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

VINÍCIUS VALDUGA DE ALMEIDA CAMARGO

Evaluating the impact of charge traps on MOSFETs and Circuits

Thesis presented as a partial fulfillment of the requirements for the degree of Doctorate in Microelectronics.

Advisor: Prof. Dr. Gilson Inácio Wirth

Porto Alegre
2016

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Valduga de Almeida Camargo, Vinícius
Evaluating the impact of charge traps on MOSFETs and Circuits / Vinícius Valduga de Almeida Camargo. – 2016.
77 f.:il.
Orientador: Gilson Inácio Wirth;
Tese (Doutorado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Microeletrônica. Porto Alegre, BR – RS, 2016.
1.RTN. 2.BTI. 3.Ensemble Monte Carlo. 4.Simulação de circuitos. 5.TCAD. I. Wirth, Gilson. II. Evaluating the impact of charge traps on MOSFETs and Circuits.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. Rui Oppermann
Vice-Reitora: Profa. Jane Tutikian
Pro-Reitor de Pos-Gradução: Prof. Celso Giannetti Loureiro Chaves
Diretora do Instituto de Informática: Profa. Carla Maria Dal Sasso Freitas
Coordenador do Pós-Graduação em Microeletronica: Profa. Fernanda Gusmão de Lima Kastensmidt
Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

ACKNOWLEDGEMENTS

I am very grateful to all my masters, specially to my advisor Gilson Wirth, to professor Dragica Vasileska and to Ben Kaczer, without the discussions with them and the knowledge they passed to me, nothing in this thesis would be possible.

I also grateful to my family and my beloved ones for all the love and support along this journey.

I am grateful my colleges in LAPROT, at ASU and at IMEC, specially to Maurício Banaszeski da Silva. Not only for making the day-to-day in the laboratory a pleasure, but also for the daily discussions in each other research. That is how many great ideas begin.

Análise do impacto de armadilhas em MOSFET's e circuitos

RESUMO

Nesta tese são apresentados estudos do impacto de armadilhas no desempenho elétrico de MOSFETs em nível de circuito e um simulador Ensamble Monte Carlo (EMC) é apresentado visando a análise do impacto de armadilhas em nível de dispositivo. O impacto de eventos de captura e emissão de portadores por armadilhas na performance e confiabilidade de circuitos é estudada. Para tanto, um simulador baseado em SPICE que leva em consideração a atividade de armadilhas em simulações transientes foi desenvolvido e é apresentado seguido de estudos de caso em células SRAM, circuitos combinacionais, ferramentas de SSTA e em osciladores em anel. Foi também desenvolvida uma ferramenta de simulação de dispositivo (TCAD) atomística baseada no método EMC para MOSFETs do tipo p. Este simulador é apresentado em detalhes e seu funcionamento é testado conceitualmente e através de comparações com ferramentas comerciais similares.

Palavras-chave: RTN. BTI. Ensemble Monte Carlo. Simulação de circuitos. TCAD.

Evaluating the impact of charge traps on MOSFETs and Circuits

ABSTRACT

This thesis presents studies on the impact of charge traps in MOSFETs at the circuit level, and a Ensemble Monte Carlo (EMC) simulation tool is developed to perform analysis on trap impact on PMOSFETs. The impact of charge trapping on the performance and reliability of circuits is studied. A SPICE based simulator, which takes into account the trap activity in transient simulations, was developed and used on case studies of SRAM, combinational circuits, SSTA tools and ring oscillators. An atomistic device simulator (TCAD) for modeling of p-type MOSFETs based on the EMC simulation method was also developed. The simulator is explained in details and its well function is tested.

Keywords: Traps. RTS. BTI. Ensemble Monte Carlo. TCAD. Circuit simulations.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS AND ACRONYMS

ASIC	Application Specific Integrated Circuit
ASU	Arizona State University
BSIM	Berkeley Short-Channel IGFET Model
BTI	Bias Temperature Instability
CAD	Computed Aided Design
CDF	Cumulative Distribution Function
CIC	Cloud In Cell
CLT	Central Limit Theorem
CMOS	Complementary Metal-Oxide-Semiconductor
DC	Direct Current
EMC	Ensemble Monte Carlo
LER	Line Edge Roughness
NBTI	Negative Bias Temperature Instability
NEC	Nearest Element Center
NGP	Nearest Grid Point
NMOS	Negative Metal Oxide Semiconductor
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
PBTI	Positive Bias Temperature Instability
PDF	Probability Density Function
PMOS	Positive Metal Oxide Semiconductor
PTM	Predictive Technology Model
PVT	Process, Voltage and Temperature
RDF	Random Dopant Fluctuation
RO	Ring Oscillator
RTN	Random Telegraph Noise
RTS	Random Telegraph Signal
SIP	Strong Implicit Method
SPICE	Simulation Program With Integrated Circuit Emphasis
SRAM	Static Random Access Memory
STA	Static Timing Analysis
SSTA	Statistical Static Timing Analysis
TCAD	Technology Computed Aided Design

TABLE OF CONTENTS

1 INTRODUCTION

Charge trapping and de-trapping at localized states at the interface or in the gate dielectric is a significant reliability issue in state of the art CMOS applications (VAIDYANATHAN, 2012). It is known to be a source of low-frequency noise, and also considerably contributes to the Bias Temperature Instability (BTI). These localized states are known as traps. Trapping and de-trapping are stochastic events, and a trap may be either populated or empty. A trap becomes populated when capturing a charge carrier and becomes empty after emitting it. These capture and emission events dynamically impact the electrical parameters of the transistor, generating time dependent reliability issues at the circuit level (KACZER, 2010).

According to the characteristics of each trap, it may impact the MOSFET in a different way (TOLEDANO-LUQUE, 2012). If a trap is more likely to capture a carrier than to emit it, this trap will tend to remain occupied, producing an aging effect. BTI is an aging phenomenon attributed to this trap activity effect (GRASSER, 2011). On the other hand, a trap which is equally likely to capture and to emit carriers will likely change its state frequently generating a noise effect. This effect is known as Random Telegraph Signal (RTS) or Random Telegraph Noise (RTN).

In a MOSFET device, a carrier might be trapped at the Si/SiO_2 interface or deeper into the gate dielectric. To first order, the impact of a trapped carrier on the electrical behavior of the transistor is inversely proportional to the channel area regardless of its position. The miniaturization of the transistor's channel, hence, leads to an increase of the relevance of charge trapping and de-trapping phenomena. Furthermore, Random Dopant Fluctuations (RDF) will induce percolation paths, further increasing the impact of a single trap and leading to variability issues in circuit electrical response (GERRER, 2014). Another trend is the increase of the nominal threshold voltage and/or the decrease of the operation voltage, aiming to reduce the power consumption. When combining these effects, the increase of the impact of trap activity on transistor variability becomes evident, leading to serious circuit reliability issues. Atomistic (TCAD) simulation results show that the impact of a single trap on a 45nm channel length NMOS transistor might cause an increase as high as 16% on the device threshold voltage (ASHRAF, 2011). This kind of simulation, however, cannot be used to study the traps impact in circuits with more than a couple of transistor due to its computational cost.

Aiming to better evaluate the impact of traps in the circuit level and aid designers to cope with effects related to the trap activity, a simulation tool capable of simulating the trap

kinetics and of estimating its impact on transient electrical simulations was developed. Chapter 2 of this thesis further motivates the need of such a simulation tool comparing to others in the literature. The model used and the simulation framework are meticulously described and case studies show the wide range of applications in which it can be used. Each case study is presented providing insights on how the trap activity affects different circuit types.

Chapter 3 changes the focus from the circuit level to the device level. In this chapter it is presented an Ensemble Monte Carlo device simulation tool which was developed for PMOS transistor. Details of the implementation of the code are provided beginning from the scattering probabilities calculation and the bulk simulation to the device simulation. Results showing that it conforms to results from commercial tools are also presented. The Ensemble Monte Carlo device code is then used to evaluate the impact charges trapped in the dielectric or in the Si/SiO_2 cause in the behavior of the device. Chapter 4 discusses the results achieved in this thesis and presents future works that may be obtained with the two simulation tools developed and presented here.

2 TRAP KINETICS ON TRANSIENT CIRCUIT SIMULATIONS

As RTN has become a critical effect for circuit reliability, its impact on circuits has been widely investigated. Traditionally RTN was evaluated only in the frequency domain, therefore the circuit analysis techniques to evaluate the RTN impact on circuits were developed based on frequency domain analysis (LEYRIS, 2007). For digital and low power applications, however, it is of most interest to make studies in the time-domain (YE, 2010), but up until now there are still no commercial SPICE tools available that allow this type of simulation.

Recently many authors presented simulation methodologies capable to simulate RTN effect on transient SPICE analysis thus allowing the investigation of its impact at the circuit level. A cyclo-stationary simulation method for RTN was first presented by Van Der Wel (2003) and further developed by Kolhatkar (2005) and Wirth (2010) and (2012). Other simulation methodologies were developed at the electrical level (SPICE level). These methodologies are based on i) equivalent electrical models for each trap (REISINGER, 2012) and (YE, 2010), where trap behavior is modeled by equivalent circuits; or ii) by evaluating the trap activity outside the electrical simulation tool in a single simulation as in (TANG, 2010) and in MUSTARD, or in multiple simulations as in SAMURAI, (AADITHYA, 2013) aiming to consider the bias dependency.

The goal of this kind of simulation is to obtain a reliable result for circuit simulations considering the trap activity on transient simulations. A cyclo-stationary simulator can be used to perform case studies on specific circuits that might give insights to designers in order to develop newer, more reliable design techniques, and also to evaluate the traps impact on small critical blocks widely used on designs such as SRAM cells, ring oscillators and analog blocks. Another possible use of this kind of simulations is to provide a golden simulation result for the development of more computing efficient CAD tools that would be used during the design of large real circuits, such SSTA tools.

