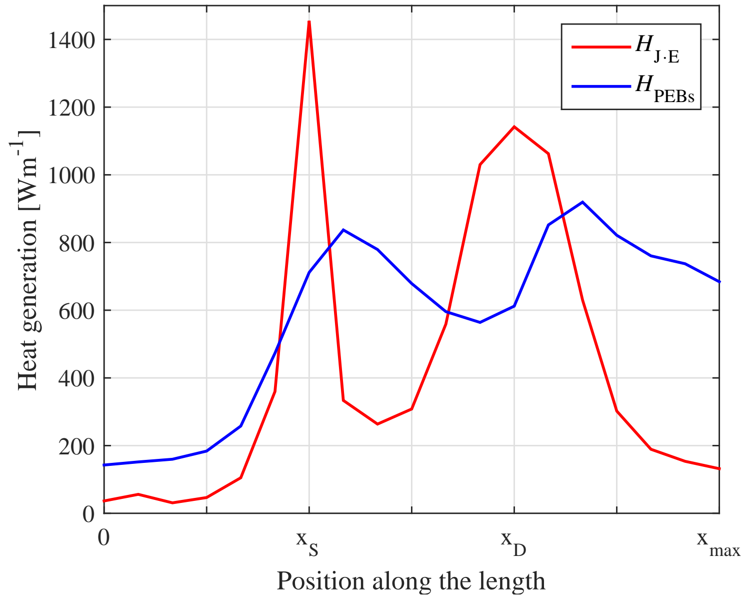


(b)

Source: author.

carrier concentration, up to the drain region. Conversely, for the Joule heating model the heat generation also peaks at the interface, however, it is more pronounced in the vicinity of the source-channel, where the carrier concentration peaks, and channel-drain junction,

Figure 6.11: Average heat generation along the transistor length according to the PEB model and Joule heating model for 100 case study bulk MOSFETs at $V_{\mathrm{D}} = V_{\mathrm{G}} = -1.0$ V.



Source: author.

where the electric field and the carrier drift velocity peak. For both models, the volumetric heat generation is on the order of 10^{18} Wm^{-3}, which is consistent with literature data (POP; SINHA; GOODSON, 2006; NI et al., 2012).

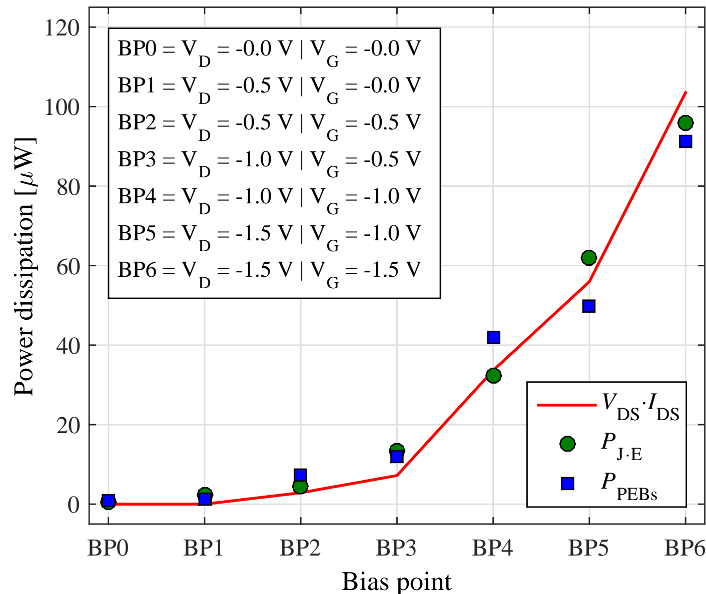
In order to understand the reason why the lattice temperature rise peaks at the drain side of the transistor even though the volumetric heat generation may peak at the source side, one has to calculate the heat generation along the transistor length (x-axis), which is given by

$$H(x) = \int_0^{y_{\max}} \int_0^{z_{\max}} H(x,y,z) \cdot dydz. \tag{6.7}$$

The calculated curves are depicted in Figure 6.11 for both the PEB model and the Joule heating model.

When represented in 1D, one may observe that the heat generation no longer peaks at the source end of the channel. For the PEB model, the yielded curve has two well-defined peaks; one where the carrier concentration peaks and another where the carrier temperature peaks (see Figures 6.5 and 6.6). Similarly, the curve evaluated through the Joule heating model presents also two well-defined peaks; one where the carrier concentration peaks, and another where the carrier drift velocity and electric field peaks. Note that, for both models, the larger portion of the heat is being generated in the drain side of the transistor, where the temperature hot spot was found to be located.

Figure 6.12: Average power supplied to the transistor (line) and average power dissipated as heat (squares) for each thermal model as a function of the applied bias for 100 case study bulk MOSFETs. Error bars are not shown.



Source: author.

Similarly to the integration performed to evaluate the curves depicted in Figure 6.11, by integrating the volumetric heat generation H over the total volume of the device V, one can obtain the total power P that is being dissipated as heat, i.e.,

$$P = \int H \cdot dV. \tag{6.8}$$

For the sake of energy conservation, however, the power dissipated must be approximately the same, in spite of the model utilized for the computation of the heat generation term. Moreover, the power dissipated as heat has to be also very close to the electric power supplied to the device. In this context, Figure 6.12 depicts the transistor power dissipation for several bias points, from accumulation to strong inversion region. The electric power consumed by the transistor is shown as the product $V_{\mathrm{DS}} \cdot I_{\mathrm{DS}}$, whereas P_{PEBs} and $P_{\mathrm{J \cdot E}}$ stand for the total power dissipated as heat calculated via phonon energy balance model and Joule heating model, respectively. For the entire range of simulated biases, the power dissipated yielded from both models is in agreement with power consumed by the device.

Table 6.1: Average drain current for isothermal and non-isothermal simulations as a function of the applied bias for 1000 case study bulk MOSFETs. The current variation for each case is calculated over the current extracted from the isothermal simulation with $T_{\text{sink}} = 300$ K.

Simulation Type	Average drain current [μA/μm] vs Bias					
	$V_G = -1.0$ V				$V_G = -1.5$ V	
	$V_D = -1.0$ V		$V_D = -1.5$ V			
Isothermal $T_{\text{sink}} = 300$ K	339.29	—	376.75	—	696.60	—
Non-iso (J·E model)	338.59	−0.21%	375.19	−0.41%	691.39	−0.75%
Non-iso (PEB model)	338.12	−0.34%	375.35	−0.37%	691.40	−0.75%
Isothermal $T_{\text{sink}} = 320$ K	335.98	−0.97%	373.03	−0.99%	688.06	−1.22%

6.1.4 Impact of Self-heating on the bulk MOSFET On-Current

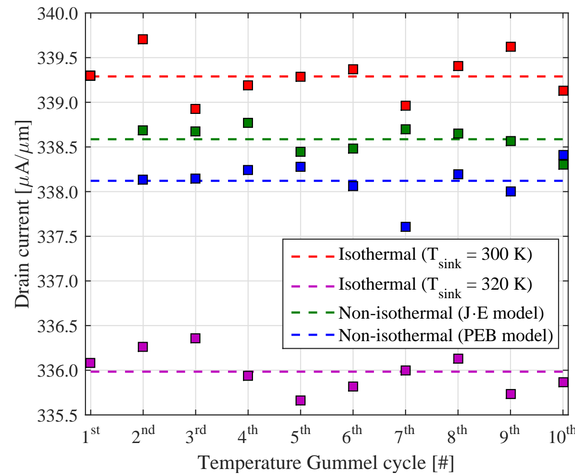
The impact of self-heating on the current capability of the transistor was also extracted using both the PEB model and Joule heating model for the evaluation of the phonon temperatures. To do that, first, isothermal simulations with $T_{\text{sink}} = 300$ K were carried out for 1000 case study transistors, in order to provide a reference current I_{D0}. The statistical seeds of such simulations were preserved and, then, non-isothermal simulations with $T_{\text{sink}} = 300$ K were performed for the same ensemble of devices and a correspondent current distribution I_D was extracted. These currents and their average value for three different bias points, namely $V_D = -1.0$ V and $V_G = -1.0$ V; $V_D = -1.5$ V and $V_G = -1.0$ V; and $V_D = -1.5$ V and $V_G = -1.5$ V, are depicted in Figure 6.13. Current values are normalized by the device width and expressed in units of μA/μm.

It is clear in Figure 6.13 that the non-isothermal simulations yield less drain current than isothermal simulation for the same population of devices. From these currents, a current deviation ΔI_D can be calculated as

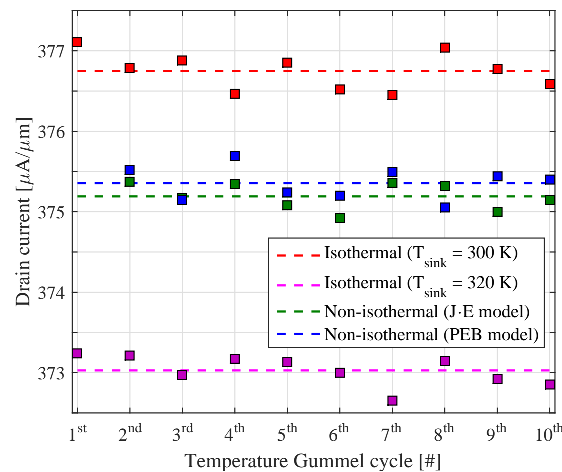
$$\Delta I_D = I_D - I_{D0}. \tag{6.9}$$

The average current as well as the average current deviation according to the applied bias are summarized in Table 6.1. The average current deviation, normalized by the isothermal current at $T_{\text{sink}} = 300$ K, is also presented in Figure 6.14 for a 95% confidence interval.

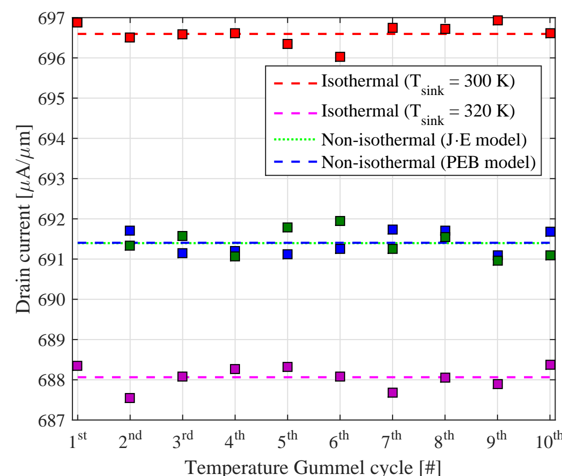
Figure 6.13: Drain current as a function of the temperature Gummel cycle (i.e., time) for isothermal and non-isothermal simulations at $V_D = -1.0$ V and $V_G = -1.0$ V (a), $V_D = -1.5$ V and $V_G = -1.0$ V (b), and $V_D = -1.5$ V and $V_G = -1.5$ V (c). Dashed lines represent the average value for each case.



(a)



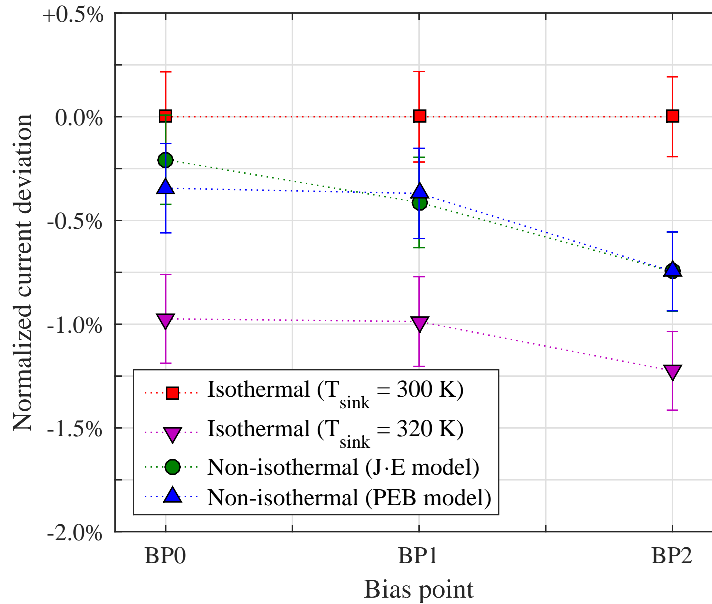
(b)



(c)

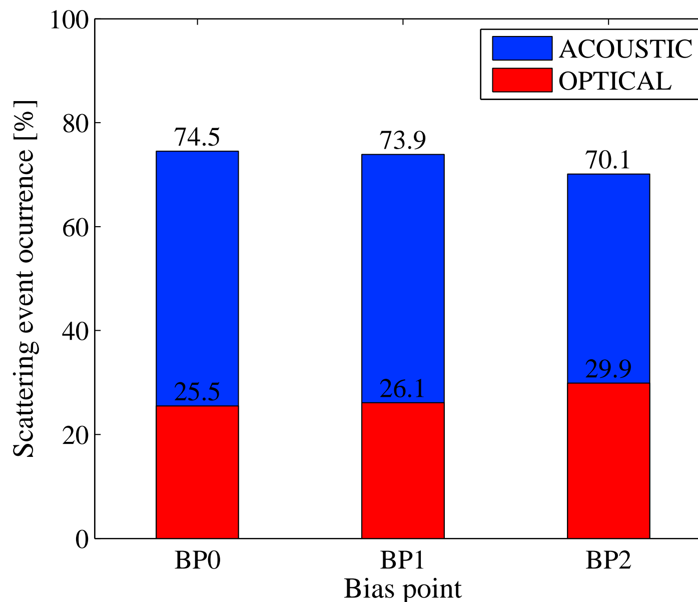
Source: author.

Figure 6.14: Average drain current deviation for isothermal and non-isothermal simulations according to the applied biases BP0, BP1, and BP2, which correspond to $V_D = -1.0$ V and $V_G = -1.0$ V, $V_D = -1.5$ V and $V_G = -1.0$ V, and $V_D = -1.5$ V and $V_G = -1.5$ V, respectively. The values are normalized by the average current extracted from isothermal simulations with $T_{sink} = 300$ K. Error bars stand for the 95% confidence bounds.



Source: author.

Figure 6.15: Percentage of acoustic and optical scattering events as a function of the applied bias, where BP0, BP1, and BP2 correspond to $V_D = -1.0$ V and $V_G = -1.0$ V, $V_D = -1.5$ V and $V_G = -1.0$ V, and $V_D = -1.5$ V and $V_G = -1.5$ V, respectively.



Source: author.

Nonetheless, the current deviation due to self-heating effect in bulk devices was found to be quite modest. The main reason for the small current degradation is the slight acoustic phonon (lattice) temperature rise, which is, at most, on the order of 10 K for $V_D = V_G = -1.5$ V. For lower biases, the increment in the temperature is even less pronounced. Note that these temperature rises are for the hot spot regions (see Figure 6.9); at the conduction channel–where the most significant part of the charge transport occurs–, the temperature rise might be even lower.

The optical phonon temperature rise, in turn, was found to be much larger than the acoustic phonon one. However, even though it is taken into account in the non-isothermal simulations performed via PEB model, optical phonon scattering events represent less than 30% of the total scattering events, as presented in Figure 6.15. Thus, the current degradation is still mainly governed by the acoustic phonon temperature. As a figure of merit, isothermal simulations performed at $T_{\text{sink}} = 320$ K yielded more current degradation than the non-isothermal ones for all the three biases.

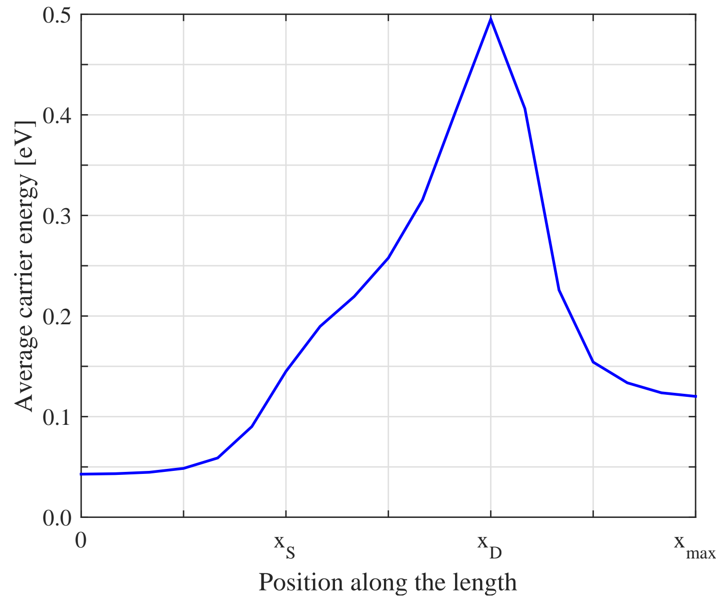
6.2 Electro-thermal Simulations in FD-SOI MOSFETs

6.2.1 Carrier Profiles

Similarly to the quantities extracted for the bulk case study device, some general parameters pertaining to the transistor operation were also extracted for the FD-SOI case study structure. These parameters are presented for the isothermal condition, considering $T_{\text{sink}} = T_{\text{gate}} = 300$ K. Figure 6.16 depicts the average carrier energy along the transistor length. The discussion presented in Subsection 6.1.1 for bulk devices is also valid here. At the source region, the carriers possess an energy level close to $3/2k_BT$. Their energy increases as they are accelerated in the channel region, and it peaks at the channel-drain junction. Upon entering the drain region, carriers start to lose energy due to the enhancement of the collision rate, and their energy abruptly drops, remaining slightly higher than the carrier energy at the source due to the applied bias. The carrier energy can be directly translated into carrier temperature, which is depicted in Figure 6.20b.