The previous methodologies presented in the literature are either not accurate enough because they do not support models which depend on the bias conditions or are not computationally efficient, becoming unviable to run Monte Carlo simulations on realistic circuits under usual switching conditions. The existing methodologies also were focused on considering just RTN neglecting other effects caused by the trap activity, such as BTI.

This chapter presents in depth a circuit simulation methodology capable of performing workload dependent trap simulations in the time-domain on arbitrary circuit designs under arbitrary bias conditions. It is thus capable to evaluate the effects caused by the traps, such as

BTI and RTN, at the circuit level. As many studies on the trap kinetics and its impact on transistors are still being made, many new models and enhancements to previous models, which may be more suitable for different technology nodes, are still being developed. On this kind of simulation, the complexity of the model is tightly attached to the computational cost of the simulation. The methodology presented in this manuscript was developed so that the model used by the simulation tool can be easily changed, thus allowing the user to choose the model that best fits their needs.

This chapter first presents the model of trap kinetics which is used in the case studies, then the simulation tool developed is presented and case studies on a ring oscillator, on a SRAM cell and a comparison with commercial SSTA tools on combinational circuits are presented and discussed. Finally, a conclusion section is presented summarizing the main contributions of this work.

2.1 Trap Kinetics Model

Oxide and interface traps capture and emit charge carriers responsible for the drain current of the MOSFET. When a trap captures a charge carrier, the drain current is affected due to the electrostatic effect, which leads to reduction in number of carriers in the channel and also affects mobility. This is here modeled as a V_{th} fluctuation.

For a constant bias condition i.e., in a steady state condition, the traps whose energy level is close to the Fermi level will present larger activity. By changing the transistor's bias point, the difference between the trap's energy level and the Fermi level is also changed, thus changing the occupation probability of the trap (GRASSER, 2012). Responding to this change, the traps tend to change their states, adapting to the new occupation probability. However, this change is not instantaneous, but a function of the time constants of each trap. The same model can be used to describe the charge trapping activity, both for steady state (DC) condition as well as for arbitrary bias conditions (GRASSER, 2010).

Capture and emission of a charge carrier by a trap is modeled as a Poisson processes with rates τ_c and τ_e. State 1 stands for the populated trap, while state 0 stands for the empty trap. τ_c and τ_e are then the average residence times in states 0 and 1 in a steady state condition, respectively (MACHLUP, 1954).

Equations (2.1) and (2.2) shown the probability of a particular trap to change its state after an elapsed time Δt (WIRTH, 2011).

$$P_{01}(\Delta t) = \left[1 - \exp\left(\frac{-\Delta t}{\tau_{eq}}\right)\right] \cdot \frac{\tau_e}{(\tau_c + \tau_e)} \quad (2.1)$$

$$P_{10}(\Delta t) = \left[1 - \exp\left(\frac{-\Delta t}{\tau_{eq}}\right)\right] \cdot \frac{\tau_c}{(\tau_c + \tau_e)} \quad (2.2)$$

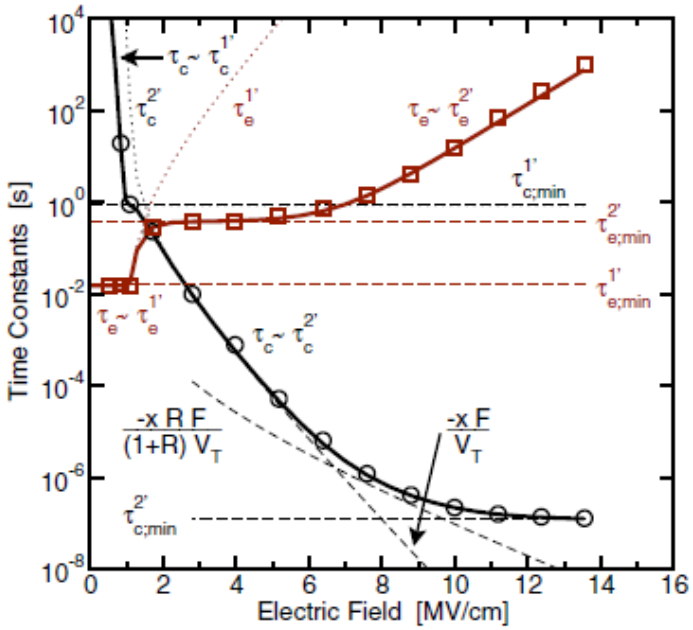
Where $1/\tau_{eq} = 1/\tau_c + 1/\tau_e$.

The equations (2.1) and (2.2) can be summarized into equation (2.3). The bias point condition and the process are included in the time constant parameter. For our circuit simulation in the time domain, Eq. (2.3) is defined as the time step of the simulation, 'p' is the process, either capture ('c') or emission ('e'), 'p*' is the complimentary process, 'v' is the bias condition on that given instant of time and is the trap time constant.

$$P_{p,v} = \frac{\tau_{p^*,v}}{\tau_{e,v}+\tau_{c,v}}\left\{1 - exp\left[-\left(\frac{1}{\tau_{e,v}} + \frac{1}{\tau_{c,v}}\right)\Delta t\right]\right\} \quad (2.3)$$

These equations apply to any bias condition. However, as depicted by Figure 2.1, the time constants are a function of the bias point. This dependency is included in the model by a numerical fitting algorithm which aims at finding the correct value of the time constants for a given bias point from known time constants defined for a known bias point.

Figure 2.1 - Bias point dependency of the capture and emission times for a particular trap.



Source: (GRASSER, 2010)

For the emission time, the simulator is fed with two parameters, $\tau_{e;min}^{1'}$ and $\tau_{e;min}^{2'}$, which are defined in (GRASSER, 2010). The emission time used during the simulations is given by equation (2.4). For the capture time, the simulator is fed with the capture time for two different gate voltages in the range of operation of the device. Using a linear approximation on the log-scale, the capture time of this given trap is calculated for the other gate voltages. The linear approximation presented to be satisfactory for electric fields in the range of 1-6MV/cm as shown in Figure 2.1. For a more accurate, but more computing cost simulation, the voltage dependency model can be changed to follow the model presented in (GRASSER, 2010).

$$\tau_e = \begin{cases} \tau_{e;min}^{1'} & V_g \leq V_T \\ \tau_{e;min}^{2'} & V_g > V_T \end{cases} \tag{2.4}$$

2.2 Simulation Framework

This section presents the simulator framework developed and used in this study. Because not only the charge trapping kinetics are a stochastic phenomenon, but also the number of traps and their properties are described as random variables, the simulation framework is decomposed into two different levels. One level evaluates the trap activity during the electrical simulation and the other one controls the simulation flow externally.

In order to include the impact of the trap activity into the electrical simulator, the transistor model used in the simulations was enhanced by adding the trap kinetics equations presented on section 2.1. This enhancement was performed by creating a Verilog-A component which contains the BSIM4 description. The Verilog-A standard provides a flexible way to model devices and integrate new models into electrical simulations (CHALKIADAKI, 2012).

All the information regarding the trap properties for a given technology are included in a configuration file. This file also contains the information about the running options, such as the number of runs of a Monte Carlo Simulation; and analysis options, such as if time zero variability is to be considered in the analysis.

The input parameters for each instance of the Verilog-A component are: i) number of traps for each transistor in the circuit; ii) the time constants for each trap; iii) the impact of each trap on the threshold voltage (when populated); iv) the parameters that define the dependence of capture and emission probability on temperature and bias point, for each trap; and v) the initial state of each trap. All these parameters are defined by the control script based on the configuration file information. The values of the parameters are defined for each instance prior to each electrical simulation.

Once that every instance has all of its parameters defined, the electrical simulation is performed. At each time step of the transient simulation, equation (2.3) is used to evaluate the probability of the trap to change its state. Based on this information the trap's next state is randomly defined. For each populated trap, its impact on Vth is added to the threshold voltage parameter of the transistor model.

The impact on V_{th} caused by each trap is considered to be a constant value independent of bias and temperature. Even though recent studies show that the trap's impact on V_{th} can have dependency on bias (FRANCO, 2012), this is not a strong one and has a negligible impact during the on-state conditions. There is still no conclusive study detailing this dependency nor models capable of describing it. The methodology is capable of considering this dependency by adding models in the Verilog-A model (as done to include the bias dependency of time constants).

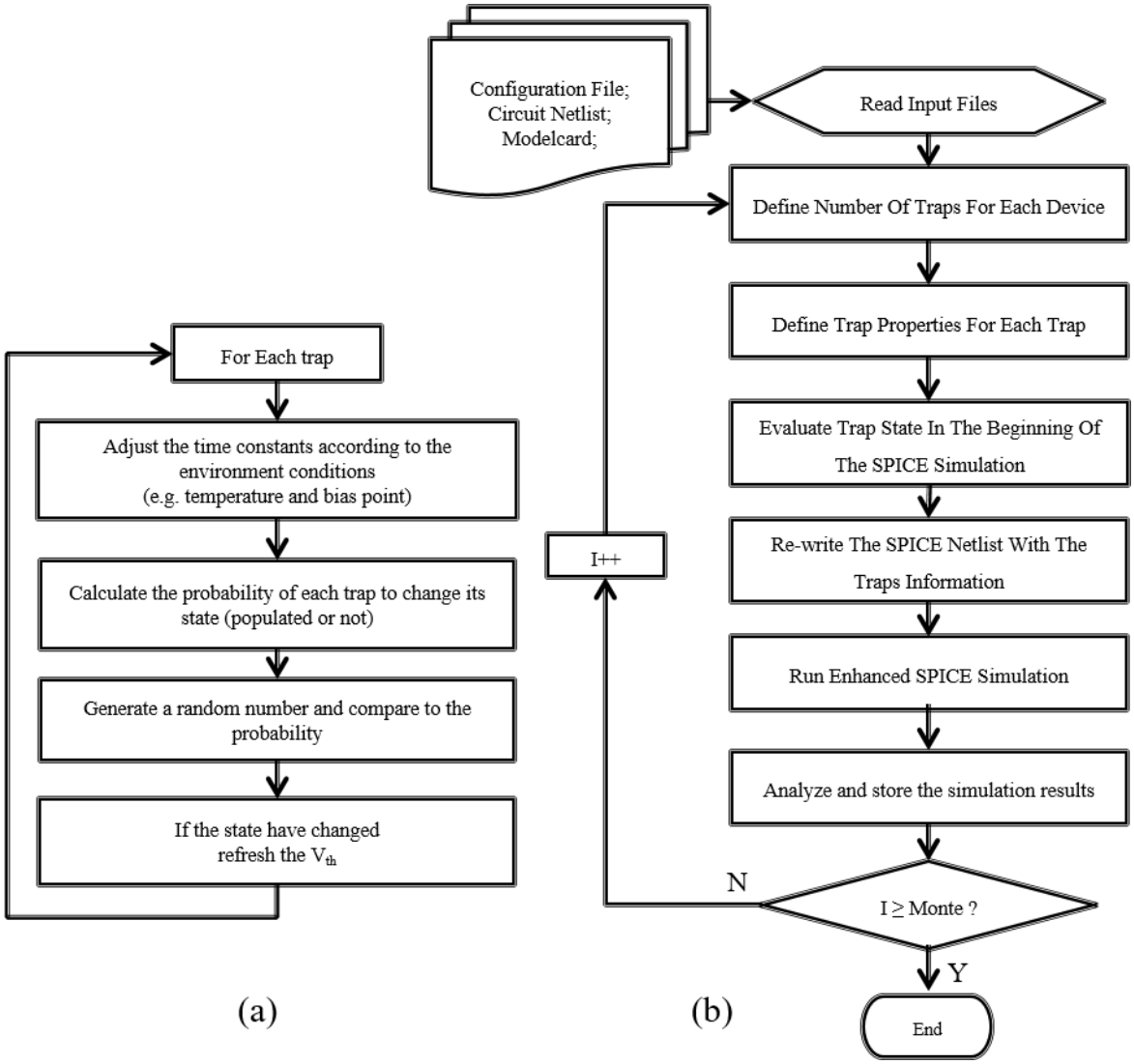
The time constants used in equation (2.3), however, are a function of environmental variables such as the electrical field at the Si/SiO_2 interface as discussed for instance in (GRASSER, 2010). To include these dependencies, during each step, prior to evaluate the next state of the trap the time constants are adjusted numerically, as presented in section 2.1, based on the environmental parameters of each device and the conditions at which the trap was characterized. For a more precise consideration of these dependencies, one might also use the equations (15) and (16) presented in (GRASSER, 2010). A summary of this flow is presented on Figure 2.2(a).

By performing an electrical simulation using the Verilog-A component, it is possible to simulate a given circuit with pre-defined traps considering the trap activity during the transient simulation. Aiming to improve the simulation process, a control script was written. Based on a configuration file that contains the information about the probability density functions of the trap related parameters and on the netlist of the studied circuit, this control script writes a new netlist changing the MOSFETs devices by the Verilog-A component and adding, to each instance of this component, its parameters (number of traps, time constants, and so on).

The control script starts by reading the configuration file and the netlist of the circuit of interest. For each Monte Carlo run, the control script defines the number of traps for each Verilog-A component in the circuit. For each MOSFET, the number of traps is randomly chosen according to a Poisson distribution as shown in (KEYES, 1975). The parameters of this distribution are defined in the configuration file and are adjusted based on the design parameters of each given component, such as its channel area (WIRTH, 2012). During this stage the control

script also writes the experimental characterization information related to the traps, which is also found in the configuration file, into every instance of the Verilog-A component.

Figure 2.2- High level simulation framework (b) including the detail of the Trap kinetics on Verilog-A component (a).



Once the number of traps on each MOSFET is defined, the control script defines the impact of the trap on the threshold voltage (ΔV_{th}) and the time constants (τ_c and τ_e). Both parameters are random variables where the first is shown to be described by an exponential distribution with average η, where η scales with the device properties as presented in (KACZER, 2010). For simplicity the time constants are described by uniform distributions in a logarithmic scale. In general, a more accurate bi-variate Gaussian distribution can be used (BINA, 2012). Studies show that there is no correlation between the step-height and the time constants of each trap (KACZER, 2010). The definitions of these parameters for each trap are randomly chosen according to the statistical distribution of each random variable. The

distributions and their properties are defined in the configuration file and are adjusted for each transistor based on its design dependent parameters, such as the channel area.

After all trap related information of each device in the design is defined, it is possible to run an electrical simulation which will consider the trap kinetics. Since all the equations were added to the transistor models and not to the simulation tool, any commercial tool that supports the Verilog-A standard is able to perform the simulation. Post simulation scripts aiming to analyze the results may be added after the electrical simulation to avoid losing information as other electrical simulations run in the same Monte Carlo loop.

The same trap kinetics mechanisms can be used to simulate BTI effects. Because BTI is an aging effect we are interested in evaluating its impact after a given time of operation, that may be in the range of years, and no longer in short time as we did for RTN. It is, however, impracticable to run transient electrical simulations in the range of years.

Aiming to allow users to simulate BTI effects, we implemented an analytical function capable of predicting the trap state after a certain period of time under a predefined stress condition. This way the user may describe the stress signal by its voltage, frequency and duty cycle, and define the duration of this stress condition. This function will, thus, evaluate the state of each trap for the stress condition defined and will feed the electrical simulator which will use these as its initial condition. The analytical solution is based on the trap kinetics equations presented on section 2.1. A method of predicting the trap state after an arbitrary stress time under known conditions is presented in (TOLEDANO-LUQUE, 2011-a).

In contrast to the method that evaluates the trap activity on a transient SPICE simulation, the analytical method presented in (TOLEDANO-LUQUE, 2011-a) is not capable to consider the impact that degraded devices have on the other devices surrounding it. A circuit with degraded devices produces slightly different waveforms which will be the stress signals for other circuits, for the case of a logic gate the rise and fall times will change, also changing how the next stage is stressed. A comparison between the two methods was performed, validating the methodology of (TOLEDANO-LUQUE, 2011-a) for long stress times where the BTI phenomena dominates, for short periods of time, where RTN dominates, the correlation between the impact of traps is more significant and thus the enhanced SPICE simulation considering the trap kinetics should be used.

2.3 Case Studies

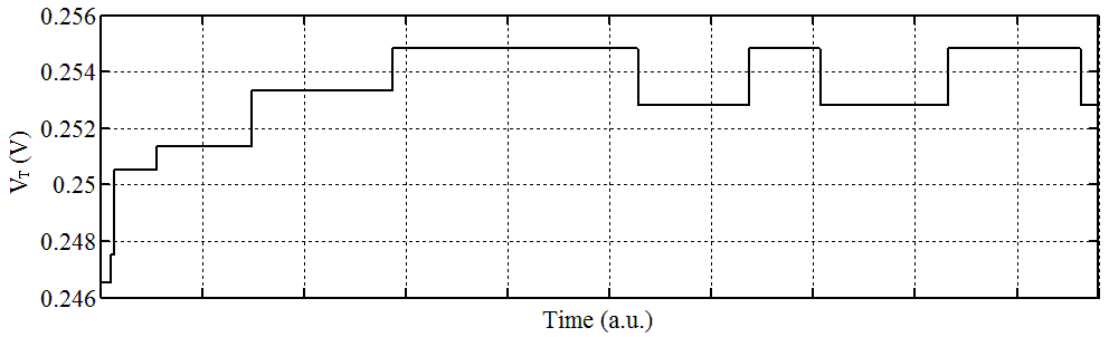
Aiming to better illustrate the behavior of the simulation tool and its capabilities as to study circuits of interest for designers, some case studies were performed and are presented

here. For all case studies the test circuit was designed on a 45nm CMOS technology. A PTM 45nm model card for high-performance applications was used for the SPICE simulation (ZHAO, 2006). The trap properties used in the simulation are described in detail in (TOLEDANO-LUQUE, 2011-b). Since trap related properties (as for instance number of traps on each device) are random variables, a statistical analysis method is mandatory.

2.3.1 Ring Oscillators

For an initial study, a transient simulation was run and the V_{th} of a PMOS transistor was measured right after the transistor started to be stressed. The result is plotted on Figure 2.3. In this figure the trap activity and its impact on V_{th} are clearly seen. As each trap causes a particular impact on V_{th} it is possible to identify each trap and when it changed its state. The increase of V_{th} in time is caused by the BTI effect. In this figure it is possible to identify four traps responsible for causing the BTI effect. It is also possible to see one trap responsible for the RTN effect as it suffers both capture and emission events three times on this simulation.

Figure 2.3- Trap activity impact on the threshold voltage of a PMOS transistor in time where both BTI and RTN effects are present.



As next step towards a ring oscillator we performed a simulation on an inverter chain. The inverters used were sized with w_p=135nm and w_n=90nm and L=50nm. A Monte Carlo simulation with 10000 samples was then run and the delay of a given inverter in the center of the chain was analyzed. For this case study we considered not only the trap activity but also the time zero variability. The time zero variability in the transistors occurs due to effects such as Random Dopant Fluctuation (RDF) and Line Edge Roughness (LER) and, in this work, was modeled as a Normal distribution of the threshold voltage with the standard deviation equal to 10% of the mean (CARTHIGNOL, 2008).

Figure 2.4- a) QQ-plot of the delay of an inverter considering the time zero variability and b) considering also the trap activity after 1000s of stress on nominal operating voltage (0.8V) and duty factor of 0.5.