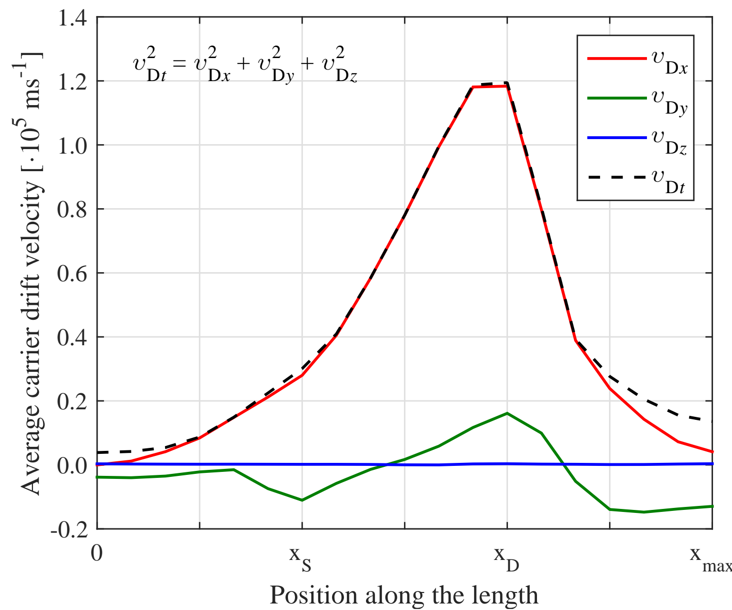
The average carrier drift velocity along the transistor length is presented in Figure 6.17. Note that the total drift velocity υ_{Dt} is basically dominated by x-component υ_{Dx}. Within the conduction channel, the slope of the drift velocity is more aggressive for SOI in comparison with bulk devices (see also Figure 6.2). Hence, it peaks and apparently

Figure 6.16: Average carrier energy along the transistor length for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V.



Source: author.

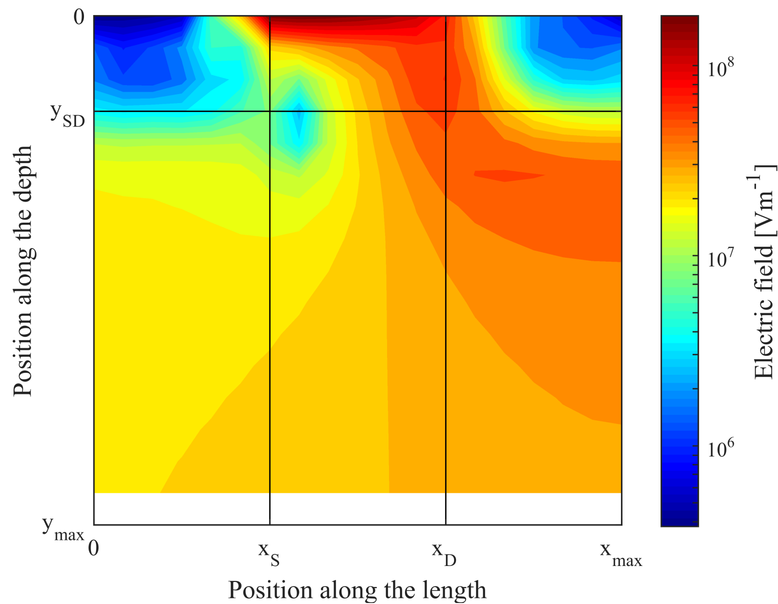
Figure 6.17: Average carrier drift velocity along the transistor length for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V.



Source: author.

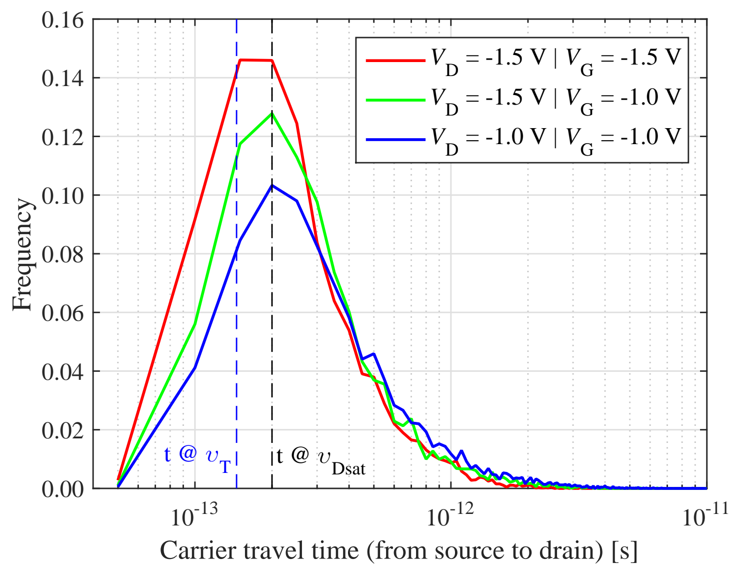
saturates at the pinch-off region at $\upsilon_{Dt} \approx 1.2 \cdot 10^5$ ms^{-1}. The total carrier drift velocity characteristic is closely related to the electric field profile, depicted in Figure 6.18. At the drain, the y-component of the velocity becomes significant, since the carriers are being vertically attracted to the metal contact. This helps to explain the energy level the carriers

Figure 6.18: Average electric field profile (absolute value) along the transistor length and depth for 100 case study FD-SOI MOSFETs at $V_{\mathrm{D}} = V_{\mathrm{G}} = -1.0$ V.



Source: author.

Figure 6.19: Distribution of the carrier travel time from source to drain as a function of the applied bias for 100 case study FD-SOI MOSFETs.

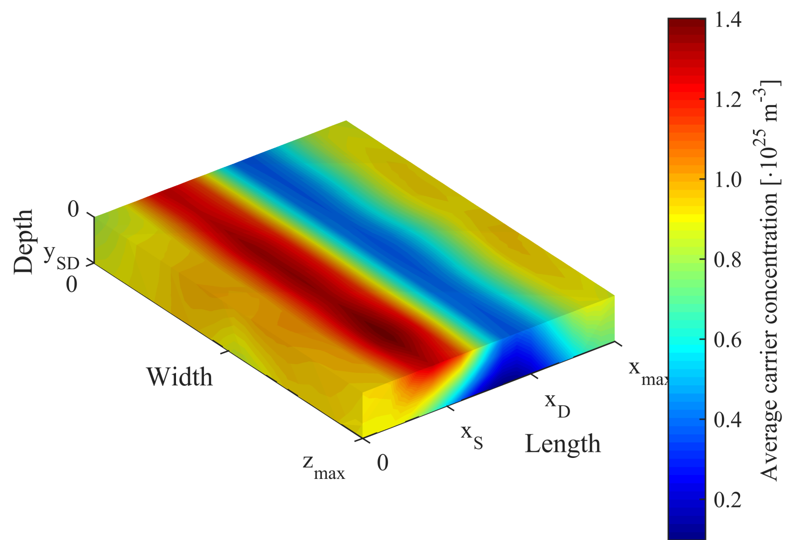


Source: author.

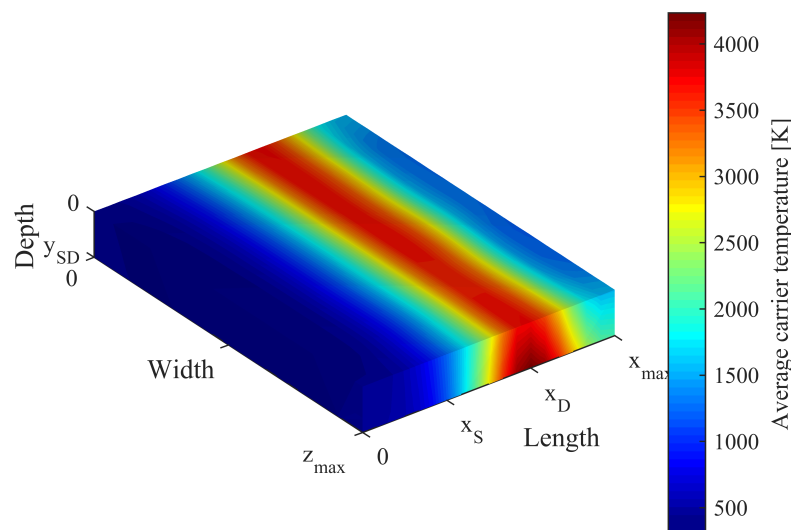
possess at that region.

The carrier travel time was also extracted for FD-SOI MOSFETs, and it is depicted in Figure 6.19 as a function of the applied bias. The corresponding times for a carrier traveling at the saturation velocity υ_{Dsat} and at the thermal velocity υ_{T} limit are high-

Figure 6.20: Average carrier concentration (a) and average carrier temperature (b) along the transistor silicon layer for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V.



(a)



(b)

Source: author.

lighted in the plot. The saturation velocity for this case was extracted from Figure 6.17 as $\upsilon_{Dsat} = 1.2 \cdot 10^5$ ms^{-1}. Similarly to the behavior observed for the bulk case study device (see Figure 6.4), as the applied bias is increased, travel times do not reduce significantly in average. On the other hand, an increase occurs in the number of particles traveling at higher velocities, i.e., the carrier travel time distribution shift vertically. Hence, in SOI

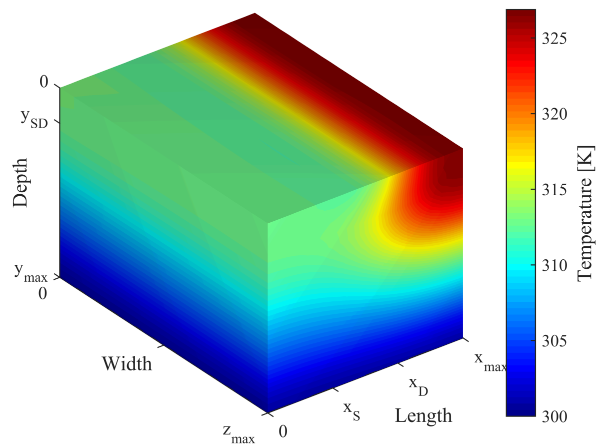
devices the carriers are confined within a thin layer of silicon, so they are more likely to move straightly from source to drain. As a consequence, the number of "sluggish" carriers (at the tail of the distribution) reduces, and the number of "fast" carriers (at the peak of the distribution) enlarges in comparison with bulk devices for a same given bias. In other words, the carrier travel time distribution becomes narrower and sharper for SOI transistors, and this behavior can be related to the enhancement of the drain current in such devices.

Finally, the average carrier concentration along the transistor silicon layer is depicted in Figure 6.20a. In spite of typical fluctuations caused by the RDF effect, one may note that, at the source and drain regions, the carrier concentration is close to the net impurity concentration, as expected. Since the device operates in strong inversion, both the inversion layer — more pronounced at the source side of the channel — and the pinch-off region are clearly visible.

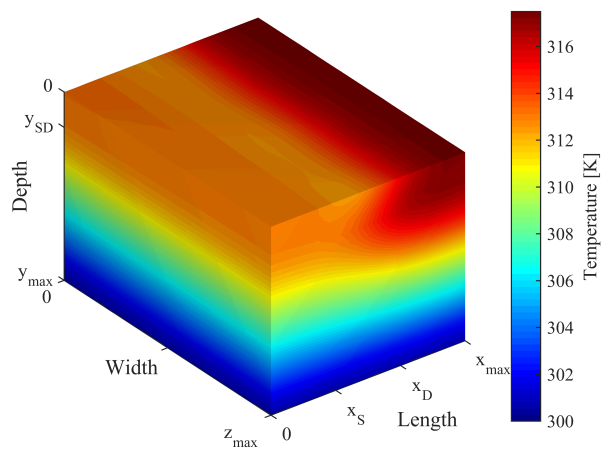
6.2.2 Temperature Profiles

The acoustic phonon (lattice) and optical phonon temperature distribution across the study FD-SOI MOSFET structure are presented in Figure 6.21. The acoustic phonon temperature was extracted via PEB model and Joule heating model, presented respectively in Figure 6.21a and Figure 6.21b. In both cases, the temperature hot spot was found to be located at the drain side of the transistor, as expected. The temperature rise at the hot spot, however, differs from model to model. As the temperature extracted via PEB model remains close to 327 K at the hot spot for $V_{\mathrm{D}} = V_{\mathrm{G}} = -1.0$ V, the one computed via Joule heating model is roughly 10 K lower. Such discrepancy occurs because the heat generation evaluated via Joule heating model is mostly concentrated at the junctions, i.e., in the vicinity of the gate (as presented in the following section). Since in SOI devices simulation the gate terminal is considered to be a heat sink, lower temperature rises are expected for this case. The optical phonon temperature, in turn, is depicted in Figure 6.21c. Even though it follows the acoustic phonon temperature profile, optical phonon temperature is around 5 K to 8 K higher at the transistor active area, characterizing the non-equilibrium condition between the phonon modes.

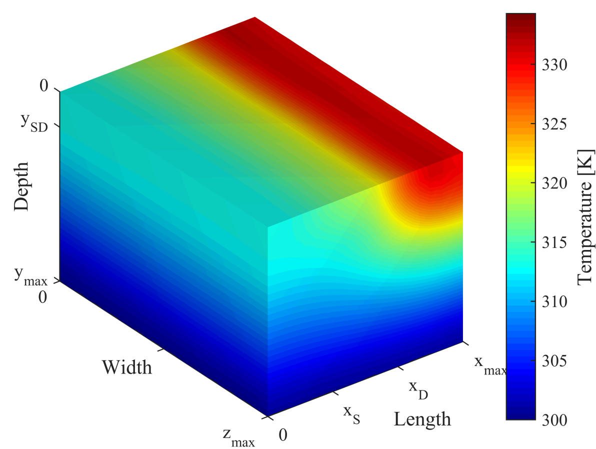
Figure 6.21: Average acoustic phonon (lattice) temperature profile evaluated via PEB model (a) and Joule heating model (b), and optical phonon temperature profile (c) for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V and $T_{gate} = 300$ K.



(a)



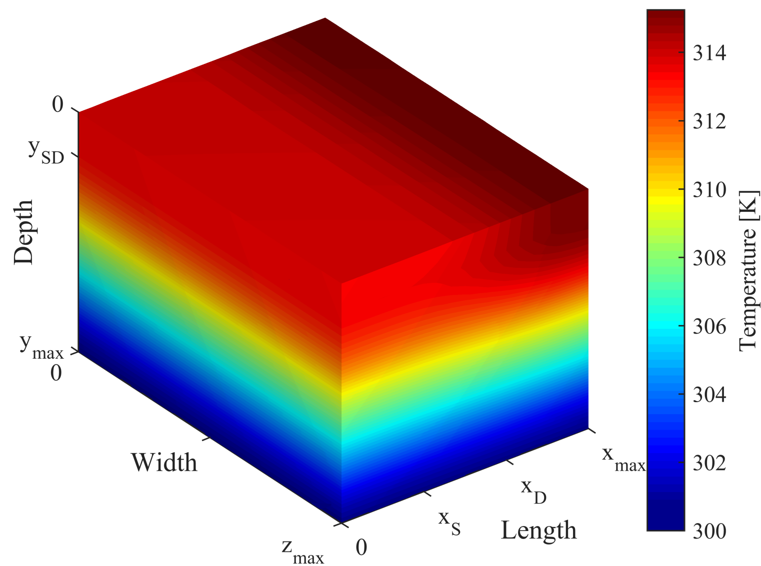
(b)



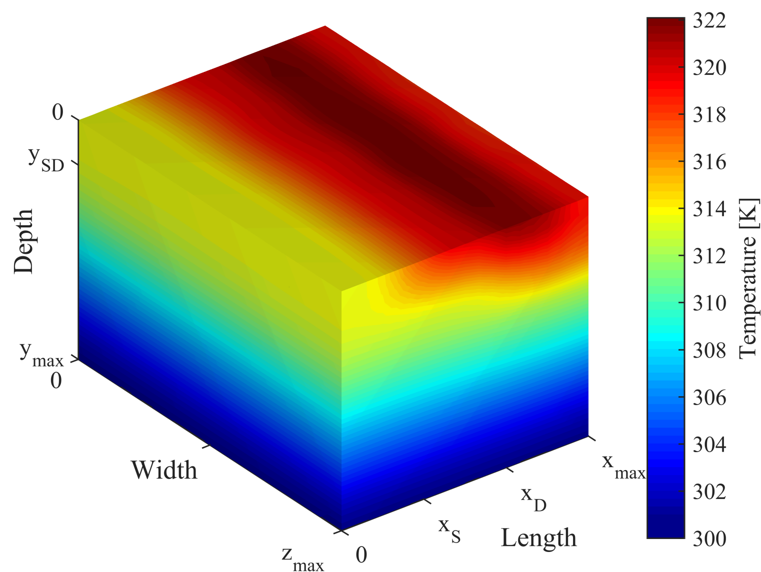
(c)

Source: author.

Figure 6.22: Average acoustic phonon (lattice) (a) and optical phonon (b) temperature profiles for 100 case study FD-SOI MOSFETs with silicon thermal conductivity depending only on the temperature at $V_{\mathrm{D}} = V_{\mathrm{G}} = -1.0$ V and $T_{\mathrm{gate}} = 300$ K.



(a)

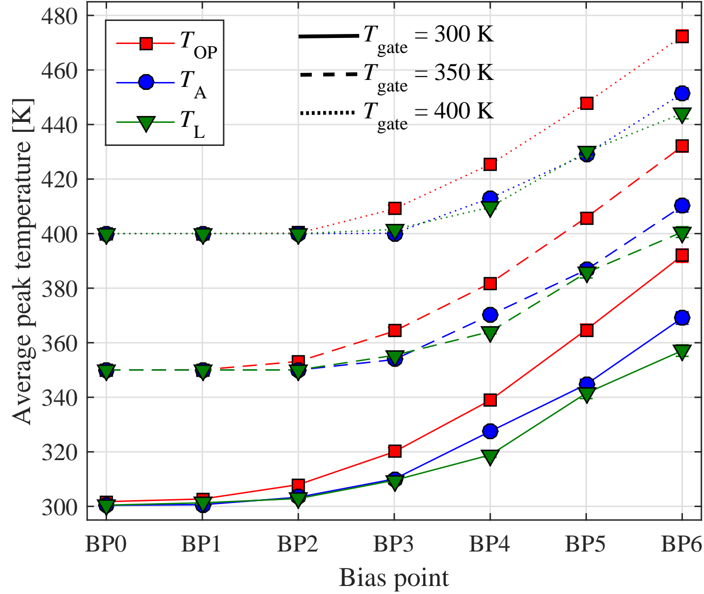


(b)

Source: author.

Another important characteristic of the FD-SOI profiles depicted in Figure 6.21 is the significant difference between the temperature at the drain and the temperature at the source region, taking into account the dimensions of the transistor. For bulk devices, as seen in Figure 6.7, such difference does not exceed 2 K, whereas, for SOI devices,

Figure 6.23: Average optical phonon (T_{OP}), acoustic phonon (T_A), and lattice (T_L) peak temperature for 100 case study FD-SOI MOSFETs as a function of the applied bias and gate temperature. Error bars stand for the 95% confidence bounds. The bias points were set as follows: BP0: $V_D = V_G = -0.0$ V; BP1: $V_D = -0.5$ V and $V_G = -0.0$ V; BP2: $V_D = V_G = -0.5$ V; BP3: $V_D = -1.0$ V and $V_G = -0.5$ V; BP4: $V_D = V_G = -1.0$ V; BP5: $V_D = -1.5$ V and $V_G = -1.0$ V; BP6: $V_D = V_G = -1.5$ V.



Source: author.

it can surpass 10 K for the acoustic phonon temperature, for instance. This behavior is related to the modulation of the silicon thermal conductivity by the temperature and silicon layer thickness. As presented in Section 4.4, the thin silicon layer of SOI devices possesses much lower thermal conductivity than the thick silicon layer of bulk structures[4]. In this way, the temperature spread and the heat dissipation across the SOI silicon film are severely reduced. As a consequence, the temperature rise is more localized at the heat generation points and the hot spots are intensified.

Aiming to capture the impact of the silicon thermal conductivity reduction due to the thickness of the silicon film on the temperature rise, a set of 100 case study FD-SOI devices was simulated utilizing a thermal conductivity model that depends only on the temperature, i.e., using Equation 4.14. The results are depicted in Figure 6.22. Note that the acoustic phonon temperature at the hot spot is roughly reduced by 40% in comparison with the value for the temperature and thickness-dependent thermal conductivity model (at the same bias and power dissipation). Hence, the temperature spread across the silicon

[4]For an SOI device with silicon layer thickness around 10 nm, and whose temperature varies between 300 K and 400 K — which is the case here —, a good approximation for the thermal conductivity is 13 $Wm^{-1}K^{-1}$(RALEVA et al., 2008; VASILESKA; RALEVA; GOODNICK, 2010a).

layer for the former case is significantly more uniform. The optical phonon temperature, in turn, does not depend directly on the thermal conductivity, but it also varies since it depends on the acoustic phonon temperature.

As the temperature rise is expected to vary with respect to the applied bias and the gate temperature, a set of 100 case study FD-SOI devices was simulated for several bias points and for three different gate temperatures. The average optical phonon, acoustic phonon, and lattice peak temperatures[5] are depicted in Figure 6.23. Note that the peak temperatures depend directly on the applied bias and on the gate temperature. Basically, by increasing the bias, the power dissipation also increases and, consequently, more pronounced temperature rises are expected (TENBROEK et al., 1996). In the same way, by raising the gate temperature, less heat is removed from the structure and, thus, larger hot spots are expected.

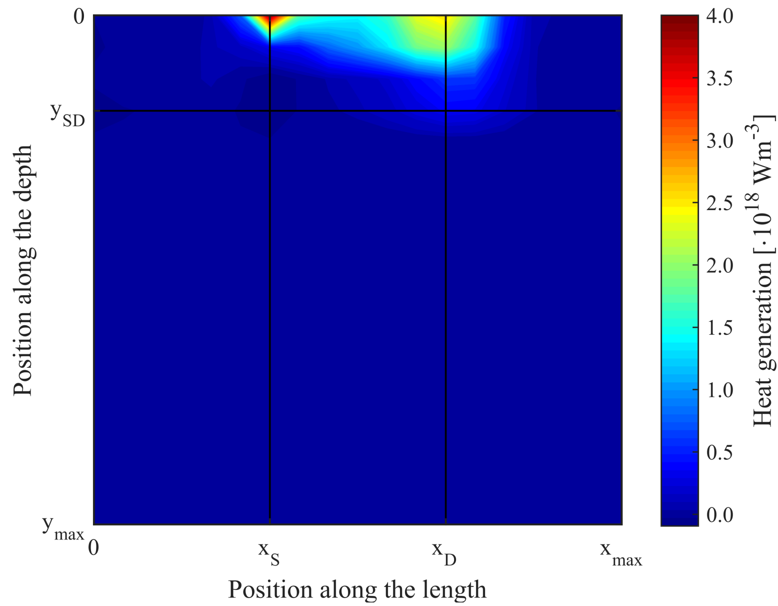
6.2.3 Heat Generation Profile

Figure 6.24 depicts the average volumetric heat generation profiles for 100 case study FD-SOI MOSFETs. Aiming to present these profiles in great detail, they were plotted in 2D (averaged along the transistor width) rather than in 3D. The heat generation evaluated via Joule heating model, shown in Figure 6.24a, is concentrated at the junctions, as expected. It peaks at the source end of the channel–due to the peak of the carrier concentration–, and at the drain end, where both the electric field and the carrier drift velocity are the most. At the latter region, however, the heat generation is more spread out over the junction. Entering in the drain region, the heat generation abruptly drops due to the strong reduction on the carrier drift velocity and electric field, as seen in Figure 6.17 and Figure 6.18.

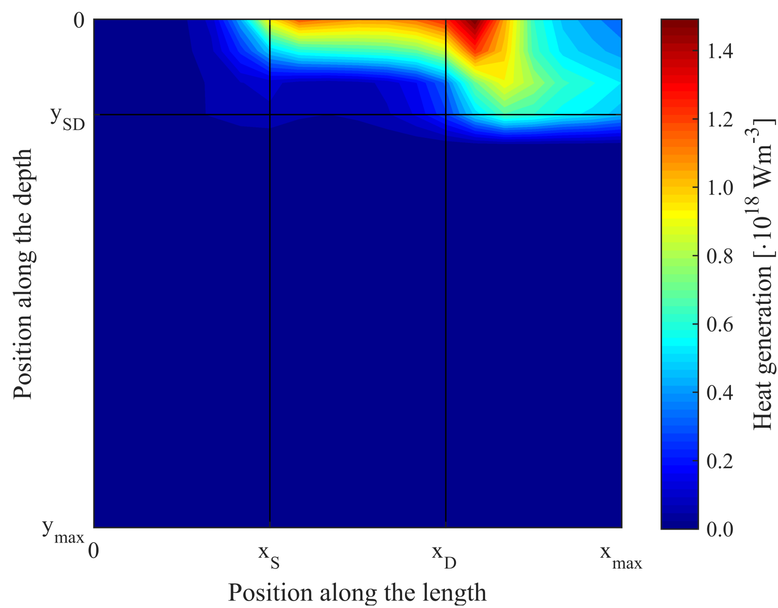
The heat generation computed via PEB model, shown in Figure 6.24b, depicts a more reasonable profile. One may observe that the heat is mostly generated near the interface, extending itself from the source end to drain end of the channel. The peak generation takes place within the drain region since that is the area where the carriers dissipate the energy gained from the electric field upon crossing the channel. Hence, the

[5] As presented in Chapter 3, the acoustic phonon temperatures and the lattice temperatures are assumed to be equal. In this plot, however, the nomenclature *acoustic phonon temperature* refers to the lattice temperature computed via PEB model, which differentiates between phonon modes. Since the Joule heating model does not differentiate phonon modes, the nomenclature *lattice temperature* refers to the lattice temperature extracted via Joule heating model.

Figure 6.24: Volumetric heat generation profile (averaged along the transistor width) evaluated via Joule heating model (a) and PEB model (b) for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V and $T_{gate} = 300$ K.



(a)

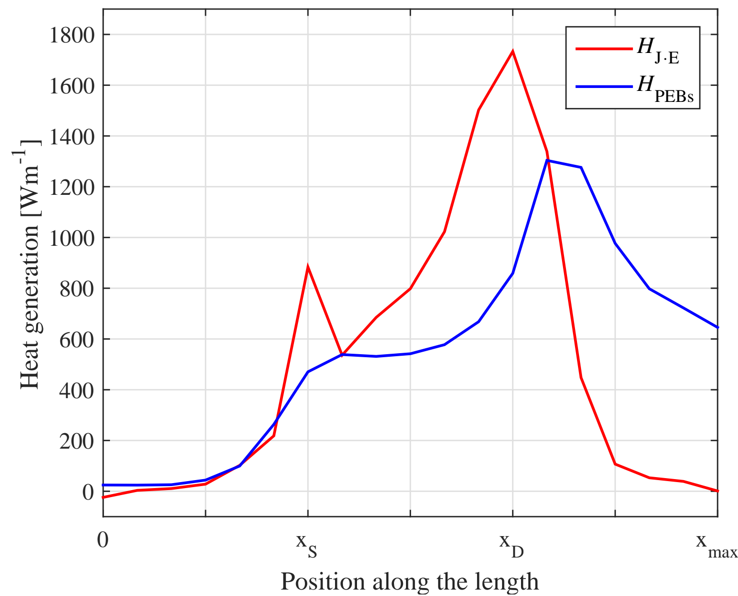


(b)

Source: author.

energy dissipation occurs all over the drain region, unlike the profile calculated via Joule heating model. The difference between the heat generation along the transistor length computed via each thermal model is clearly seen in Figure 6.25. Nevertheless, note that for both models the volumetric heat generation remains on the order of 10^{18} $\mathrm{Wm^{-3}}$, which is similar to the values found for bulk devices, and in agreement with literature predictions

Figure 6.25: Average heat generation along the transistor length for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V and $T_{gate} = 300$ K.



Source: author.

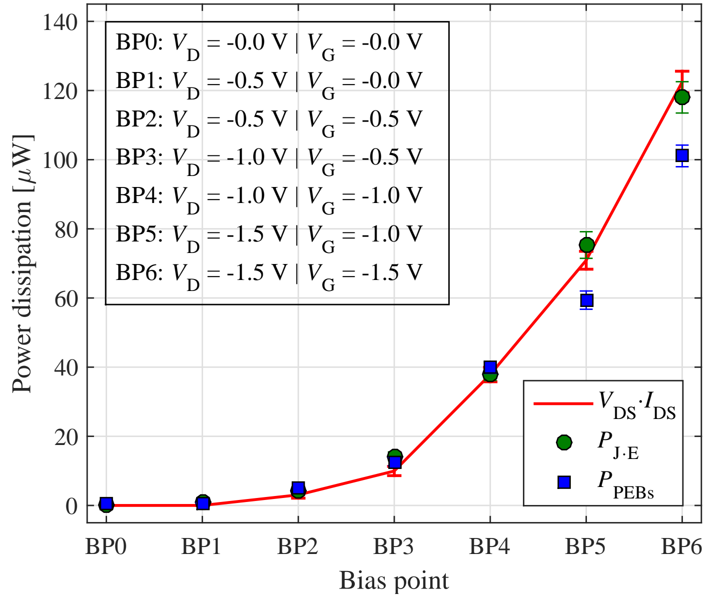
(POP; SINHA; GOODSON, 2006; NI et al., 2012).

As presented in Subsection 6.1.3, by integrating the volumetric heat generation over the total device volume, one may estimate the total power dissipation as heat, which has to be close to the electric power supplied to the device for the sake of energy conservation. In this way, Figure 6.26 depicts the electric power supplied to the device and the power dissipated as heat for each thermal model as a function of the applied bias. Note that the power dissipation estimated via Joule heating model is in good agreement with the electric power consumed by the device for all biases. This behavior is expected since the Joule heating model only depends on electrical quantities to compute the power dissipation–i.e., the current density and the electric field–, which are gathered directly from the simulation.

The estimation provided by the PEB model, in turn, is in good agreement with the electric power for biases up to $V_D = V_G = -1.0$ V. For higher biases, the model slightly underestimate the power dissipation as heat. This behavior is due to the fact the power dissipation computed via PEB model relies on the carrier-optical phonon relaxation time τ_{c-OP}, which in this thesis, it is an input parameter and considered to be constant for the entire range of biases[6]. A better fitting might be achieved by using energy-dependent carrier-phonon relaxation times, but there is a lack of data in the literature addressing

[6]This is an approximation intrinsic to the model.

Figure 6.26: Electric power supplied to the device (line) and power dissipated as heat (markers) for each thermal model according to the applied bias for 100 case study FD-SOI MOSFETs considering $T_{\text{gate}} = 300$ K. Error bars stand for the 95% confidence bounds.



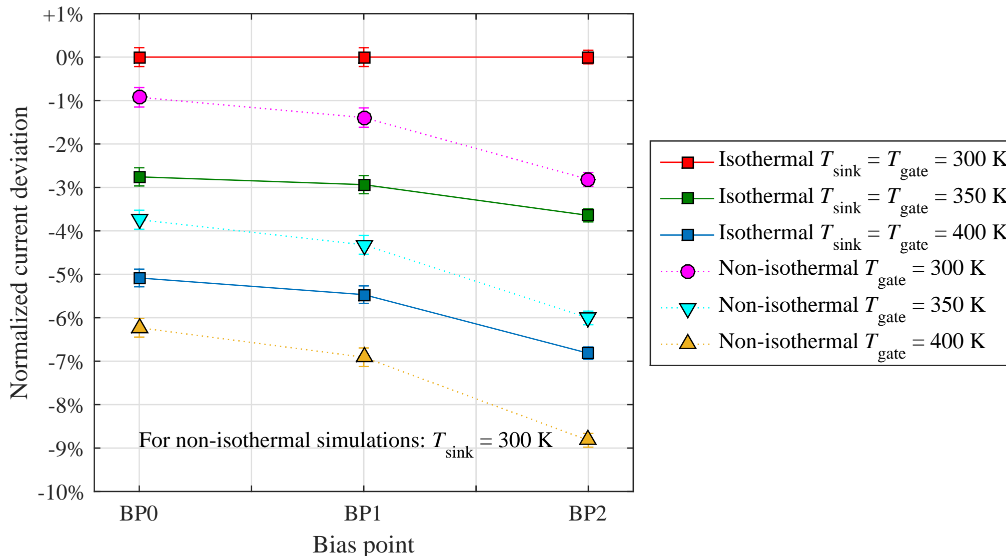
Source: author.

such dependence. The discussion presented in this subsection is also applicable to the bulk case study device simulation (see Figure 6.12).