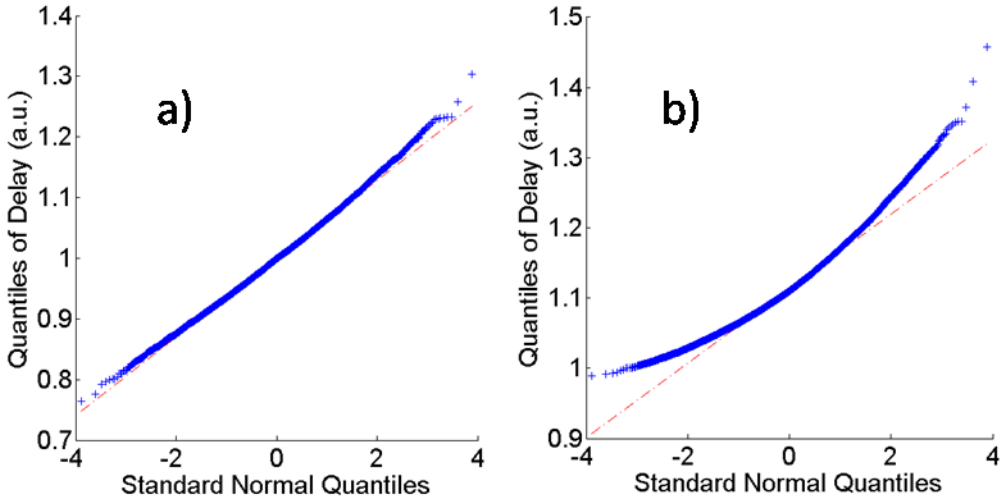


Figure 2.4 shows the results obtained from these simulations. In a) the delay of the inverter only considering the time zero variability is shown, while b) shows the delay considering also the trap activity in the simulation. Comparing both figures we see not only the increase of the delay of the test inverters due to BTI but also that the right side tail obtained in the simulation considering the trap activity decreases much more smoothly. It is important to notice that the tail of the delay distribution is of high importance since it is in that region where circuits that fail to meet the constraints are located.

Ring oscillators are simple digital structures traditionally used as the base structure for voltage controlled oscillators and for test structures aiming at the extraction of basic parameters of the technology. The inverters used in the ring oscillator simulated were sized with w_p=135nm and w_n=90nm and L=50nm.

The ring oscillator was submitted to a series of Monte Carlo analyses under different conditions. In each of these 1000 transient simulations were performed, using the enhanced electrical simulator presented on section 2.2. On each simulation, the period of the first 100 cycles was measured. The standard deviation of the period on each simulation was considered to be the signal cycle to cycle jitter, or just jitter.

First, we evaluated the impact RTN on the ring oscillator's jitter through a comparison between the initial impact of the RTN traps, and the impact of the RTN traps after 10^4s of activity. These simulations were performed both with and without considering the time zero variability. It was then performed a new set of simulations increasing the operation voltage

(Vdd) in order to adjust the oscillation period. All the simulations were performed under the conditions described above. The results are summarized at Table 1.

Table 2.1- Summary of the7 stage RO analysis.

	Vdd	Time	Avg. Period	Avg. Jitter
No Time Zero Variability	0.8	0s	0.37360ns	4.6665ps
	0.8	10^4s	0.40492ns	5.6084ps
	0.83	10^4s	0.37692ns	4.6822ps
Time Zero Variability	0.8	0s	0.37608ns	5.0754ps
	0.8	10^4s	0.40767ns	6.0284ps
	0.83	10^4s	0.37797ns	4.8370ps

The results presented on Table 2.1 show that the jitter caused by the traps cannot be neglected with its average reaching a variation larger than 1.25% of the period of the signal. When including the long-term BTI effects caused by the traps in the simulation after 10^4s of operation, we obtained an increase of 8.5% in the average period and of 18.8% in the average jitter over all the measured samples with time variability included.

The increase in the period is a natural cause of the increase of the V_{th} caused by the BTI effect. As the impact of each trap is also modeled as an increase on the V_{th} and, according to basic transistor models, it has a crescent impact on the drain current as it gets larger. The increase on the jitter is caused mainly by this increase of the impact of each trap in the transistor's drain current, which is occurs due to the increase on the V_{th} caused by the BTI. An increase of 0.3V in the operation voltage was needed to compensate the impact of the BTI on the RO. The increase in the operation voltage also reduced the jitter as it also decreased the impact of each trap in the transistor's drain current.

Figure 2.5- Cumulative Density Function of 1000 periods of a (a) 3-stage ring oscillator and (b) of a 7-stage ring oscillator under the effect of trap activity.

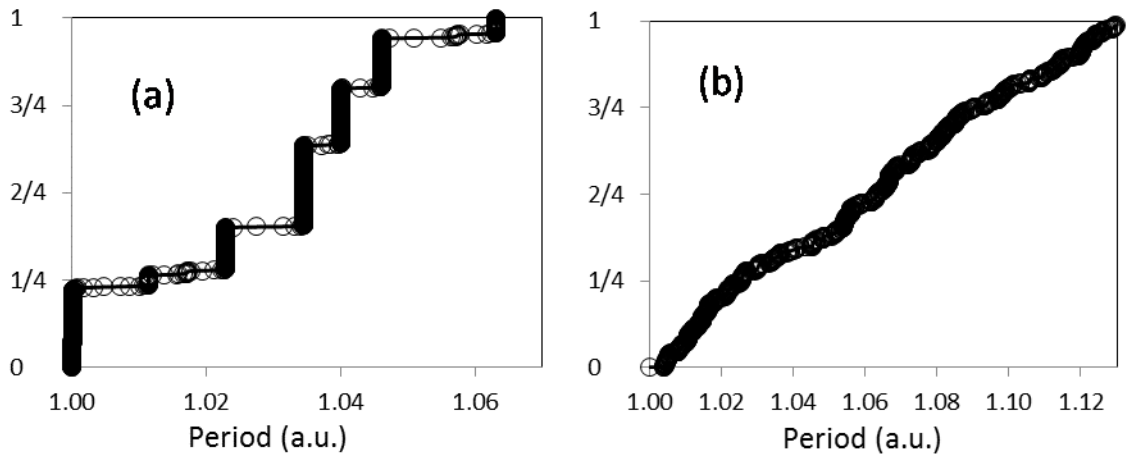


Figure 2.5(a) shows the cumulative distribution function (CDF) of oscillation period, measured in the simulation of 1000 successive cycles of a 3-stage ring oscillator with the same size as the test circuit. This simulation was performed considering that the transistors were already stressed by 10^{-4}s in the ring oscillator with a duty factor of 0.5. In this figure it is clearly seen that a few traps dominate the behavior, with clear discrete steps in the CDF. This happens because the cause of the jitter in this simulation is the trap activity which varies between discrete levels characterizing the RTN effect. For larger ring oscillators, i.e. ring oscillators formed by larger transistors or with a higher number of stages, this step like distribution is not as clearly noticed due to the increase of the number of traps affecting the signal period as the decrease of the impact of each trap. This case is shown in Figure 2.5(b) where the studied circuit was a 7-stage ring oscillator.

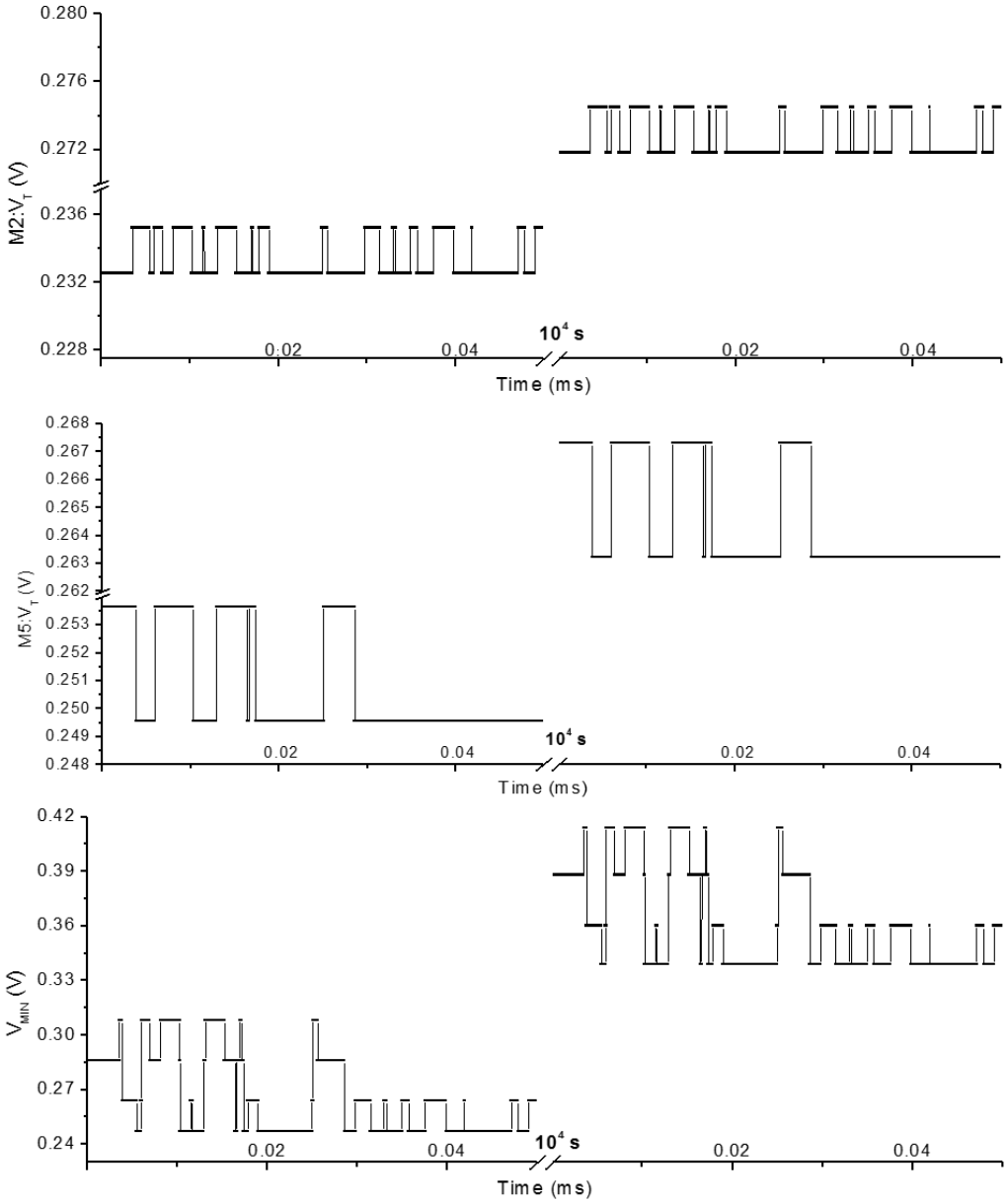
2.3.2 SRAM

A case study consisting of a simulation of the V_{min} in a SRAM cell during the stand by condition under the effect of trap activity was performed. Both the PMOS and NMOS core transistors were sized with (W/L)=(65nm/45nm) and the NMOS pass transistor were sized with (W/L)=(90nm/45nm). The V_{min} was considered the minimum supply voltage at which the design has a positive Read Noise Margin.

Figure 2.6 shows the V_{min} of a SRAM cell in a transient simulation. The transient simulation was run for two different BTI stress times. The first simulation is for a BTI stress of 100µs, and the second simulation corresponds to a stress time of 10ks. Due to the small number of traps in the devices and the constant bias, RTN trap activity was detected only on a PMOS

core transistor (M2) and on a NMOS core transistor (M5) during the transient simulation, each of these transistors showing a single RTN trap. From the BTI stress time of 100µs to 10ks, the duty cycle on the memory was 0.5, leading to a V_{th} shift due to BTI in transistor M2 of 25.7mV and of 9.96mV in M5. The duty cycle of 0.5 was chosen to avoid degrading some transistors of the cell much more than others which would generate an unbalance leading to a much larger V_{min}.

Figure 2.6- Threshold voltage as a function of time for the transistors M2 and M5 of the 6T-SRAM case study cell and how these impact SRAM V_{min}. The right hand side depicts the values after an elapsed time of 10^4 s (indicated by the break of 10^4 s shown in the time axis). During this time the cell suffered BTI, as seen in the threshold voltages and SRAM V_{min}.



To allow easy comparison between the BTI stress conditions, the same seed was chosen for the random number generator in the RTN trap simulation for both cases. In other words, RTN activity is the same in both simulations.

Comparing the V_{min} of the SRAM cell with the trap activity seen in the transistor's V_{th}, we can see that the trap activity translates its impact to the V_{min} parameter. Experimental data showing the impact of RTN on SRAM reliability is shown in (AGOSTINELLI, 2005), (TAKEUCHI, 2010), (TAKEUCHI, 2011) and (YAMAOKA, 2011). This effect is very relevant for the test of SRAM cells as it shows that a cell tested working for a given V_{min} condition might not be reliably working for the same condition after just a few instants later due to the trap activity. It is possible to see that both BTI and RTN traps have an important impact on the cell's V_{min}.

2.3.3 Combinational Circuits and SSTA

The Standard cell methodology is a method widely used for designing the digital functionalities of ASICs. The Standard cell design flow consists of creating, for instance, through the full custom method, a library containing a certain number of cells with the same height so that they can be placed one next to the other in order to form a more complex circuit. Based on these cells the designer translates a high level circuit description into a logic gate level description of the circuit limited by the cells present in the library. Once the designer has the circuit described in the logic gates level, he places the cells into a predefined grid and route the connections among them. During the entire process the designer is aided by CAD tools. One of these tools is the Static Timing Analysis (STA), which is the one responsible to verify after each step if the time constraints are still being accomplished (RABAEY, 2003).

The principle of the STA tool is to sum the delay of each cell and connection in a path (usually the critical path) of the circuit according to its position in the path. To that end each cell is characterized by its delay for different input signals with different fan-outs. This is done for different PVT (Process, Voltage and Temperature) corners. As the design advances from the circuit description to the placement and to the routing, the STA result becomes progressively more precise.

The advances in device technologies lead to transistors that are faster, cheaper, and/or with less power consuming. To convert these gains to the circuit level and finally to the system level, cells containing these devices are designed, characterized and ultimately used as the base for complex high level designs.

The downscale of the transistors present in the new technologies, together with physical limitations during the process stage, leads to new challenges that must be taken in account during the design stage. Among the main challenges there are the increase of the relative variability of the transistor physical dimensions, Line Edge Roughness (LER), effects created by the discreteness of charge and matter such as Random Dopant Fluctuation (RDF). All these sources of physical variability convert to variability in the electrical characteristics of the transistor and propagate to the higher levels of the design (ASENOV, 2003).

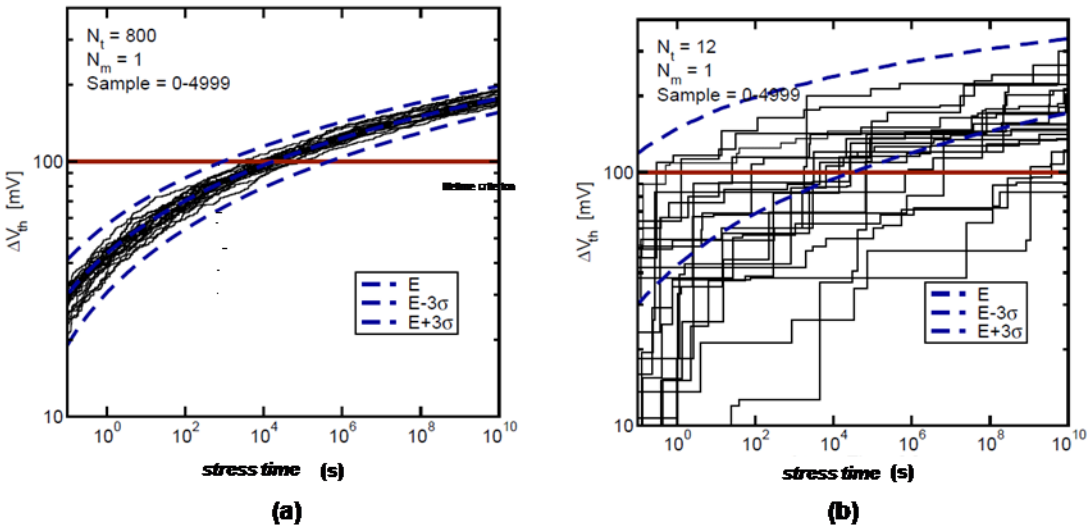
The traditional corner analysis, which consisted of selecting the worst and the best case scenarios for characterizing the cells, is not anymore able to reproduce well the high variability generated by the process of the transistors, leading to over-pessimistic designs. As the distributions become wider, in some circumstances it may be advantageous to sacrifice yield to gain in performance. To solve this problem during the design flow the STA tools were improved so that they could be able to deal with these issues and started to be called Statistical STA tools, or just SSTA.

The SSTA tools, together with its compatible characterization tool, analyze the circuit considering the parameters variability as distributions through statistical methods, such as linear sensitivity analysis (SILICON, 2007), (SILICON, 2008). They make a prediction of the probability density function of the electrical performance of the cell.

Because aging effects were not in the past considered to be stochastic, the design of the SSTA tools was originally done to consider only process variations at the starting time. These process variations are usually described following Normal distributions in parameters such as the threshold voltage and the transcondutance. This way the linear sensitivity analysis method presents a satisfactory response.

Bias Temperature Instability (BTI) is a critical reliability mechanism that affects device electrical parameters such as the threshold voltage (SCHRODER, 2003). It is caused by charging of defects in the gate oxide. In large devices of older CMOS technologies, BTI was considered to be a deterministic effect for a given circuit and stress condition. New, “atomistic” studies show that the BTI degradation (aging) in state-of-the-art technologies based on deeply-scaled transistors will lead to a huge increase in time-dependent variability (KACZER, 2010). This trend is illustrated in Figure 2.7, discussed in detail in the next section.

Figure 2.7- (a) The random properties of many defects N in large devices average out, resulting in a well-defined lifetime while (b) the stochastic nature of a handful of defects in deeply-scaled devices becomes apparent.



Source: (KACZER, 2010)

In the literature it is possible to find suggestions of how to include the Bias Temperature Instability and other aging effects in the (S)STA tools including works from (SANGWOO, 2010), (WANG, 2010), (WENPING, 2007) and (KUMAR, 2007).

We, however, could not find any commercial tool that is already using any of the approaches suggested in the literature. Due to this, in order to consider aging effects in the (S)STA tools available in the market, a ‘time’ corner is defined for the desired live expectancy of the circuit and the characterization is performed for each aged cell.

This case study presents a characterization method for the linear sensitivity analysis based SSTA tools and compare those with the result obtained with the simulation tool presented in this chapter. In order to compare the precision of the SSTA simulation method to evaluate the BTI phenomena on circuits we first characterized a standard cell library. The library chosen was the Nangate Open Cell Library for 45nm technology node used together with a PTM model for 45nm (ZAHO, 2006). Both the library and the modelcard were chosen because they are open, allowing anyone to reproduce the same results obtained in this case study.

The linear sensitivity analysis is a method used by the most of the commercial SSTA tools. It is based on a simulation of a circuit with their parameters all at nominal values and one new simulation for each random parameter in the circuit. In each of these simulations one of the random parameters is shifted while the others are kept the same. This way it is possible to see the impact of each parameter in the output of the cell individually (PARRAT, 1961). This

information together with the probability density function (PDF) information of the random parameters leads to an approximation of the PDF of the output. This method assumes the PDF of each parameter as a Normal distribution and a linear dependency between the parameter and the output which is a valid approximation when dealing with process variability issues because they are traditionally modeled by Normal distributions.