6.2.4 Impact of Self-heating on the FD-SOI MOSFET On-Current

The impact of self-heating effects on the device current capability was also extracted for FD-SOI MOSFETs. In this regard, a set of 1000 case study transistors was simulated for the isothermal condition at three different applied voltages, i.e., $V_{\text{D}} = V_{\text{G}} = -1.0$ V, $V_{\text{D}} = -1.5$ V and $V_{\text{G}} = -1.0$ V, and $V_{\text{D}} = V_{\text{G}} = -1.5$ V, with $T_{\text{gate}} = T_{\text{sink}} = 300$ K, providing a reference current distribution I_{D0} for each bias. Then, non-isothermal simulations were performed for the same ensemble of devices considering $T_{\text{sink}} = 300$ K but varying the gate temperature T_{gate}, aiming to extract the current degradation dependence on the gate performance as a heat sink. These simulations were carried out considering $T_{\text{gate}} = 300$ K, $T_{\text{gate}} = 350$ K, and $T_{\text{gate}} = 400$ K. A respective current distribution I_{D} and its average value $\overline{I_{\text{D}}}$ were extracted for each case, allowing one to compute the average current degradation $\Delta\overline{I_D}$, i.e., $\Delta\overline{I_D} = \overline{I_{\text{D}}} - \overline{I_{\text{D0}}}$. The normalized current degradation for isothermal and non-isothermal simulations is depicted in Figure 6.27 for each bias point and

Figure 6.27: Average drain current deviation for isothermal and non-isothermal simulations according to the applied biases BP0, BP1, and BP2, which correspond to $V_D = -1.0$ V and $V_G = -1.0$ V, $V_D = -1.5$ V and $V_G = -1.0$ V, and $V_D = -1.5$ V and $V_G = -1.5$ V, respectively. The values are normalized by the average current extracted from isothermal simulations with $T_{sink} = T_{gate} = 300$ K. Non-isothermal simulations were carried out with $T_{sink} = 300$ K. Error bars stand for the 95% confidence bounds.



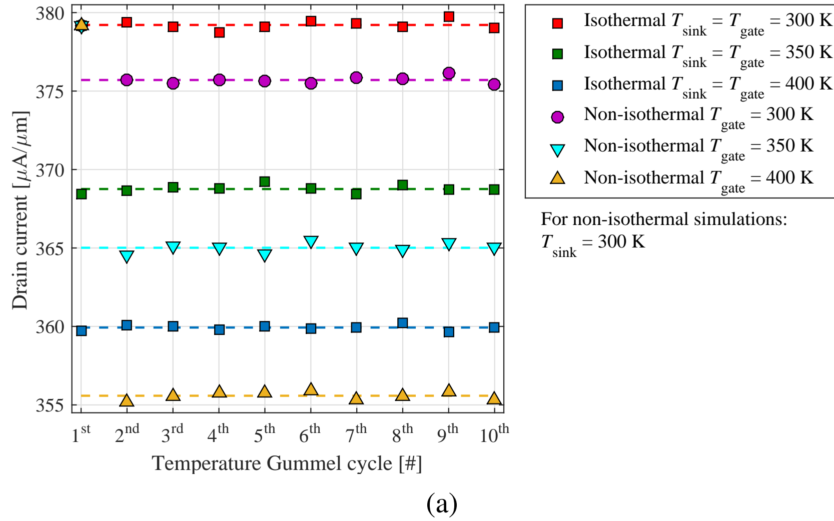
Source: author.

multiple heat sink and gate temperatures. Figure 6.28, in turn, presents the device current as a function of the temperature Gummel cycle for the same conditions mentioned above. Isothermal currents for $T_{gate} = T_{sink} = 350$ K and $T_{gate} = T_{sink} = 400$ K are also presented for comparison. These results are summarized in Table 6.2.

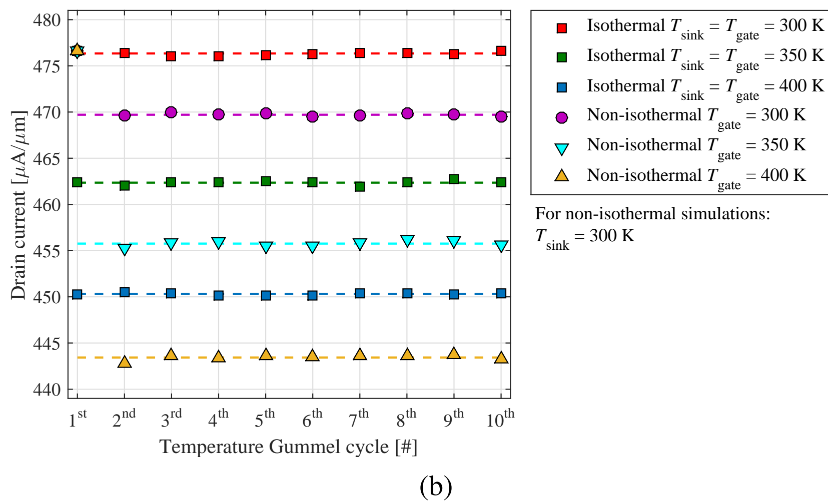
Table 6.2: Average drain current for isothermal and non-isothermal simulations as a function of the gate temperature and applied bias. The current variation for each case is calculated over the current extracted from the isothermal simulation with $T_{gate} = T_{sink} = 300$ K. Non-isothermal simulations were performed with $T_{sink} = 300$ K.

Simulation Type	Average drain current [μA/μm] vs Bias					
	$V_G = -1.0$ V				$V_G = -1.5$ V	
	$V_D = -1.0$ V		$V_D = -1.5$ V			
Iso $T_{gate} = T_{sink} = 300$ K	379.21	—	476.34	—	826.26	—
Iso $T_{gate} = T_{sink} = 350$ K	368.76	−2.76%	462.36	−2.93%	796.16	−3.64%
Iso $T_{gate} = T_{sink} = 400$ K	359.93	−5.08%	450.30	−5.47%	769.93	−6.72%
Non-iso $T_{gate} = 300$ K	375.70	−0.93%	469.71	−1.39%	803.01	−2.81%
Non-iso $T_{gate} = 350$ K	365.01	−3.74%	455.76	−4.32%	776.65	−6.00%
Non-iso $T_{gate} = 400$ K	355.58	−6.23%	443.43	−6.91%	753.38	−8.82%

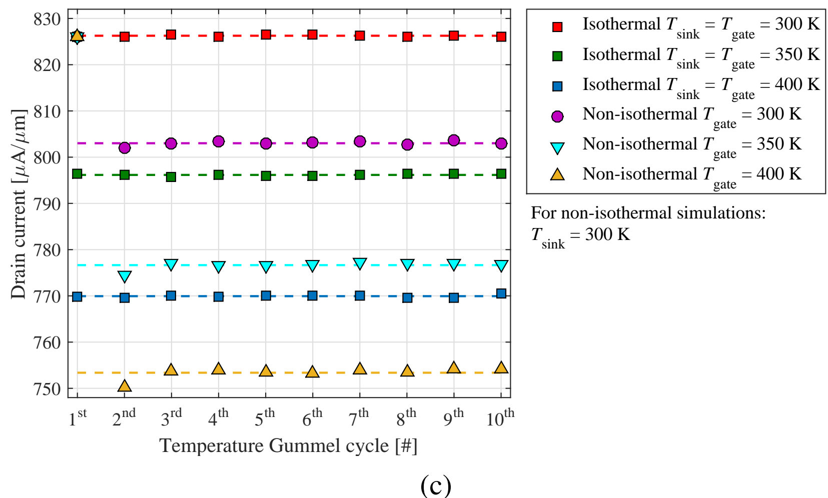
Figure 6.28: Drain current as a function of the temperature Gummel cycle (i.e., time) for isothermal and non-isothermal simulations for multiple heat sink and gate temperatures at $V_D = -1.0$ V and $V_G = -1.0$ V (a), $V_D = -1.5$ V and $V_G = -1.0$ V (b), and $V_D = -1.5$ V and $V_G = -1.5$ V (c). Dashed lines represent the average value for each case.



(a)



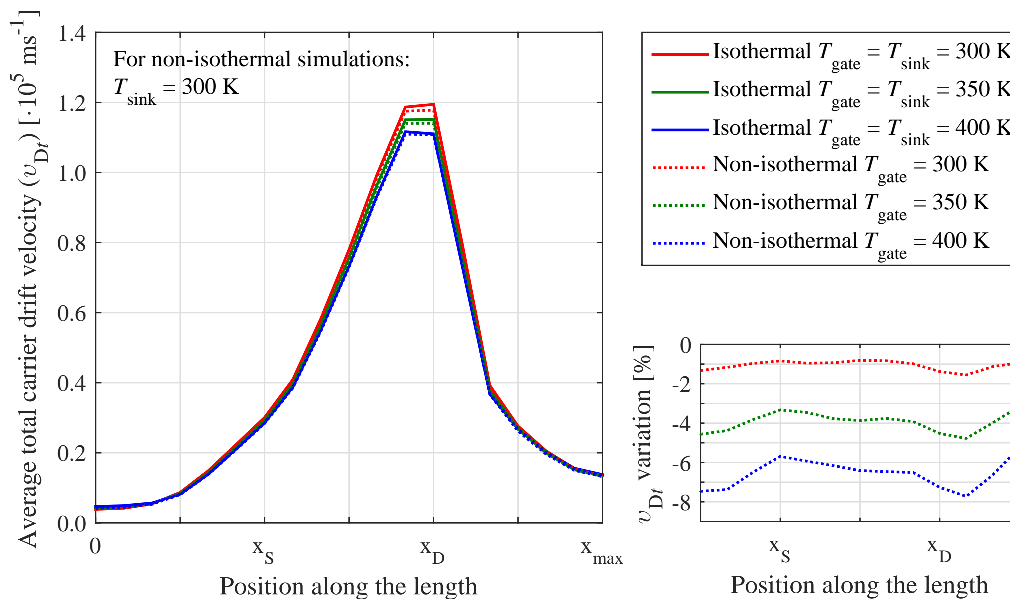
(b)



(c)

Source: author.

Figure 6.29: *Left*: average total carrier drift velocity v_{Dt} for isothermal (lines) and non-isothermal (dashed lines). *Bottom right*: total carrier drift velocity variation over its isothermal value ($T_{gate} = T_{sink} = 300$ K) along the transistor length as a function of the gate temperature for non-isothermal simulations. These plots were extracted for 100 case study FD-SOI MOSFETs at $V_D = V_G = -1.0$ V and $T_{gate} = 300$ K.



Source: author.

It is clear from the results that isothermal simulations carried out at a given gate temperature notoriously overestimate the device current when compared to the non-isothermal analysis performed for the same circumstances. In addition, considering both the gate and the back contact at room temperature with an applied voltage of $V_D = V_G = -1.0$ V, the current degradation due to self-heating is less than 1%, but it can be as high as 6.23% by assuming $T_{gate} = 400$ K. If the applied bias is increased up to $V_D = V_G = -1.5$ V, the current degradation is even larger: 8.82% considering $T_{gate} = 400$ K.

In the literature, however, 2D electro-thermal simulations of *n*-type FD-SOI devices via PEB model predict a current degradation due to self-heating in the order of 10% for a transistor under similar geometry, gate temperature, and bias conditions (i.e., $V_D = V_G = -1.2$ V) (RALEVA et al., 2008; RALEVA et al., 2012). Moreover, the hot spot is reported to exceed 500 K. On the other hand, note that the current capability of Raleva's case study device is roughly 4 times higher than the current capability of the structure under analysis here and, thus, so is the power dissipation. In this way, a larger current degradation is obviously expected for that case.

As presented in Chapter 3, the increase on the transistor temperature intensifies the carrier-phonon scattering rate. The enhancement of the scattering rate degrades the

carrier mobility, which, in turn, is directly dependent on the carrier drift velocity. In this context, Figure 6.29 depicts the total carrier drift velocity υ_{Dt} for isothermal and non-isothermal simulations carried out at $V_D = V_G = -1.0$ V. As expected, the carrier drift velocity reduces as the temperature increases for both isothermal and non-isothermal cases. Figure 6.29 also presents the variation on the total carrier drift velocity along the transistor length as a function of the gate temperature for the non-isothermal cases. Note that the total carrier drift velocity variation is, on average, in the order of the current degradation presented in Table 6.2 for each respective case. Hence, a peak in the drift velocity degradation can be seen at the drain region of the transistor, at the point where the heat generation is the highest.

6.3 Particle-based Simulations of Charge Traps in MOSFETs

The particle-based characteristic of the tool proposed in this thesis is also suitable for the simulation of the impact of charge traps on the performance of *p*-type MOS devices. The real-space treatment of the charge transport and Coulomb interactions allows one to account for the particle-like (discrete) nature of carriers and impurities. Moreover, the placement of the dopant atoms in 3D properly reproduces the RDF effect and also facilitates the generation of unique random dopant profiles. The aforementioned features are key for the analysis of the impact of charge traps in MOSFETs via computational simulations. The RDF effect, for instance, directly affects the trap impact, since, for inhomogeneous doping configurations, charge traps may induce random telegraph noise (RTN) with larger magnitude (MUELLER; SCHULZ, 1998; VANDAMME; SODINI; GINGL, 1998).

In this context, the present section covers the study of the impact of charge traps in MOSFETs via particle-based device simulations. In Section 6.3.1, results pertaining statistical simulations of traps in bulk MOSFETs are presented and discussed. By employing the electro-thermal feature of the tool, Section 6.3.2, in turn, addresses the interplay between self-heating and trap activity for both bulk and FD-SOI MOSFET devices.

6.3.1 Statistical Simulations of Charge Traps in Bulk MOSFETs

In order to extract the current degradation caused by a single charge trap, isothermal simulations were carried out for a bulk case study MOSFET device. The physical dimensions of such a device slightly differ from the structure introduced in Section 5.3. Basically, the transistor has a channel length of 25 nm, a width of 30 nm, and gate oxide thickness of 1.2 nm (a full description of the case study device characteristics is presented in Rossetto et al. (2018)). At first, a set of $N = 100$ devices were simulated in absence of a fixed oxide charge, in order to estimate a reference current distribution I_{D0}. Subsequently, a single charge trap was incorporated into the case study structures. Given the three dimensional character of the simulation, the trap was placed in a position $p_t(x_t, d_t, z_t)$, where x_t, d_t, and z_t stand for the trap position along the transistor length, its depth within the gate oxide, and its position along the transistor width, respectively. Here, x_t ranges from 0 to L_G, and d_t ranges from 0 nm (i.e., Si/SiO_2 interface) to t_{OX}. Along the width, the trap was placed at $z_t = 0.5W$, since in that position the current degradation due to a trap is expected to be the most[7]. The aforementioned simulations were carried at $V_D = V_G = -0.5$ V.

The device current will then depend on the trap position. In this way, for each device i simulated with a trap in a given position $p(x_t, d_t, z_t)$, the impact of the trap $\delta I_{D,i}(x_t, d_t, z_t)$ was extracted as

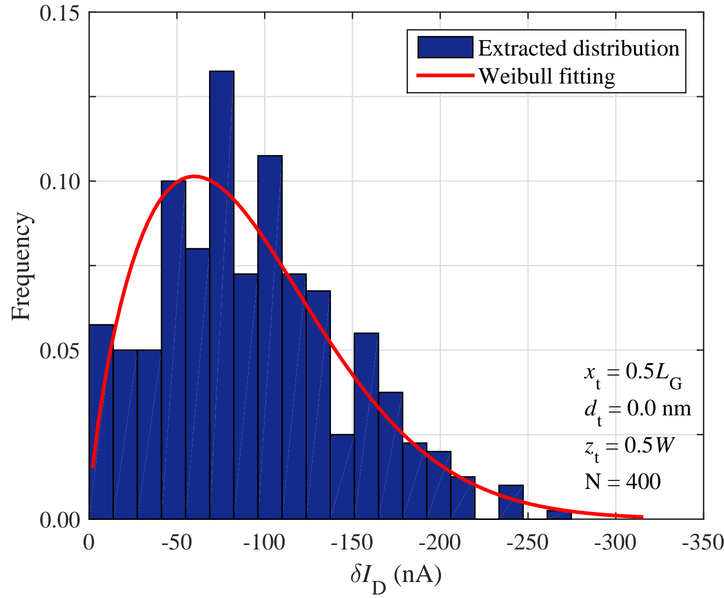
$$\delta I_{D,i}(x_t, d_t, z_t) = I_{D,i}(x_t, d_t, z_t) - I_{D0,i} \ , \tag{6.10}$$

where $I_{D,i}(x_t, d_t, z_t)$ and $I_{D0,i}$ are the time averaged drain currents of device i with and without a trap, respectively. By taking all possible x_t, d_t and z_t, δI_D provides a statistical distribution of trap impacts for the analyzed geometry.