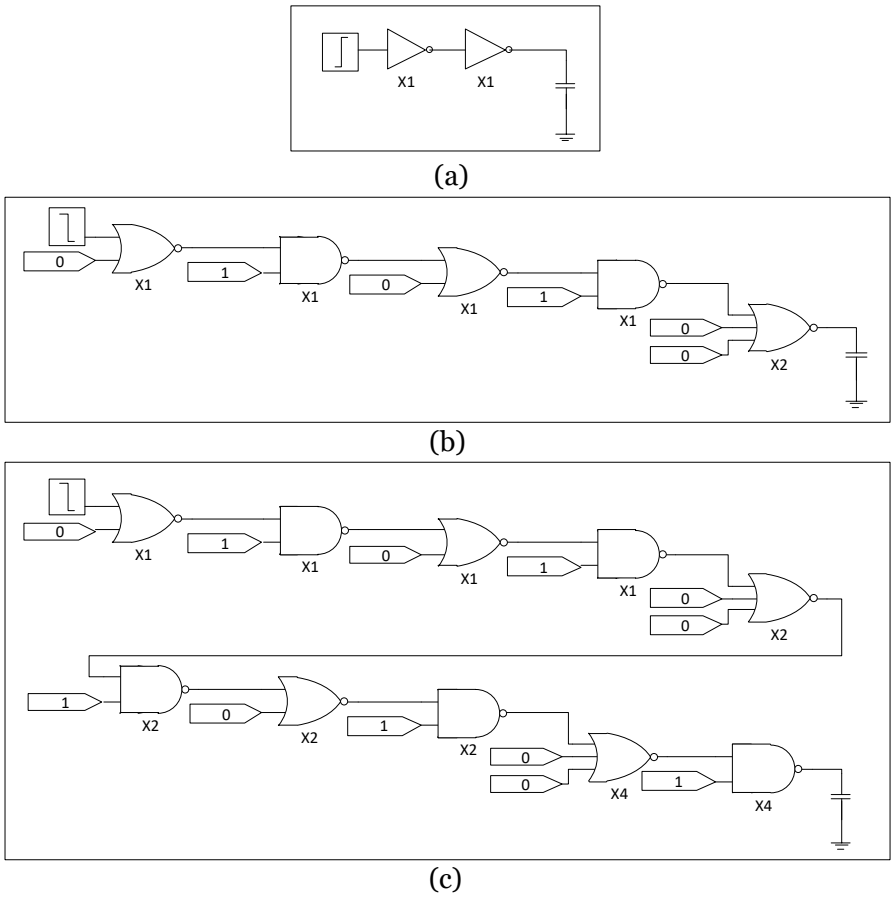
Once that the tool has evaluated the delay PDF of each cell it makes a sum of the delay on each stage finding the possible critical paths. Changes in the waveform and power consumption are also evaluated.

In order to characterize the cells to make them compatible to a linear sensitivity analysis based SSTA tool it is needed to obtain not only the Fan-out versus Slope table for the cell with its parameters with nominal values but it is also needed to obtain a new version of this table after changing each random nature parameter in the design. All this information is translated into a Liberty™ file which is later used by the SSTA tool.

Once the library is properly characterized, we define some case study circuits on which we will compare the SSTA results with the enhanced SPICE ones. In order to study the behavior of circuits with different complexity, we selected three circuits with different logic depth, they are shown in the Figure 2.8. The first one is an inverter, a simple circuit making it easy to see the effect of each trap in the general result. The second circuit is a critical path in a circuit with five logic stages. This circuit was selected because it is a typical path used in a standard cell method and is still small enough making possible to run a Monte Carlo simulation with the enhanced SPICE tool in an acceptable amount of time. The third path is ten logic stages path. This path has the objective to show the behavior of the distributions in comparison with the complexity of the path.

The third step of the process is to obtain the random parameters distribution in a transistor after the amount of time that will be defined in the time corner of the SSTA analysis. This information is used to feed the SSTA tool. In our case we defined this amount of time to be 1s, 10^4s and 10^8s. For the sake of simplicity and to have a result that can be easily interpreted, the only random variable reported on each transistor will be the Vth. In the first analysis only trap dependent variability (BTI) will be considered, on further analysis the time zero variability will be also included.

Figure 2.8- The schematic of the analyzed paths. (a) Path1, (b) Path2 and (c) Path3.



Once the mean and the standard deviation of the random parameters of the circuit are known it's possible to run the linear sensitivity analysis based SSTA and the enhanced SPICE simulation for the same circuits defined previously. The electrical simulations were performed using the Monte Carlo method with 1000 runs.

Figure 2.9 compares the delay obtained by the enhanced electrical simulation to the delay obtained by the SSTA method presented before, after a stress time of 10^4 seconds. The delay is normalized by its mean value. The result obtained by the SSTA method is a normal distribution, and hence represented by the dashed lines (straight lines) in the Q-Q plots. It is important to pay attention to the tails of the distribution, where the failing circuits are located. The Q-Q plot format helps detailing the tails of the distributions.

Looking at the tails one can see a significant difference between both methods for the case study with less complex paths, while for the more complex paths the error is less severe.

Table 2.2 presents both the Skewness and the Kurtosis of the distributions versus the complexity (measured in logic stages) and the different stress times. The Kurtosis value was chosen because is the most representative value of the weight of the tails of a distribution, while

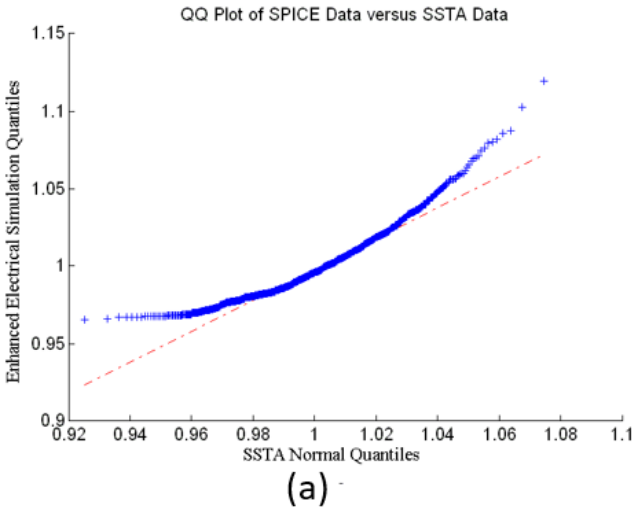
the Skewness is the measure of the PDF´s asymmetry. This table shows a tendency of the delay´s PDF to become closer to a Normal distribution as path complexity increases. The accuracy of the SSTA tool to evaluate the chance of a timing violation is correlated to the Normality of the path´s delay.

Table 2.2-Delay distributions characteristics for the three studied Paths for different stress times.

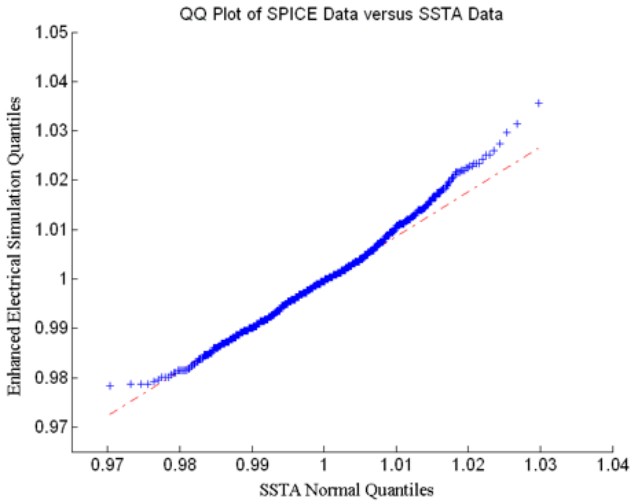
	Stress time	Skewness	Kurtosis
Path 1 (1 Logic Stage)	1s	1.5094	6.3276
	10^4s	1.1068	4.6823
	10^8s	1.2947	6.1172
Path 2 (5 Logic Stages)	1s	0.5641	3.4397
	10^4s	0.3751	3.3063
	10^8s	0.5394	3.2249
Path 3 (10 Logic Stages)	1s	0.4501	3.2773
	10^4s	0.5382	3.7532
	10^8s	0.3885	3.1765

Let us focus first on the Figure 2.9 (a), which is the one regarding to the path 1 (a single inverter). That is the case where the error of the SSTA tool is the largest. Since there is just one logic stage, the SSTA tool is given only by the linear sensitivity analysis of that given cell. The linear sensitivity analysis assumes that the input random variables are all given by Normal distributions. In the BTI case, these random variables are the threshold voltage of both the PMOS and NMOS transistors in the gate. Because these random variables are not normally distributed (RAUCH, 2007), they lead to an error in the shape of the distribution. The other important error generated by the linear sensitivity analysis is the fact that it considers the delay as a linear function of the input random variables leading to minor errors both in the mean and the variance of the result distribution. The result presented by the enhanced electrical simulator is considered to be correct since its distribution is automatically generated based on experimental data for the distributions of the trap properties and in first principles.

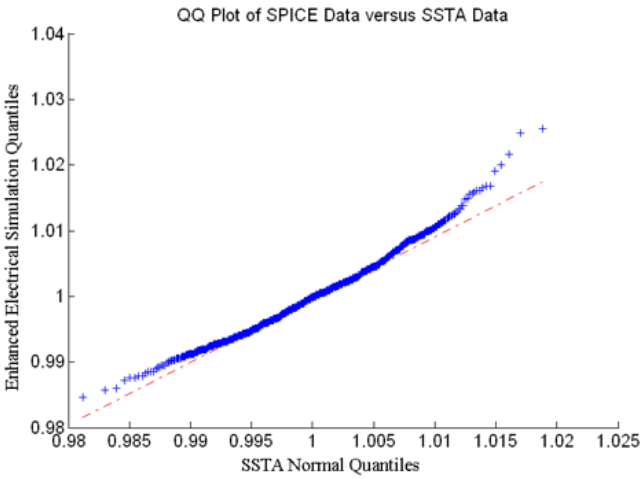
Figure 2.9- The Q-Q plot of the normalized delay generated by the electrical simulation in comparison with the one generated by the SSTA analysis in the Path 1after 10^4s (a), in Path 2 after 10^4s (b) , and in Path 3 after 10^4s (c).