A typical current deviation distribution $\delta I_D(x_t, d_t, z_t)$ extracted for bulk MOSFETs is depicted in Figure 6.30. In this particular case, the trap was placed at $x_t = 0.5L_G$, $d_t = 0.0$ nm, and $z_t = 0.5W$. In addition, the distribution was extracted for $N = 400$ instead of $N = 100$ case study devices, in order to capture better distribution tail statistics. Note that the effect of the random dopant distribution on the trap impact is evidenced by the spread of the current deviations. Even though the charge trap was placed at the same location for all transistors, its impact on the transistor current varies, as the impurity (and carrier) concentration profile slightly differs from device to device. For those tran-

[7] Statistically, the trap impact along the width is expected to play a minor role on the current deviation distribution.

Figure 6.30: Current deviation distribution from $N = 400$ transistors at $V_D = V_G = -0.5$ V with a single charge trap located at the position $x_t = 0.5L_G$, $d_t = 0.0$ nm, and $z_t = 0.5W$.



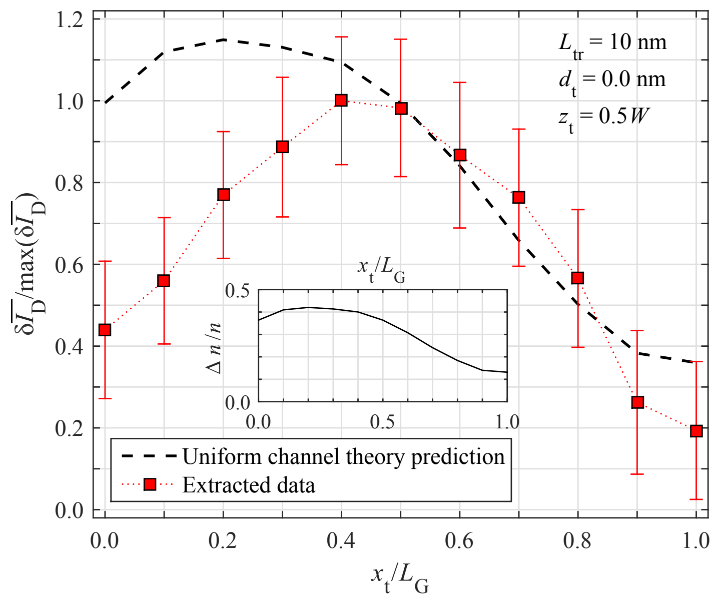
Source: author.

sistors in which the impurity concentration right beneath the charge trap is large, the trap impact is minimized and, thus, smaller current deviations are expected. Conversely, for devices in which the impurity concentration at the same position is reduced, trap impact is maximized and, thus, more pronounced deviations are observed.

Since a charge trap always impacts negatively the transistor current, only negative values of current deviations are expected. In practice, however, positive current deviations were observed for a few simulations due to the uncertainty of the time-averaged drain currents. These anomalous observations are few and close to zero; thus, they do not significantly affect the overall distribution. The fitting of the extracted deviation was performed using a Weibull distribution, even though other single-sided distributions (such as the log-normal) might be possible.

The impact of a trap on the device current was also extracted as a function of the trap position along the channel length. To do that, a single charge trap was placed at $d_t = 0.0$ nm and $z_t = 0.5W$, and its position along the transistor channel was swept from 0 to L_G, in steps of $0.1L_G$. The normalized trap impact along the channel length is depicted in Figure 6.31. Here, it is clear that the influence of the charge trap on the device current varies depending on the trap position, which is expected from number fluctuation theory. Figure 6.31 also depicts the expected current degradation according to the uniform channel theory. It postulates that the deviation of the transistor current due

Figure 6.31: Normalized average drain current deviation $\delta\overline{I_D}$ caused by a single trap as a function of its position along the channel length x_t with $d_t = 0.0$ nm and $z_t = 0.5W$ at $V_D = V_G = -0.5$ V. Each point represents the average value for $N = 100$ transistors, and the error bars stand for the 95% confidence bounds. Theoretical prediction is calculated by Equation 6.11 with $L_{tr} = 10$ nm. *Inset*: Carrier concentration variation $\Delta n/n$ along the transistor length.



Source: author.

to a trap $\delta I_{D,calc}$, assuming that the carrier concentration n directly impacts the transistor channel conductance σ, is given by (SIMOEN et al., 1992)

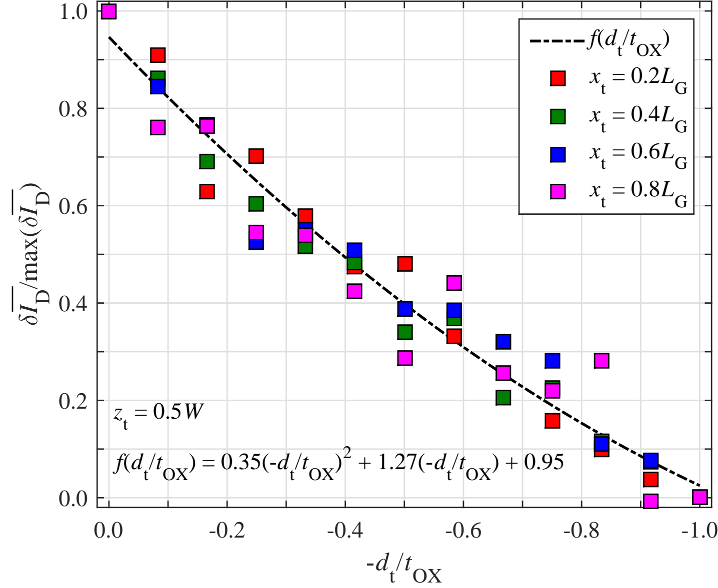
$$\delta I_{D,calc} = \frac{\overline{I_{D0}}}{WL} \cdot L_{tr}^2 \cdot \frac{\Delta n}{n}, \tag{6.11}$$

where L_{tr}^2 stands for the trap impact area[8], and $\Delta n/n$ is the change on the carrier concentration due to a trap (in cm^{-2}). In Figure 6.31, the uniform channel theory prediction was plotted considering $L_{tr} = 10$ nm for whole extension of the transistor channel.

Similarly, the impact of a trap on the transistor current was extracted as a function of the trap distance from the Si/SiO_2 interface. For these simulations, the trap depth into the oxide was swept from 0 nm to t_{OX}, in steps of 0.1 nm. Along the width, the trap was placed at $z_t = 0.5W$, and four different position along the channel were considered, namely $x_t = 0.2L_G$, $x_t = 0.4L_G$, $x_t = 0.6L_G$, and $x_t = 0.8L_G$. The results are presented in Figure 6.32. For each curve, the plot is normalized by the respective impact of a trap at

[8]The trap impact area regards solely to the uniform channel theory prediction, i.e., it is not used in the simulation.

Figure 6.32: Normalized average current deviation $\delta\overline{I_D}$ caused by a single trap as a function of the its depth into the oxide d_t for four distinct positions along the channel length, namely $x_t = 0.2L_G$, $x_t = 0.4L_G$, $x_t = 0.6L_G$, and $x_t = 0.8L_G$, and with $z_t = 0.5W$ at $V_D = V_G = -0.5$ V. Each point represents the average value for $N = 100$ transistors. Error bars are not presented for the sake of graph clarity.



Source: author.

the interface. Note that, as the absolute value of d_t increases, trap impact decreases in a quasi-linear fashion. This behavior can be related to the trap impact on the device carrier concentration. Simulations have demonstrated that as the trap position is swept deeper into the gate oxide, the carrier concentration beneath the trap recovers linearly from its reference (without a trap) value, following the theoretical prediction by Jindal and Ziel (1978). From Figure 6.32, the trap impact dependence on the trap depth into de oxide can be fitted to the expression

$$f(d_t/t_{OX}) = 0.35(-d_t/t_{OX})^2 + 1.27(-d_t/t_{OX}) + 0.95. \quad (6.12)$$

Owing to the charge trap impact dependence on the trap position along the channel length and on its depth into the oxide for a few discrete points, one can extrapolate a distribution for δI_D for any point $p_t(x_t, d_t, z_t = 0.5W)$ within the structure as[9]

$$\delta I_{D,i} = \delta I_D(X_i) f(Y_i), \quad (6.13)$$

[9]Under the assumption that $\delta I_D(X)$ and $f(Y)$ are statistically independent, which is the case.

where

$$X_i = \frac{x_t}{L_G}, \tag{6.14}$$

and

$$Y_i = -\frac{d_t}{t_{OX}}. \tag{6.15}$$

Here, x_t and d_t are uniformly distributed random variables representing the trap position. The distribution $\delta I_D(X)$ stands for the trap impact variation along the channel length, and it also accounts for the impact of the RDF on the current deviation, providing Weibull distributed values. The shape and scale parameters of the Weibull distribution vary according to the position x_t. Thus, the trap impact dependence on its position along the channel length is taken into account by modulating the Weibull distribution with the parameters according to the correspondent trap position. Since such distribution is only extracted for a few discrete positions along the channel length ($x_t = 0$, $x_t = 0.1L_G$, ... , $x_t = L_G$), an interpolation scheme is adopted. The function $f(Y)$, in turn, accounts for the attenuation of the trap influence as it goes deeper into the gate oxide. Note that the function arises from the curve fitting presented in Figure 6.32. This statistical analysis is presented in details in Rossetto et al. (2018).

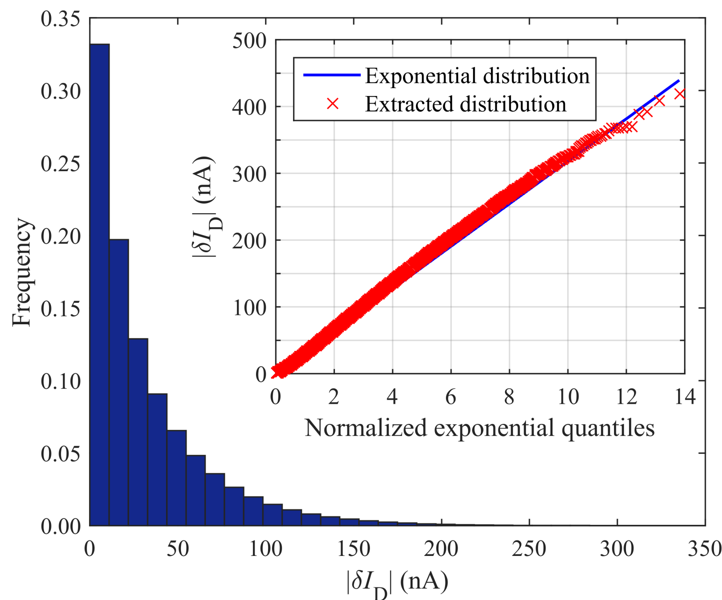
From Equation 6.13, one may estimate the distribution δI_D for any number of traps. For instance, Figure 6.33 presents the distribution of current deviations δI_D for 10^6 traps randomly positioned within the gate oxide. Note that the distribution of current deviation follows an exponential behavior, as theoretically predicted and observed by experimental studies (KACZER et al., 2010; FRANCO et al., 2011; WECKX et al., 2017). This result strongly supports the idea that particle-based simulations are suitable for the study of current degradation due to charge trap activity in MOSFETs.

6.3.2 Electro-thermal Simulations of Charge Traps in Bulk and FD-SOI MOSFETs

The electro-thermal feature of the simulator allows one to study the interplay between self-heating effects and trap activity. In this regard, non-isothermal simulations were carried out for bulk and FD-SOI MOSFETs considering the influence of a single charge trap. These simulations, however, were performed for the case study structures presented in Section 5.3, and the charge trap was placed at $x_t = 0.5L_G$, $d_t = 0.0$ nm, and $z_t = 0.5W$. The results are depicted in Figure 6.34.

For bulk devices, the current degradation due to self-heating increases with the

Figure 6.33: Distribution of drain current deviations estimated using Equation 6.13 considering 10^6 charge traps randomly positioned within the transistor gate oxide.



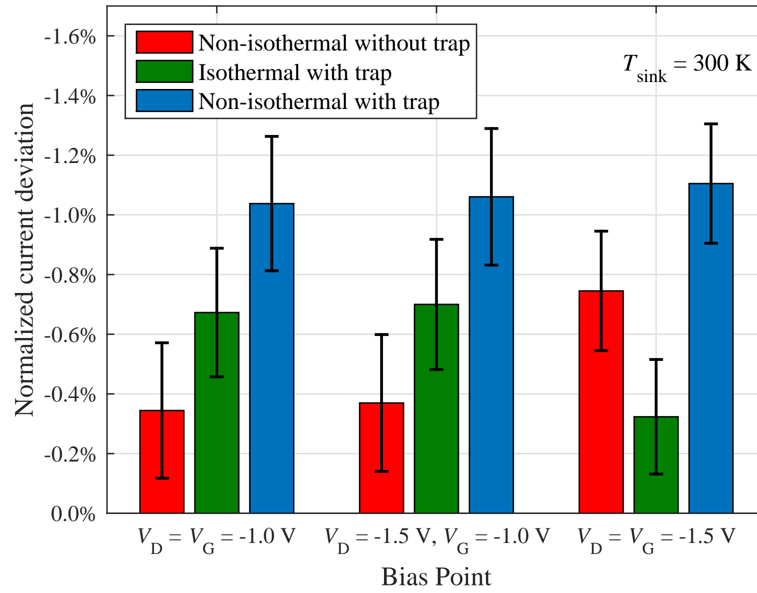
Source: author.

applied bias (TENBROEK et al., 1996). The degradation due to charge traps, on the other hand, does not follow the same behavior. For the same gate voltage, the trap's impact slightly depends on the drain voltage; for the same drain bias, its impact significantly reduces with the gate voltage (BUISSON; GHIBAUDO; BRINI, 1992). The total degradation is apparently composed by the individual degradations due to self-heating and due to charge traps, and it marginally increases with the bias. One should observe that, in a real device, the temperature directly affects the trap's time constants (GRASSER et al., 2010; WIRTH; SILVA; KACZER, 2011). Since the tool proposed here performs only static steady-state simulations, i.e., the trap remains occupied during the entire simulation, such an effect cannot be captured.

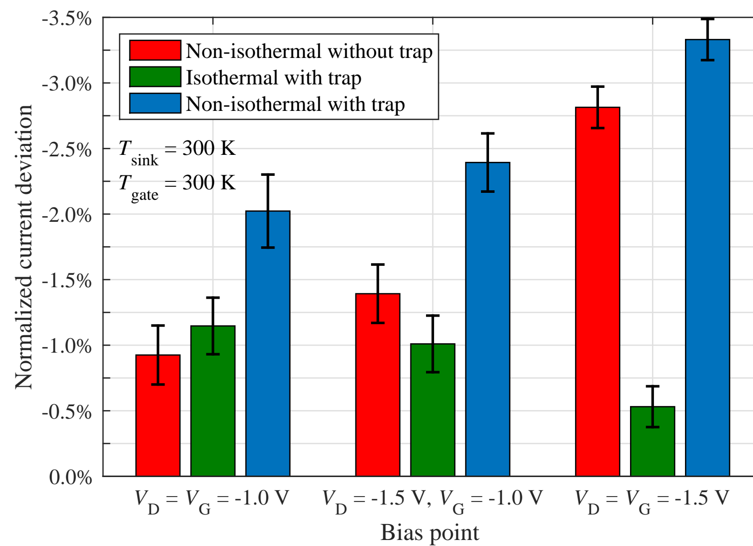
For FD-SOI MOSFETs, the current degradation due to self-heating also increases with the applied bias, but in a more aggressive manner, since these structures suffer from more pronounced hot spots than the bulk ones. Considering the error bars, the degradation due to charge traps follows the same behavior already discussed for bulk devices. Hence, the same cumulative current degradation behavior was observed, being the total degradation composed by the self-heating and charge trap individuals contributions.

One important aspect is that, for the same bias point, the current degradation exclusively due to a trap differs from FD-SOI to bulk transistors, even though the structures share the same substrate doping. Basically, bulk devices exhibit less degradation due to

Figure 6.34: Normalized current degradation due to self-heating and trap activity for bulk (a) and FD-SOI (b) MOSFETs as a function of the applied bias. Error bars stand for the 95% confidence bounds.



(a)



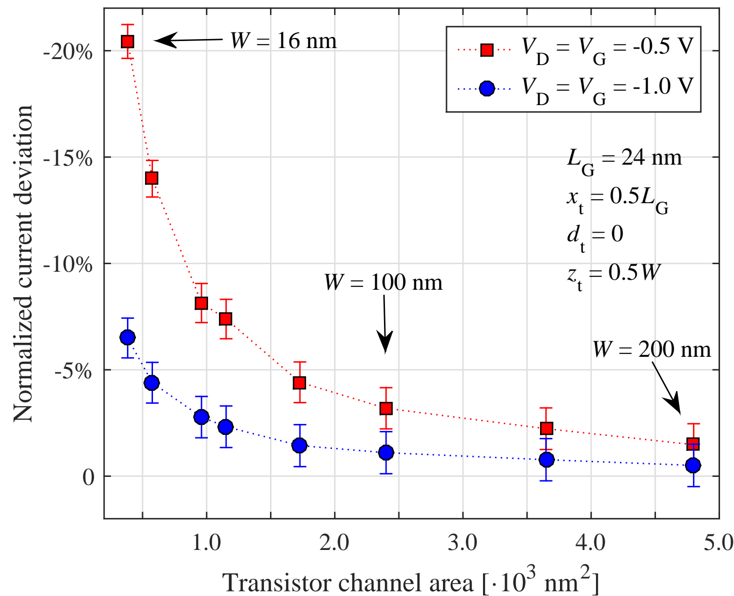
(b)

Source: author.

the smaller oxide thickness (1.2 nm, in comparison with 2.0 nm for FD-SOI); the gate electrode has more control of the conduction channel, due to the larger oxide capacitance. Since FD-SOI devices have lower gate capacitance — due to the larger gate oxide thickness —, more degradation due to traps is expected (SIMOEN et al., 1992).

Analyzing the trap's impact on the device current separately, one can note that it is,

Figure 6.35: Normalized current deviation due to a single trap at the Si/SiO_2 interface as a function of the transistor channel area and applied bias. These plots were extracted from isothermal simulations of 100 case study FD-SOI MOSFETs with $T_{\text{gate}} = T_{\text{sink}} = 300$ K. Error bars stand for 95% confidence bounds.



Source: author.

at most, in the order of −0.7% and −1.2% for bulk and FD-SOI transistors, respectively. This modest degradation might be related to the relatively-large device width, which is as much as 100 nm for both cases. For comparison, the impact of a single trap on the drain current of FD-SOI MOSFETs was extracted as a function of the transistor channel area, and it is depicted in Figure 6.35. Note that the trap impact becomes more significant as the device width is reduced, which is in agreement with literature data (see Asenov et al. (2003, Figure 6), for instance). A 16 nm-wide FD-SOI transistor, for example, may present over 20% of current degradation due to a single charge trap when biased at $V_D = V_G = -0.5$ V, while a 100 nm-wide one presents less than 4%. The current deviation amplitude due to a trap also varies with the applied voltage, as expected (SIMOEN et al., 1992; BUISSON; GHIBAUDO; BRINI, 1992), and it becomes more significant for lower biases (SHI; MIEVILLE; DUTOIT, 1994).

7 CONCLUSION

The CMOS scaling process of the recent decades has turned self-heating into a major reliability issue for deeply-scaled devices. The effect itself arises from the heat generation and heat confinement within the semiconductor structure. In regions with elevated temperature, some physical and electrical characteristics are degraded, affecting the overall device performance. As self-heating changes the charge transport dynamic, the operation of a modern transistor cannot be treated as an isothermal problem. In this way, the temperature is no longer an input variable, but an output parameter that must be computed from the simulation. Therefore, a rigorous simulation tool must be capable to compute the non-isothermal temperature profile, as well as properly account its effect on the device performance.

In this context, this work presented a novel electro-thermal particle-based device simulator that was developed and utilized for the simulation of self-heating effects in bulk and FD-SOI p-type MOS transistors. It self-consistently couples the solution of the carrier BTE via the Ensemble Monte Carlo method with the resolution of the phonon energy balance equations. By adopting temperature-dependent scattering tables, the phonon temperatures are accounted during the charge transport phase, allowing the impact of self-heating on the transistor current capability to be captured. The tool also provides detailed profiles of the acoustic and optical phonon temperatures along the transistor structure, as well as the heat generation picture within the device. Extracted data for general device parameters, phonon temperature, heat generation, and power dissipation profiles of case study structures were found to be qualitatively in agreement with the expected behavior and with literature data, ensuring the proper operation of the simulator.

Electro-thermal simulations performed for bulk devices demonstrated that self-heating plays a minor role in these structures. Due to the relatively high thermal conductivity of the substrate, most part of the heat generated by the transistor can be removed by the heat sink in the bottom of the device. In this way, bulk MOSFETs exhibited very low lattice temperature rises over the heat sink temperature; at most 4 K and 10 K for $V_D = V_G = -1.0$ V and $V_D = V_G = -1.5$ V, respectively, considering $T_{sink} = 300$ K. Such a modest temperature rise translates itself into a modest current degradation, which is, respectively, -0.34% and -0.75% for the biases presented earlier. The optical phonon temperature was found to be higher than the acoustic phonon one, which characterizes the temperature non-equilibrium between the phonon modes.

For FD-SOI devices, in turn, electro-thermal simulations revealed much higher temperatures at the hot spots, as well as all over the device active area. Assuming $T_{\text{gate}} = T_{\text{sink}} = 300$ K, the extracted lattice temperature rise over the heat sink temperature was on the order of 27 K and 69 K for $V_D = V_G = -1.0$ V and $V_D = V_G = -1.5$ V, respectively; but it can be as high as 113 K and 150 K for the respective biases presented earlier when $T_{\text{gate}} = 400$ K. These more pronounced temperatures are the result of the heat confinement within the silicon layer due to the poor thermal dissipation via buried oxide. Moreover, the silicon thermal conductivity lowering due to the layer's reduced thickness also hampers the heat flow to the boundaries, further increasing the hot spot temperatures. Consequently, the current degradation due to self-heating was found to be much more significant for FD-SOI than for bulk devices. Assuming $T_{\text{gate}} = T_{\text{sink}} = 300$ K, it may reach -0.93% and -2.81% for $V_D = V_G = -1.0$ V and $V_D = V_G = -1.5$ V, respectively; whereas for $T_{\text{gate}} = 400$ K, the current degradation is respectively -6.23% and -8.82% for the same bias points. The non-equilibrium between the acoustic and optical phonon modes was also observed for FD-SOI MOSFETs.

The particle-based characteristic of the simulator was also used to study the impact of charge traps in MOSFETs. From statistical simulations performed for a bulk case study device, the impact of a single charge trap was extracted as a function of the trap position along the channel length and its depth within the gate oxide for a few discrete points. These results were used to compose a statistical model, allowing one to estimate the impact of a charge trap randomly-positioned at any point within the gate oxide structure. From such statistical model, estimated current deviations due to random traps were observed to be exponentially distributed, which is in agreement with experimental observations from the literature.

Electro-thermal simulations of charge traps were also performed for the case study structures at multiple biases. For both bulk and FD-SOI devices, the largest current degradation for each bias was observed when the effects of self-heating and trap activity take place simultaneously. In addition, the impact of charge traps on the device current is more pronounced for lower biases, whereas the degradation due to self-heating dominates at larger biases, in agreement with literature predictions. For a case study transistor with $L_G = 25$ nm and $W = 100$ nm, $T_{\text{gate}} = T_{\text{sink}} = 300$ K, and biased at $V_D = V_G = -1.5$ V, the total current degradation (i.e., self-heating plus trap activity) reached -1.11% and -3.33% for bulk and FD-SOI MOSFETs, respectively. The trap-induced current degradation also showed to depend on the gate oxide thickness and the channel area, as expected.

FD-SOI devices, which possess larger oxide thickness, presented more pronounced current deviations due to a single trap than bulk ones for the entire range of biases. Similarly, the trap impact was found to be more severe as the device channel area is reduced. Simulations of 16 nm-wide 25 nm-long FD-SOI transistors carried out at $V_D = V_G = -0.5$ V presented over 20% of current degradation due to a single charge trap, while 100 nm-wide ones presented less than 4% at the same conditions.

Future work will focus on (I) performing relevant case study simulations of self-heating and trap activity — and the interplay between them — with the software that is already established; and (II) expanding the capabilities of the simulation tool, allowing one to study even more relevant transistor operation scenarios and more complex device geometries. Regarding to I, additional simulations will be carried out for both short- and long-channel transistors to characterize the trap impact as a function of the bias, doping, channel area and width, gate oxide material, etc..., evaluating for each case the trap-induced variation on the surface potential, surface concentration, and trap cross-sectional area. In II, the simulator will be modified to deal with non-planar technologies, allowing one to simulate self-heating and trap activity effects in state-of-the-art Tri-gate and FinFET devices. From the process point of view, the transistor structure under consideration will also be improved by adding non-uniform doping profiles, gate oxide stacks, and pocket implants.

REFERENCES

ADLER, M. S. Accurate calculations of the forward drop and power dissipation in thyristors. **IEEE Transactions on Electron Devices**, [s.l.], v. 25, n. 1, p. 16–22, 1978.

ANDERSON, C. L.; CROWELL, C. R. Threshold energies for electron-hole pair production by impact ionization in semiconductors. **Physical Review B**, College Park, U.S.A.: American Physical Society, v. 5, n. 6, p. 2267–2272, 1972.

ARNOLD, E.; PEIN, H.; HERKO, S. P. Comparison of self-heating effects in bulk-silicon and SOI high-voltage devices. In: **Proceedings of the IEEE International Electron Devices Meeting**. San Francisco, U.S.A.: [s.n.], 1994. p. 813–816.

ASENOV, A. et al. RTS amplitudes in decananometer MOSFETs: 3-D simulation study. **IEEE Transactions on Electron Devices**, [s.l.], v. 50, n. 3, p. 839–845, 2003.

ASHEGHI, M. et al. Thermal conduction in doped single-crystal silicon films. **Journal of Applied Physics**, [s.l.], v. 91, n. 8, p. 5079–5088, 2002.

ASHEGHI, M. et al. Phonon-boundary scattering in thin silicon layers. **Applied Physics Letters**, [s.l.], v. 71, n. 13, p. 1798–1800, 1997.

ASHEGHI, M. et al. Temperature-dependent thermal conductivity of single-crystal silicon layers in SOI substrates. **Journal of Heat Transfer**, New York, U.S.A.: American Society of Mechanical Engineers, v. 120, n. 1, p. 30–36, 1998.

BARLOW, P. S.; DAVIS, R. G.; LAZARUS, M. J. Negative differential output conductance of self heated power MOSFETs. **IEEE Proceedings I — Solid-State and Electron Devices**, [s.l.], v. 133, n. 5, p. 177–179, 1986.

BEECHEM, T. et al. Contribution of optical phonons to thermal boundary conductance. **Applied Physics Letters**, [s.l.], v. 97, n. 6, p. 061907–1–061907–3, 2010.

BERGER, M.; CHAI, Z. Estimation of heat transfer in SOI-MOSFETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 38, n. 4, p. 871–875, 1991.

BHAVNAGARWALA, A. J.; TANG, X.; MEINDL, J. D. The impact of intrinsic device fluctuations on CMOS SRAM cell stability. **IEEE Journal of Solid-State Circuits**, [s.l.], v. 36, n. 4, p. 658–665, 2001.

BINDER, K. **The Monte Carlo Method in Condensed Matter Physics**. 2nd. ed. Berlin, Germany: Springer-Verlag, 1995. 420 p.

BRANDT, A. Multi-level adaptive technique (MLAT) for fast numerical solution to boundary value problems. In: ____. **Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics**. Berlin, Germany: Springer Berlin Heidelberg, 1973. p. 82–89.

BUISSON, O. R.; GHIBAUDO, G.; BRINI, J. Model for drain current RTS amplitude in small-area MOS transistors. **Solid-State Electronics**, [s.l.], v. 35, n. 9, p. 1273–1276, 1992.

CALLAWAY, J. Model for lattice thermal conductivity at low temperatures. **Physical Review**, College Park, U.S.A.: American Physical Society, v. 113, n. 4, p. 1046–1051, 1959.

CAMARGO, V. V. A. **Evaluating the impact of charge traps on MOSFETs and circuits**. 76 p. Tese (Doutorado) — Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, 2016.

CANALI, C.; OTTAVIANI, G.; QUARANTA, A. A. Drift velocity of electrons and holes and associated anisotropic effects in silicon. **Journal of Physics and Chemistry of Solids**, [s.l.], v. 32, n. 8, p. 1707–1720, 1971.

CARRUTHERS, J. A. et al. The thermal conductivity of germanium and silicon between 2 and 300 degrees K. **Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**, London, U.K.: The Royal Society, v. 238, n. 1215, p. 502–514, 1957.

CASHWELL, E. D.; EVERETT, C. J. **A Practical Manual on the Monte Carlo Method for Random Walk Problems**. London, U.K.: Pergamon, 1959. 164 p.

CAVIGLIA, A. L.; ILIADIS, A. A. Linear dynamic self-heating in SOI MOSFETs. **IEEE Electron Device Letters**, [s.l.], v. 14, n. 3, p. 133–135, 1993.

CHAI, J. C.; LEE, H. S.; PATNAKAR, S. V. Finite volume method for radiation heat transfer. **Journal of Thermophysics and Heat Transfer**, [s.l.], v. 8, n. 3, p. 419–425, 1994.

CHENG, A. H.-D.; CHENG, D. T. Heritage and early history of the boundary element method. **Engineering Analysis with Boundary Elements**, [s.l.], v. 29, n. 3, p. 268–302, 2005.

CHRYSSAFIS, A.; LOVE, W. A computer-aided analysis of one-dimensional thermal transients in n-p-n power transistors. **Solid-State Electronics**, [s.l.], v. 22, n. 3, p. 249–256, 1979.

DALLMANN, D. A.; SHENAI, K. Scaling constraints imposed by self-heating in submicron SOI MOSFETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 42, n. 3, p. 489–496, 1995.

DENNISON, R. T.; WALTER, K. M. Local thermal effects in high performance bipolar devices/circuits. In: **Proceedings of the Bipolar Circuits and Technology Meeting**. Minneapolis, U.S.A.: [s.n.], 1989. p. 164–167.

DOLL, J.; FREEMAN, D. L. Monte Carlo methods in chemistry. **IEEE Computational Science and Engineering**, [s.l.], v. 1, n. 1, p. 22–32, 1994.

EBINA, A. et al. Ultra low-power CMOS IC using partially-depleted SOI technology. In: **Proceedings of the IEEE Custom Integrated Circuits Conference**. Orlando, U.S.A.: [s.n.], 2000. p. 57–60.

FERRY, D. K. First-order optical and intervalley scattering in semiconductors. **Physical Review B**, College Park, U.S.A.: American Physical Society, v. 14, n. 4, p. 1605–1609, 1976.

FERRY, D. K. **Semiconductor Transport**. 1st. ed. London, U.K.: Taylor & Francis, Inc., 2000. 368 p.

FISCHETTI, M. V.; LAUX, S. E. Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects. **Physical Review B**, College Park, U.S.A.: American Physical Society, v. 38, n. 14, p. 9721–9745, 1988.

FRANCO, J. et al. On the impact of the Si passivation layer thickness on the NBTI of nanoscaled $Si_{0.45}Ge_{0.55}$ pMOSFETs. **Microelectronic Engineering**, [s.l.], v. 88, n. 7, p. 1388–1391, 2011.

FULKERSON, W. et al. Thermal conductivity, electrical resistivity, and Seebeck coefficient of silicon from 100 to 1300 K. **Physical Review**, College Park, U.S.A.: American Physical Society, v. 167, n. 3, p. 765–782, 1968.

FUSHINOBU, K.; MAJUMDAR, A.; HIJIKATA, K. Heat generation and transport in submicron semiconductor devices. **Journal of Heat Transfer**, New York, U.S.A.: American Society of Mechanical Engineers, v. 117, n. 1, p. 25–31, 1995.

GAUR, S. P.; NAVON, D. H. Two-dimensional carrier flow in a transistor structure under nonisothermal conditions. **IEEE Transactions on Electron Devices**, [s.l.], v. 23, n. 1, p. 50–57, 1976.

GERRER, L. et al. Impact of random dopant fluctuations on trap-assisted tunnelling in nanoscale MOSFETs. **Microelectronics Reliability**, [s.l.], v. 52, n. 9, p. 1918–1923, 2012.

GLASSBRENNER, C. J.; SLACK, G. A. Thermal conductivity of silicon and germanium from 3 K to the melting point. **Physical Review**, College Park, U.S.A.: American Physical Society, v. 134, n. 4A, p. A1058–A1069, 1964.

GLASSERMAN, P. **The Monte Carlo Methods in Financial Engineering**. 1st. ed. New York, U.S.A.: Springer Science+Business Media, 2003. 596 p.

GRASSER, T. et al. The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Anaheim, U.S.A.: [s.n.], 2010. p. 16–25.

GROSS, W. J. **Three-dimensional particle-based simulations of deep sub-micron MOSFET devices**. 146 p. Tese (Doutorado) — Arizona State University, Tempe, U.S.A., 1999.