(a)



(b)



(c)

Another series of simulations were run to evaluate the impact of the time zero variability in the study. Time zero variability is usually modeled by adding to the Threshold voltage a Normal distributed variable. The mean of the distribution is zero and the standard deviation is a function of the technology and the area of the transistor. In this case the standard deviation is assumed to be 10% of the threshold voltage for a transistor with the minimal dimensions for the technology (HORSTMANN, 1998). The standard deviation was considered to be inversely proportional to the area of the channel of the transistor.

Figure 2.10- Q-Q plots of the delays on paths 1 (a) and 2 (b) after 10^4s of stress considering time zero variability.

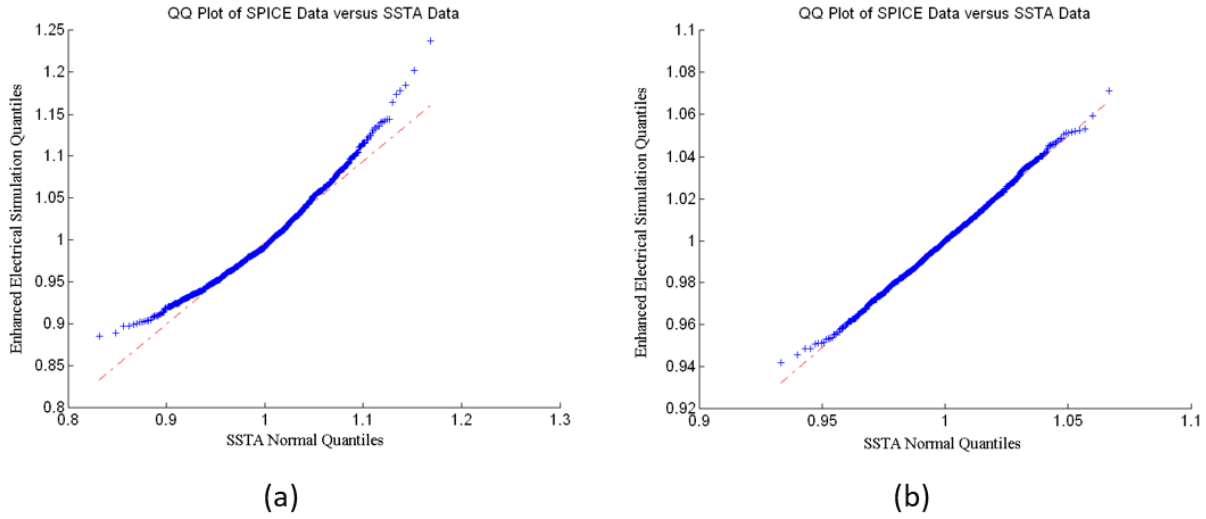
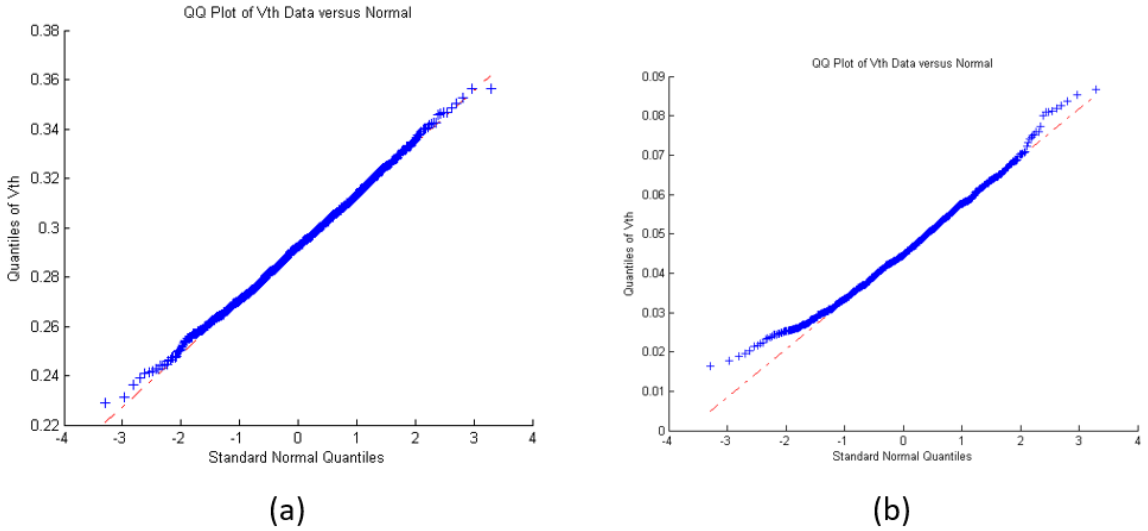


Figure 2.11- Threshold voltage distribution due to NBTI considering (a), and not considering (b) time zero variability.



The simulation of the delay on path 1 and path 2 after 10^4s is presented on Figure 10. There it is possible to see a smaller difference between the results obtained by the electrical simulation and by the SSTA method. This happens because the PDF of the threshold voltage becomes the sum of two distributions, one of those is a Normal, making the threshold voltage PDF more Normal like and then more suitable for studies using the linear sensitivity analysis method. Figure 2.11 shows the distribution of the threshold voltage due to NBTI in a transistor considering (a) and not considering (b) the time zero variability.

Despite of the error for path 1, paths 2 and 3 presented a result much closer to the ones obtained by the SSTA method. It is clear an increase in the accuracy of the method with the increase of the complexity of the path being studied. This can be understood if one considers that the Central Limit Theorem (CLT) of the Statistics, which states that "that if S_n is the sum of n mutually independent random variables, then the distribution function of S_n is well-approximated by a certain type of continuous function known as a normal density function (GRINSTEAD, 1997). This theorem is applicable to the SSTA case since the total delay is given by the sum of the delays of each stage, which are considered to be mutually independent random variables.

Two other properties of this system of sum of mutually independent random variables are presented in equations (2.5) and (2.6). These properties are valid no matter the shape of the PDF of the random variables.

$$\mu_{S_n} = \sum_{i=1}^{n} a_i \mu_i \tag{2.5}$$

$$\sigma_{S_n}{}^2 = \sum_{i=1}^{n} a_i{}^2 \sigma_i{}^2 \tag{2.6}$$

Based on these properties we can tell that for a long path the delay will tend to be Normally distributed with its mean and variance given by the equations (2.5) and (2.6) regardless of the shape of the individual cell delays.

SSTA tools are used for two main reasons, to look for setup time violations and hold time violations. The most probable paths to do these violations are the long paths for the setup time and the short paths for the hold time. BTI effects are responsible foran increase of the threshold voltage which leads to a larger delay in the logical path, changes in the flip-flops parameters and an increase in the variability. The increase of the delay of the logic path increases the probability of a setup time violation and reduces the chance of a hold time

violation. The change in the flip-flop parameters changes with the circuit's topology so we cannot present a general case of the BTI's impact on those. For setup time violation analysis, it is important to consider the critical paths, which are the paths with the largest propagation delays, so usually are the ones with the largest number of logic stages. This way we can assume that when simulating paths to look for setup time these paths are going to fulfill the long path condition.

2.4 Conclusions

A circuit simulation method capable of considering the trap activity during transient electrical simulations was presented. Because it is directly obtained from detailed defect (trap) studies and adjusted with experimental data, this approach leads to an accurate tracking of the trap states during the simulation. This allows the analysis of the impact of trap activity in arbitrary circuits. Because the trap kinetics are included in the transistor model, it is possible to use this method with any electrical simulation tool and it also allows the user to extend this methodology to include other time dependent effects. A case study on a ring oscillator demonstrates the impact of the traps on both jitter and BTI, a case study on a SRAM cell shows how critical the trap activity can be on the reliability of SRAM cells as it affects the V_{min} of the cell, and a case study on combinational circuits and SSTA tools shows where these are adequate to use on BTI effects and where they are not reliable. Other case studies performed but not presented here includes (KACZER, 2011) and (RODOPOULUS, 2011).

The overhead in the runtime of the simulation method presented here is slightly larger than a traditional electrical simulation. Overall, this simulation methodology proves to be an efficient tool for analyzing both RTN and BTI. The models used in these simulations, however, are based only experimental data which are very scarce in the literature limiting the conditions to which the simulations can be run. An alternative to experimental data is generating the data using TCAD simulations, but also in this case there are only a few case studies for specific cases and only for n-type transistors. As a step forward to solving this lack of tools in the scientific community, an EMC device simulation tool for modeling p-type transistors was developed as part of this thesis work and is described in the next chapter.

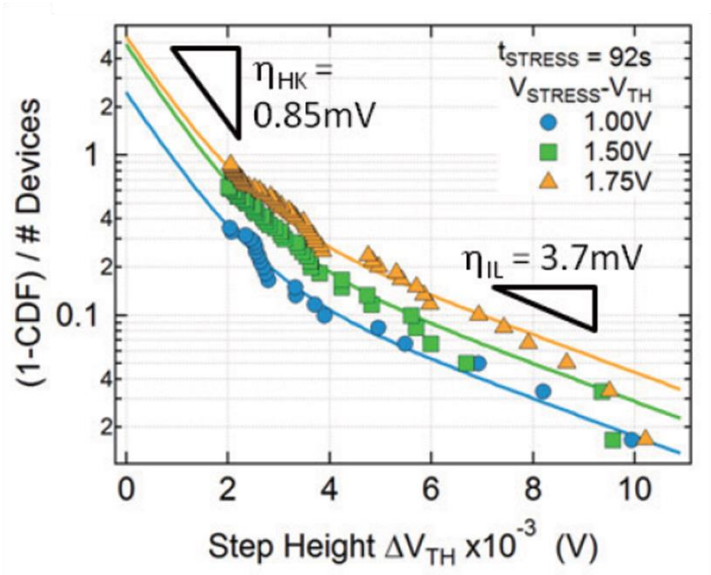
3 MONTE CARLO SIMULATIONS ON PMOSFETS

The previous chapter presented an electric simulation tool which helps to evaluate the impact that the trap activity causes in circuits through transient simulations. The relevance for translating technology characteristics related to the trap activity to circuit level variables used during the design is highlighted in chapter 2. All the simulation tools of this kind, not only the one developed in this thesis, require two inputs. They are: i) the information regarding the circuit to be evaluated and the desired output and ii) the technology information including the trap characteristics for that given technology node.