GROSS, W. J.; VASILESKA, D.; FERRY, D. K. 3D simulations of ultra-small MOSFETs with real-space treatment of the electron-electron and electron-ion interactions. **VLSI Design**, [s.l.], v. 10, n. 4, p. 437–452, 2000.

HATAKEYAMA, T.; FUSHINOBU, K. Electro-thermal behavior of a sub-micrometer bulk CMOS device: modeling of heat generation and prediction of temperatures. **Heat Transfer Engineering**, [s.l.], v. 29, n. 2, p. 120–133, 2008.

HATHWAR, R. **Generalized Monte Carlo tool for investigating low-field and high field properties of materials using non-parabolic band structure model**. 66 p. Dissertação (Mestrado) — Arizona State University, Tempe, U.S.A., 2011.

HE, X. **Two-dimensional Monte Carlo particle-based simulations of ultra-small MOSFETs**. 84 p. Dissertação (Mestrado) — Arizona State University, Tempe, U.S.A., 2000.

HESS, K. **Monte Carlo Device Simulation: Full Band and Beyond**. 1st. ed. Boston, U.S.A.: Kluwer Academic Publishers, 1991. 320 p.

HESTENES, M. R.; STIEFEL, E. Methods of conjugate gradients for solving linear systems. **Journal of Research of the National Bureau of Standards**, [s.l.], v. 49, n. 6, p. 409–436, 1952.

HOCKNEY, R. W.; EASTWOOD, J. W. **Computer Simulation Using Particles**. New York, U.S.A.: Taylor & Francis, 1988. 523 p.

HOLLAND, M. G. Analysis of lattice thermal conductivity. **Physical Review**, College Park, U.S.A.: American Physical Society, v. 132, n. 6, p. 2461–2471, 1963.

IOFFE. **Silicon electrical properties**. 2018. Disponível em: <http://www.ioffe.ru/SVA/NSM/Semicond/Si/electric.html>. Acesso em: 25/07/2018.

JACOBONI, C. et al. A review of some charge transport properties of silicon. **Solid-State Electronics**, [s.l.], v. 20, n. 2, p. 77–89, 1977.

JACOBONI, C.; LUGLI, P. **The Monte Carlo Method for Semiconductor Device Simulation**. 1st. ed. Vienna, Austria: Springer-Verlag, 1989. 356 p.

JACOBONI, C.; REGGIANI, L. The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. **Review of Modern Physics**, [s.l.], v. 55, n. 3, p. 645–705, 1983.

JENKINS, K. A.; RIM, K. Measurement of the effect of self-heating in strained-silicon MOSFETs. **IEEE Electron Device Letters**, [s.l.], v. 23, n. 6, p. 360–362, 2002.

JEON, J.; JHON, H.; KANG, M. Investigation of electrothermal behaviors of 5-nm bulk FinFET. **IEEE Transactions on Electron Devices**, [s.l.], v. 64, n. 12, p. 5284–5287, 2017.

JIANG, H. et al. Investigation of self-heating effect on hot carrier degradation in multiple-fin SOI FinFETs. **IEEE Electron Device Letters**, [s.l.], v. 36, n. 12, p. 1258–1260, 2015.

JINDAL, R.; ZIEL, A. van der. Carrier fluctuation noise in a MOSFET channel due to traps in the oxide. **Solid-State Electronics**, [s.l.], v. 21, n. 6, p. 901–903, 1978.

JU, Y. S. Phonon heat transport in silicon nanostructures. **Applied Physics Letters**, [s.l.], v. 87, n. 15, p. 153106–1–153106–3, 2005.

KACZER, B. et al. Atomistic approach to variability of bias-temperature instability in circuit simulations. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Monterey, U.S.A.: [s.n.], 2011. p. XT.3.1–XT.3.5.

KACZER, B. et al. Statistics of multiple trapped charges in the gate oxide of deeply scaled MOSFET devices–Application to NBTI. **IEEE Electron Device Letters**, [s.l.], v. 31, n. 5, p. 411–413, 2010.

KALOS, M. H.; WHITLOCK, P. A. **Monte Carlo Methods**. 2nd. ed. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co., 2008. 203 p.

KAMAKURA, Y. et al. Coupled Monte Carlo simulation of transient electron-phonon transport in small FETs. In: **Proceedings of the IEEE International Electron Devices Meeting**. San Francisco, U.S.A.: [s.n.], 2014. p. 7.3.1–7.3.4.

KAMAKURA, Y. et al. Coupled Monte Carlo simulation of transient electron-phonon transport in nanoscale devices. In: **Proceedings of the International Conference on Simulation of Semiconductor Processes and Devices**. Bologna, Italy: [s.n.], 2010. p. 89–92.

KUROSAWA, T. Monte Carlo calculation of hot electron problems. In: **Proceedings of the International Conference on the Physics of Semiconductors**. Kyoto, Japan: [s.n.], 1966. v. 21, p. 424–426.

LACROIX, D.; JOULAIN, K.; LEMONNIER, D. Monte Carlo transient phonon transport in silicon and germanium at nanoscales. **Physical Review B**, College Park, U.S.A.: American Physical Society, v. 72, n. 6, p. 064305–1–064305–11, 2005.

LAI, J.; MAJUMDAR, A. Concurrent thermal and electrical modeling of sub-micrometer silicon devices. **Journal of Applied Physics**, [s.l.], v. 79, n. 9, p. 7353–7361, 1996.

LAUX, S. E.; FISCHETTI, M. V. Monte Carlo study of velocity overshoot in switching a 0.1-micron CMOS inverter. In: **Proceedings of the IEEE International Electron Devices Meeting**. San Francisco, U.S.A.: [s.n.], 1997. p. 877–880.

LEE, C. C.; PALISOC, A. L.; MIN, Y. J. Thermal analysis of integrated circuit devices and packages. **IEEE Transactions on Components, Hybrids, and Manufacturing Technology**, [s.l.], v. 12, n. 4, p. 701–709, 1989.

LEE, Y. H. et al. Circuit-based reliability consideration in FinFET technology. In: **Proceedings of the IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits**. Chengdu, China: [s.n.], 2017. p. 1–7.

LETURCQ, P. et al. A new approach to thermal analysis of power devices. **IEEE Transactions on Electron Devices**, [s.l.], v. 34, n. 5, p. 1147–1156, 1987.

LINDEFELT, U. Heat generation in semiconductor devices. **Journal of Applied Physics**, [s.l.], v. 75, n. 2, p. 942–957, 1994.

LIU, S. E. et al. Self-heating effect in FinFETs and its impact on devices reliability characterization. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Waikoloa, U.S.A.: [s.n.], 2014. p. 4A.4.1–4A.4.4.

LUNDSTROM, M. **Fundamentals of Carrier Transport**. 2nd. ed. Cambridge, U.K.: Cambridge University Press, 2000. 440 p.

MAHMOODI, H.; MUKHOPADHYAY, S.; ROY, K. Estimation of delay variations due to random-dopant fluctuations in nanoscale CMOS circuits. **IEEE Journal of Solid-State Circuits**, [s.l.], v. 40, n. 9, p. 1787–1796, 2005.

MAKABE, M.; KUBOTA, T.; KITANO, T. Bias-temperature degradation of pMOSFETs: mechanism and suppression. In: **Proceedings of the IEEE International Reliability Physics Symposium**. San Jose, U.S.A.: [s.n.], 2000. p. 205–209.

MAKOVEJEV, S.; OLSEN, S.; RASKIN, J. P. RF extraction of self-heating effects in FinFETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 58, n. 10, p. 3335–3341, 2011.

MANLY, B. F. J. **Randomization, Bootstrap and Monte Carlo Methods in Biology**. 3rd. ed. Boca Raton, U.S.A.: Chapman & Hall/CRC, 2007. 461 p.

MAUTRY, P. G.; TRAGER, J. Investigation of self-heating in VLSI and ULSI MOSFETs. In: **Proceedings of the International Conference on Microelectronic Test Structures**. San Diego, U.S.A.: [s.n.], 1990. p. 221–226.

MAYCOCK, P. Thermal conductivity of silicon, germanium, III–V compounds and III–V alloys. **Solid-State Electronics**, [s.l.], v. 10, n. 3, p. 161–168, 1967.

MITTAL, A.; MAZUMDER, S. Monte Carlo study of phonon heat conduction in silicon thin films including contributions of optical phonons. **Journal of Heat Transfer**, New York, U.S.A.: American Society of Mechanical Engineers, v. 132, n. 5, p. 052402–1–052402–11, 2010.

MOLZER, W. et al. Self heating simulation of multi-gate FETs. In: **European Solid-State Device Research Conference**. Montreux, Switzerland: [s.n.], 2006. p. 311–314.

MUELLER, H. H.; SCHULZ, M. Random telegraph signal: an atomic probe of the local current in field-effect transistors. **Journal of Applied Physics**, [s.l.], v. 83, n. 3, p. 1734–1741, 1998.

NARUMANCHI, S. V. J.; MURTHY, J. Y.; AMON, C. H. Submicron heat transport model in silicon accounting for phonon dispersion and polarization. **Transactions of the ASME-C-Journal of Heat Transfer**, New York, U.S.A.: American Society of Mechanical Engineers, v. 126, n. 6, p. 946–955, 2004.

NEEL, O. L.; HAOND, M. Electrical transient study of negative resistance in SOI MOS transistors. **Electronics Letters**, [s.l.], v. 26, n. 1, p. 73–74, 1990.

NGHIEM, T. T. T.; SAINT-MARTIN, J.; DOLLFUS, P. New insights into self-heating in double-gate transistors by solving Boltzmann transport equations. **Journal of Applied Physics**, [s.l.], v. 116, n. 7, p. 074514–1–074514–9, 2014.

NI, C. et al. Coupled electro-thermal simulation of MOSFETs. **Journal of Computational Electronics**, [s.l.], v. 11, n. 1, p. 93–105, 2012.

ORTIZ-CONDE, A. et al. A review of recent MOSFET threshold voltage extraction methods. **Microelectronics Reliability**, [s.l.], v. 42, n. 4, p. 583–596, 2002.

PALANKOVSKI, V.; SCHULTHEIS, R.; SELBERHERR, S. Simulation of power heterojunction bipolar transistors on gallium arsenide. **IEEE Transactions on Electron Devices**, [s.l.], v. 48, n. 6, p. 1264–1269, 2001.

PELLOIE, J.-L. SOI for low-power low-voltage — Bulk versus SOI. **Microelectronic Engineering**, [s.l.], v. 39, n. 1, p. 155–166, 1997.

POP, E. Energy dissipation and transport in nanoscale devices. **Nano Research**, [s.l.], v. 3, n. 3, p. 147–169, 2010.

POP, E. et al. Localized heating effects and scaling of sub-0.18 micron CMOS devices. In: **Proceedings of the IEEE International Electron Devices Meeting**. Washington, U.S.A.: [s.n.], 2001. p. 31.1.1–31.1.4.

POP, E.; DUTTON, R.; GOODSON, K. Thermal analysis of ultra-thin body device scaling SOI and FinFET devices. In: **Proceedings of the IEEE International Electron Devices Meeting**. Washington, U.S.A.: [s.n.], 2003. p. 36.6.1–36.6.4.

POP, E.; DUTTON, R. W.; GOODSON, K. E. Analytic band Monte Carlo model for electron transport in Si including acoustic and optical phonon dispersion. **Journal of Applied Physics**, [s.l.], v. 96, n. 9, p. 4998–5005, 2004.

POP, E.; SINHA, S.; GOODSON, K. E. Heat generation and transport in nanometer-scale transistors. **Proceedings of the IEEE**, [s.l.], v. 94, n. 8, p. 1587–1601, 2006.

POPESCU, C. Selfheating and thermal runaway phenomena in semiconductor devices. **Solid-State Electronics**, [s.l.], v. 13, n. 4, p. 441–450, 1970.

PRASAD, C. et al. Self-heat reliability considerations on Intel's 22nm tri-gate technology. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Monterey, U.S.A.: [s.n.], 2013. p. 5D.1.1–5D.1.5.

PRASAD, C.; RAMEY, S.; JIANG, L. Self-heating in advanced CMOS technologies. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Monterey, U.S.A.: [s.n.], 2017. p. 6A–4.1–6A–4.7.

RALEVA, K. et al. Modeling thermal effects in nanodevices. **IEEE Transactions on Electron Devices**, [s.l.], v. 55, n. 6, p. 1306–1316, 2008.

RALEVA, K. et al. Study of self-heating effects in SOI and conventional MOSFETs with electro-thermal particle-based device simulator. **Journal of Computational Electronics**, [s.l.], v. 11, n. 1, p. 106–117, 2012.

RAMAN, A.; WALKER, D.; FISHER, T. Simulation of nonequilibrium thermal effects in power LDMOS transistors. **Solid-State Electronics**, [s.l.], v. 47, n. 8, p. 1265–1273, 2003.

RAMBO, P. W.; DENAVIT, J. Time stability of Monte Carlo device simulation. **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, [s.l.], v. 12, n. 11, p. 1734–1741, 1993.

REDDY, V. et al. Impact of negative bias temperature instability on digital circuit reliability. **Microelectronics Reliability**, [s.l.], v. 45, n. 1, p. 31–38, 2005.

REES, H. D. Calculation of distribution functions by exploiting the stability of the steady state. **Journal of Physics and Chemistry of Solids**, [s.l.], v. 30, n. 3, p. 643–655, 1969.

ROBERT, C. P. Monte Carlo Methods. In: ____. **Wiley StatsRef: Statistics Reference Online**. [S.l.]: John Wiley & Sons, Ltd, 2014. p. 1–13.

ROSENBERG, H. M. The thermal conductivity of germanium and silicon at low temperatures. **Proceedings of the Physical Society. Section A**, [s.l.], v. 67, n. 9, p. 837–840, 1954.

ROSSETTO, A. C. J. et al. Statistical analysis of the impact of charge traps in deeply-scaled MOSFETs via particle-based Monte Carlo device simulations. **IEEE Transactions on Electron Devices**, under review, 2018.

ROWLETTE, J. A.; GOODSON, K. E. Fully coupled nonequilibrium electron−phonon transport in nanometer-scale silicon FETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 55, n. 1, p. 220–232, 2008.

RUBINSTEIN, R. Y. **Simulation and the Monte Carlo Method**. 1st. ed. New York, U.S.A.: John Wiley & Sons, Inc., 1981. 414 p.

SAHA, S. Design considerations for 25 nm MOSFET devices. **Solid-State Electronics**, [s.l.], v. 45, n. 10, p. 1851–1857, 2001.

SEIDEL, T.; SCHARFETTER, D. Dependence of hole velocity upon electric field and hole density for p-type silicon. **Journal of Physics and Chemistry of Solids**, [s.l.], v. 28, n. 12, p. 2563–2574, 1967.

SELBERHERR, S. **Analysis and Simulation of Semiconductor Devices**. 2nd. ed. Vienna, Austria: Springer-Verlag, 1984. 293 p.

SESNIC, S. S.; CRAIG, G. R. Thermal effects in JFET and MOSFET devices at cryogenic temperatures. **IEEE Transactions on Electron Devices**, [s.l.], v. 19, n. 8, p. 933–942, 1972.

SHAIK, A. R. **Multi scale study of heat transfer using Monte Carlo technique for phonon transport**. 75 p. Dissertação (Mestrado) — Arizona State University, Tempe, U.S.A., 2016.

SHANKS, H. R. et al. Thermal conductivity of silicon from 300 to 1400 K. **Physical Review**, College Park, U.S.A.: American Physical Society, v. 130, n. 5, p. 1743–1748, 1963.

SHARMA, D. K.; RAMANATHAN, K. Modification of MOST I-V characteristics by self-heating. **Solid-State Electronics**, [s.l.], v. 27, n. 11, p. 989–994, 1984.

SHI, Z.; MIEVILLE, J. P.; DUTOIT, M. Random telegraph signals in deep submicron n-MOSFETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 41, n. 7, p. 1161–1168, 1994.

SILVACO. **Giga — Non-Isothermal Device Simulator**. 2018. Disponível em: <https://www.silvaco.com/products/vwf/atlas/giga/giga_br.html>. Acesso em: 02/10/2018.

SIMOEN, E. et al. Explaining the amplitude of RTS noise in submicrometer MOSFETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 39, n. 2, p. 422–429, 1992.

SINHA, S. et al. Non-equilibrium phonon distributions in sub-100 nm silicon transistors. **Journal of Heat Transfer**, New York, U.S.A.: American Society of Mechanical Engineers, v. 128, n. 7, p. 638–647, 2006.

SINITSKY, D. et al. Velocity overshoot of electrons and holes in Si inversion layers. **Solid-State Electronics**, [s.l.], v. 41, n. 8, p. 1119–1125, 1997.

STATHIS, J.; ZAFAR, S. The negative bias temperature instability in MOS devices: a review. **Microelectronics Reliability**, [s.l.], v. 46, n. 2, p. 270–286, 2006.

STONE, H. L. Iterative solution of implicit approximations of multidimensional partial differential equations. **SIAM Journal on Numerical Analysis**, [s.l.], v. 5, n. 3, p. 530–558, 1968.

SU, L. T. et al. Measurement and modeling of self-heating in SOI nMOSFETs. **IEEE Transactions on Electron Devices**, [s.l.], v. 41, n. 1, p. 69–75, 1994.

SZE, S. M. **Physics of Semiconductor Devices**. 3rd. ed. New York, U.S.A.: Wiley, 2006. 832 p.

TAKACS, D.; TRAGER, J. Temperature increase by self-heating in VLSI CMOS. In: **Proceedings of the European Solid State Device Research Conference**. Bologna, Italy: [s.n.], 1987. p. 729–732.

TAUR, Y.; WANN, C. H.; FRANK, D. J. 25 nm CMOS design considerations. In: **Proceedings of the IEEE International Electron Devices Meeting**. San Francisco, U.S.A.: [s.n.], 1998. p. 789–792.

TENBROEK, B. M. et al. Self-heating effects in SOI MOSFETs and their measurement by small signal conductance techniques. **IEEE Transactions on Electron Devices**, [s.l.], v. 43, n. 12, p. 2240–2248, 1996.

TENBROEK, B. M. et al. Measurement and simulation of self-heating in SOI CMOS analogue circuits. In: **Proceedings of the IEEE International SOI Conference**. [Fish Camp, U.S.A.: [s.n.], 1997. p. 156–157.

TENBROEK, B. M. et al. Impact of self-heating and thermal coupling on analog circuits in SOI CMOS. **IEEE Journal of Solid-State Circuits**, [s.l.], v. 33, n. 7, p. 1037–1046, 1998.

THEODOROU, C. G. et al. New LFN and RTN analysis methodology in 28 and 14 nm FD-SOI MOSFETs. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Monterey, U.S.A.: [s.n.], 2015. p. XT.1.1–XT.1.6.

TIAN, Z. et al. On the importance of optical phonons to thermal conductivity in nanostructures. **Applied Physics Letters**, [s.l.], v. 99, n. 5, p. 053122–1–053122–3, 2011.

TIEN, C. L.; MAJUMDAR, A.; GERNER, F. M. **Microscale Energy Transport**. Washington, U.S.A.: Taylor & Francis, 1998. 395 p.

VANDAMME, L.; SODINI, D.; GINGL, Z. On the anomalous behavior of the relative amplitude of RTS noise. **Solid-State Electronics**, [s.l.], v. 42, n. 6, p. 901–905, 1998.

VASILESKA, D.; GOODNICK, S. M. Computational electronics. **Materials Science and Engineering: R: Reports**, [s.l.], v. 38, n. 5, p. 181–236, 2002.

VASILESKA, D.; GOODNICK, S. M.; KLIMECK, G. **Computational Electronics: Semiclassical and Quantum Device Modeling and Simulation**. 1st. ed. New York, U.S.A.: CRC Press, 2010. 782 p.

VASILESKA, D. et al. Semiconductor device modeling. **Journal of Computational and Theoretical Nanoscience**, Valencia, U.S.A.: American Scientific Publishers, v. 5, n. 6, p. 999–1030, 2008.

VASILESKA, D.; RALEVA, K.; GOODNICK, S. M. Modeling heating effects in nanoscale devices: the present and the future. **Journal of Computational Electronics**, [s.l.], v. 7, n. 2, p. 66–93, 2008.

VASILESKA, D.; RALEVA, K.; GOODNICK, S. M. Electrothermal studies of FD-SOI devices that utilize a new theoretical model for the temperature and thickness dependence of the thermal conductivity. **IEEE Transactions on Electron Devices**, [s.l.], v. 57, n. 3, p. 726–728, 2010.

VASILESKA, D.; RALEVA, K.; GOODNICK, S. M. Heating effects in nanoscale devices. In: ____. **Cutting Edge Nanotechnology**. London, U.K.: InTech, 2010. p. 33–60.

WACHUTKA, G. K. Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling. **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, [s.l.], v. 9, n. 11, p. 1141–1149, 1990.

WEBER, L.; GMELIN, E. Transport properties of silicon. **Applied Physics A**, [s.l.], v. 53, n. 2, p. 136–140, 1991.

WECKX, P. et al. Relaxation of time-dependent NBTI variability and separation from RTN. In: **Proceedings of the IEEE International Reliability Physics Symposium**. Monterey, U.S.A.: [s.n.], 2017. p. XT–9.1–XT–9.5.

WHITE, G. K.; WOODS, S. B. Thermal conductivity of germanium and silicon at low temperatures. **Physical Review**, College Park, U.S.A.: American Physical Society, v. 103, n. 3, p. 569–571, 1956.

WIRTH, G. I.; SILVA, R. da; KACZER, B. Statistical model for MOSFET bias temperature instability component due to charge trapping. **IEEE Transactions on Electron Devices**, [s.l.], v. 58, n. 8, p. 2743–2751, 2011.

WORDELMAN, C.; RAVAIOLI, U. Integration of a particle-particle-particle-mesh algorithm with the ensemble Monte Carlo method for the simulation of ultra-small semiconductor devices. **IEEE Transactions on Electron Devices**, [s.l.], v. 47, n. 2, p. 410–416, 2000.

YOUNG, D. Iterative methods for solving partial difference equations of elliptic type. **Transactions of the American Mathematical Society**, Providence, U.S.A.: American Mathematical Society, v. 76, n. 1, p. 92–111, 1954.

ZEGHBROECK, B. V. **Principles of Semiconductor Devices and Heterojunctions**. 1st. ed. Upper Saddle River, U.S.A: Prentice Hall, 2011. 450 p.

APPENDIX A — VALIDATION OF THE THERMAL SOLVER

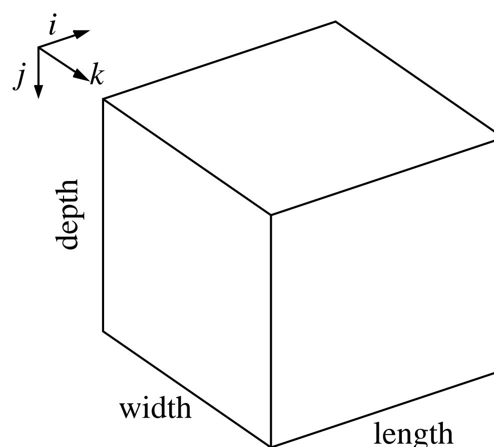
The test structure utilized for the temperature diffusion simulations was a six-sided solid of an arbitrary material with length, width, and depth equal to 100 nm. In the coordinate system, the length refers to the direction following the i-axis, whereas the depth and width follow the j- and k-axis, respectively, as depicted in Figure A.1. A total of 50 grid points were used for each direction, resulting in a uniform mesh spacing Δ of 2 nm. The thermal conductivity κ of the material was chosen to be 100 $\mathrm{Wm^{-1}K^{-1}}$. In order to cover a variety of situations, the case studies simulations were divided as follows.

Case I: the temperature diffusion was simulated considering uniform thermal conductivity and no heat source within the structure. The temperatures at the borders were set to $T_{i_{\min}} = 200$ K, $T_{i_{\max}} = 400$ K, $T_{j_{\min}} = 300$ K, and $T_{j_{\max}} = 100$ K. For the surfaces at $k_{\min}$ and $k_{\max}$, adiabatic boundary conditions (floating temperature) were applied. The temperature profile is depicted in Figure A.2a, where it is possible to observe a diagonal in the $ij-$plane whose temperature approximates to 250 K, which corresponds to the mean temperature between adjacent the surfaces $T_{i_{\min}}$ and $T_{j_{\min}}$, and $T_{i_{\max}}$ and $T_{j_{\max}}$, as expected.

Case II: the temperature diffusion was simulated also considering uniform thermal conductivity and without heat source within the structure, but applying boundary temperatures only for the surfaces at $i_{\min}$ and $i_{\max}$, i.e., $T_{i_{\min}} = 200$ K and $T_{i_{\max}} = 400$ K. Adiabatic boundaries were used for the other surfaces. The temperature distribution across the system is shown in Figure A.2b. In this case, the temperature at the center of the structure, i.e., $(i_{\max} + i_{\min})/2$, is found to be virtually 300 K, i.e., $(T_{i_{\max}} + T_{i_{\min}})/2$, as expected.

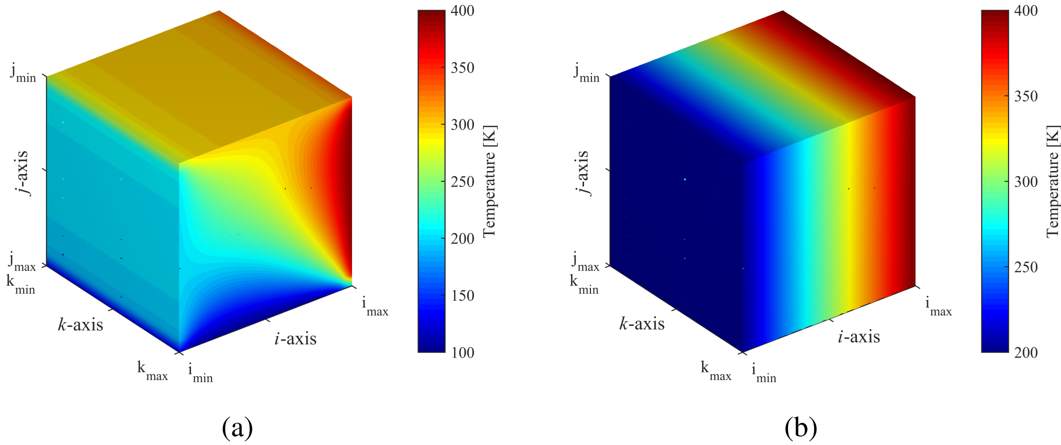
Case III: in this case, the role of the thermal conductivity was tested through a sim-

Figure A.1: Graphical representation of the test structure.



Source: author.

Figure A.2: Simulated temperature profile for Case I (a) and Case II (b).

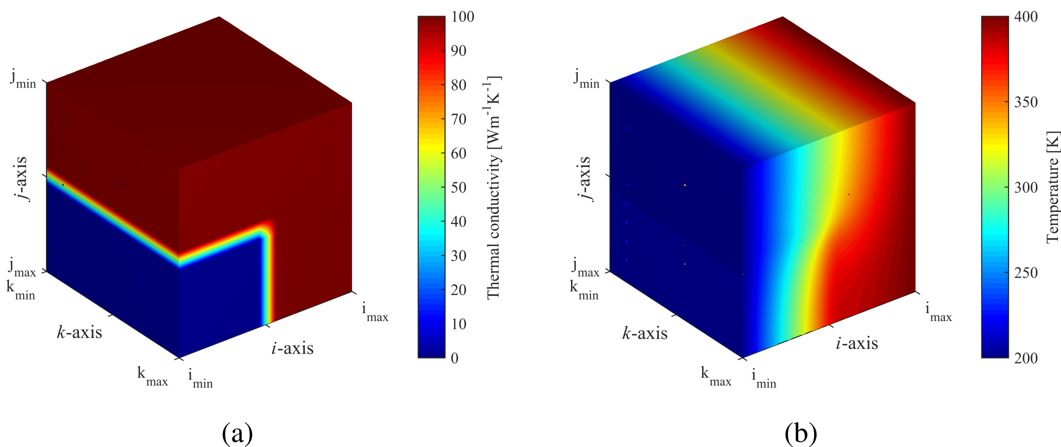


Source: author.

ulation utilizing a non-uniform thermal conductivity profile, which is depicted in Figure A.3a. The temperature boundaries were set as in Case II, and the obtained temperature profile is shown in Figure A.3b. For the top portion of the test structure, where the thermal conductivity is constant along the i-axis (κ = 100 Wm^{-1}K^{-1}), the gradient of temperature between i_{min} and i_{max} is uniform. On the other hand, for the bottom portion of the cube, where the thermal conductivity changes abruptly at the center (from κ = 1 to 100 Wm^{-1}K^{-1}), a non-uniform temperature distribution occurs. In addition, the largest "temperature drop" occurs over the region with the lowest κ, i.e., the most thermal resistive region, as expected. This behavior is analog to the voltage distribution across a resistive network.

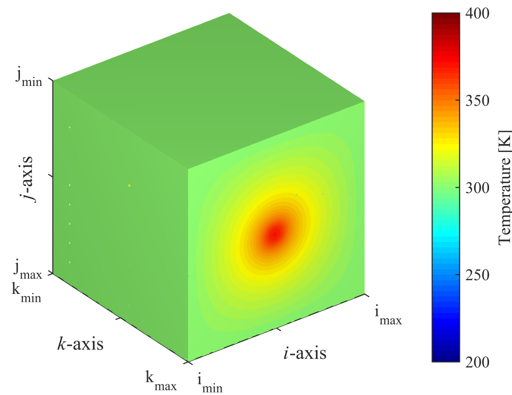
Case IV: the simulation performed in this case explores the temperature distribu-

Figure A.3: Thermal conductivity profile (a) and temperature profile (b) for Case III.



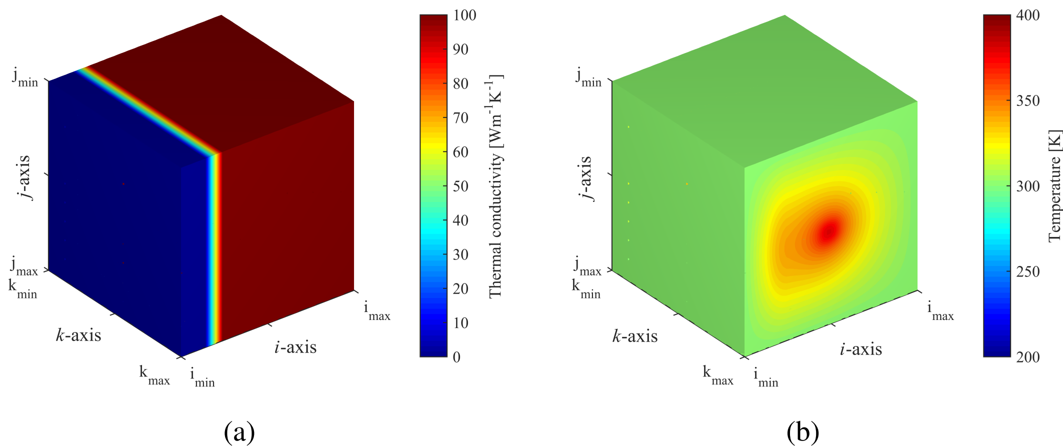
Source: author.

Figure A.4: Temperature profile for Case IV.



Source: author.

Figure A.5: Thermal conductivity profile (a) and temperature profile (b) for Case V.



Source: author.

tion for a uniform material but considering the presence of a heat source in the center the structure. A fixed temperature of 300 K was applied for the surfaces at i_{min}, i_{max}, j_{min}, and j_{max}, whereas adiabatic boundaries were applied for the surfaces at k_{min} and k_{max}. The obtained temperature profile is depicted in Figure A.4, where one can notice the uniform temperature distribution from the center to the edges of the structure.

Case V: the same conditions of Case IV were utilized in this case, except for a non-uniform thermal conductivity profile, which is depicted in Figure A.5a. The resulting thermal profile, in turn, is depicted in Figure A.5b. Now, the temperature distribution is no longer uniform within the structure, since the heat flow is partially confined in one direction. As a consequence, it is possible to observe a slight temperature rise in the hot spot of Figure A.5b when compared with the hot spot of Figure A.4.

From the qualitative results obtained at this step, the numeric solver was considered to be properly implemented.

APPENDIX B — PUBLICATIONS

A. Rossetto, V. Camargo, D. Vasileska, and G. Wirth, **Study of the Impact of Self-heating on the Performance of p-type MOSFET Transistors**. American Materials Research Society Spring Meeting. Phoenix, AZ. 2016.

A. Rossetto, V. Camargo, D. Vasileska, and G. Wirth, **Novel State-of-the-Art Monte Carlo Device Simulator for Modeling RDF, RTN and NBTI Using Real-Space Treatment of the Coulomb Interactions and Self-Heating Effects**. Variability Modeling and Characterization Workshop, ICCAD. Dallas, TX. 2016.

D. Colombo, A. Rossetto, G. Wirth, S. Bampi, and O. Gonçalez, **Total Dose Effects on Voltage References in 130-nm CMOS Technology**. IEEE Transactions on Device and Materials Reliability, [s.l.], v. 18, n. 1, p.27-36, 2018.

A. Rossetto, T. Both, V. Camargo, D. Vasileska, and G. Wirth, **Statistical Analysis of the Impact of Charge Traps in Deeply-Scaled MOSFETs via Particle-based Device Simulations**. IEEE Transactions on Electron Devices, under review, 2018.