The trap related characteristics include the probability distributions of their time constants, their impact on the threshold voltage and how these correlate with themselves and with other environmental parameters of the circuit such as the temperature and the electrical field. All these characteristics are usually obtained from experimental data. However, even for simple trap models these are hard to obtain due to the statistical and time dependent nature. This issue becomes an important limitation for using the simulation tools as the ones presented in chapter 2. In the case studies presented in section 2.3, experimental data obtained from IMEC and from other sources in the literature were used, as referenced.

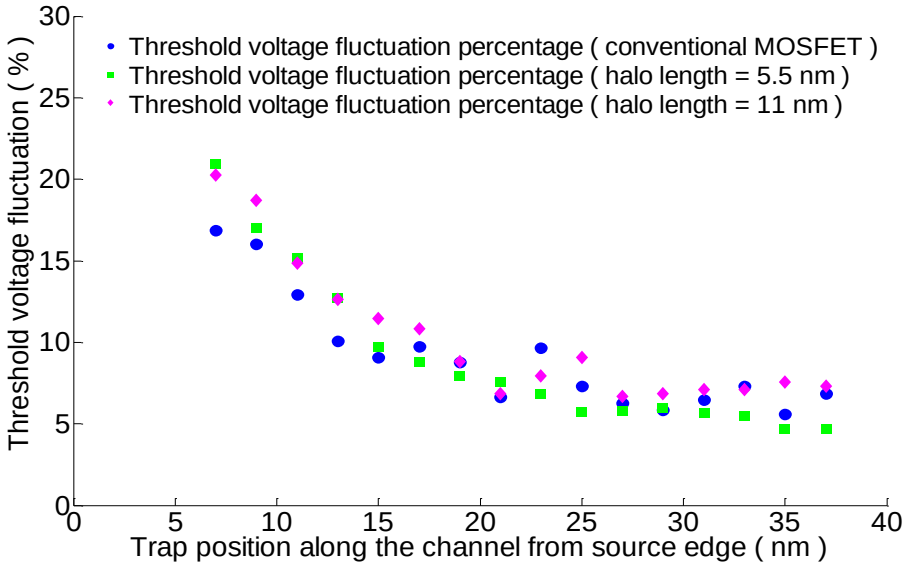
Despite of a limited number of studies, the trap characteristics dependency on the device technology is evident. Considering the trap impact on the device's threshold voltage, for instance, Toledano-Luque (2012) shows two different distributions that define the trap impact comparing the measurements made with traditional devices and with high-k as seen in figure 3.1. Ashraf (2011) shows the changes on the traps impact on the threshold voltage due to halo implantations as seen in figure 3.2. Furthermore, due to the tight relation of the impact of traps with percolation paths in the channel, one can also assume that the impact also depends strongly on the doping profile of the device near the semiconductor-dielectric interface, which is strongly affected by the doping density of the channel and V_{th} correction mechanisms. All these dependencies limit the trap characteristics to be valid only for the technology in which it was measured. All simulation tools which have trap characteristics as an input are limited to technologies which the user has access to extract that information by measurements, by simulation or both.

Figure 3.1- (1-CDF) of step heights normalized to the number of tested nFETs after positive stress. Data can be fitted with a bimodal distribution with η_{IL} = 3.7mV and η_{HK} = 0.9mV. Note that η_{IL} is similar to the η value obtained in pFETs.



Source: (TOLEDANO-LUQUE, 2012)

Figure 3.2- Relationship between the position of the trap along the channel length and the impact that a captured charge causes in the threshold voltage for devices with and without halo.



Source: (ASHRAF, 2011)

TCAD tools can be used to obtain the carriers impact on the threshold voltage without the need of performing experiments. Previous works performed at University of Glasgow, including (BUKHORI, 2010), (WANG, 2012), (AMOROSO, 2012), (AMOROSO, 2013) and (AMOROSO, 2014), and works performed at ASU as the one performed by ASHRAF (2011), show the capability of this kind of tool to evaluate the impact that a charge trapped in the

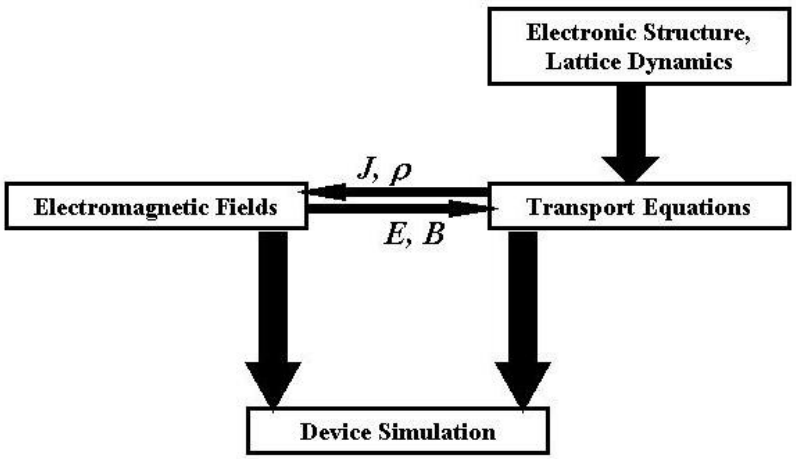
dielectric has on the device performance. Even though the computational cost of obtaining this information using TCAD tools is high, it gives the user the opportunity to obtain some of the trap characteristics for the device of interest without the need of experimental data, even before manufacturing the device. This may allow optimizing performance and reliability of future technologies. An additional upside of TCAD simulations is that they give to the user access to internal information of the device variables which one cannot access by experiments. This enables detailed studies aimed at further understanding the impact of traps. In ASHRAF study in (2011), for instance, the trap position was swept along the channel providing a better understanding on how the trap impact is as a function of its location.

A prototypical device simulation tool is, traditionally, built by the self-consistent solution of two sets of equations: the transport equations for charge flow and the Maxwell equations for the electromagnetic fields. For low frequency devices the Maxwell equations can be reduced to the Poisson equation shown in Eq. (3.1),

$$\frac{d^2\psi}{dx^2} + \frac{d^2\psi}{dy^2} + \frac{d^2\psi}{dz^2} = -\frac{\rho(x,y,z)}{K_s\varepsilon_0}, \tag{3.1}$$

Where ψ is the electrical potential, ρ is the free charge density and $K_s\varepsilon_0$is the material permittivity. Within the semi-classical limits, the equation which defines the carriers transport in a semiconductor is given by the Boltzmann transport equation (BTE). In a device simulation these two equations are self-consistently coupled as shown in Figure 3.3.

Figure 3.3- Device simulator Framework showing the coupling of the transport equations with the Electromagnetic fields calculation.



Source: (VASILESKA, 2010)

As shown previously, the trap impact on the device is highly dependent on the physical structure of the device, so the structure being simulated should be carefully matching the technology of interest. Among the physical characteristics of the device, one that should be highlighted is the doping profile in the active channel area as it is strongly responsible for the generation of percolation paths in the channel region. The percolation paths are generated by the random position of doping atoms in the channel generating a non-uniform potential landscape. To account for the random position of charges in the channel it is important for the TCAD tool used to be atomistic. The percolation path profile generates a randomness in the impact of a trap in a device, allowing a trap in the same channel position in two devices of the same technology to have different impact (ASENOV, 2003).

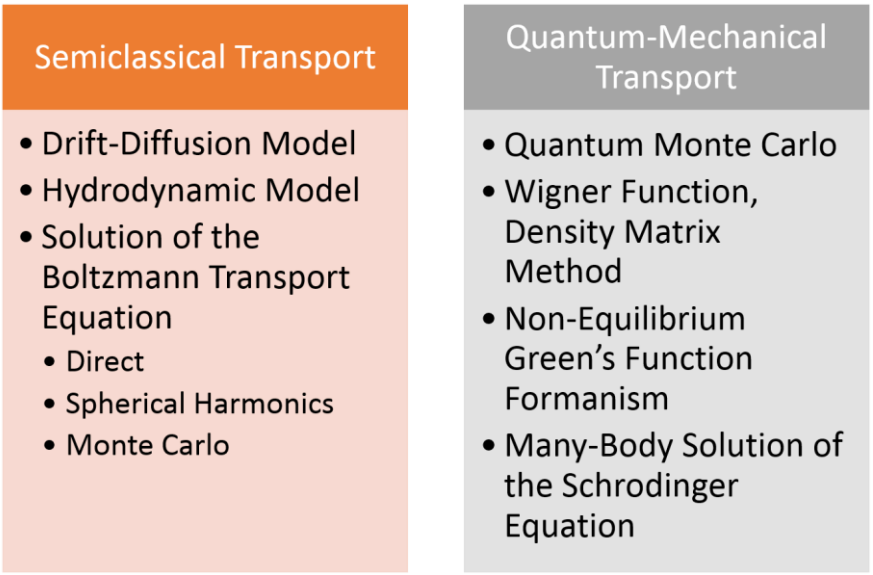
The trap impact is also known to gain relevance for the smaller devices as a consequence of having a smaller total charge in the gate capacitor. This relation also forces the TCAD tool to be suited for simulation in small devices. When dealing with small devices other effects become relevant, such as the velocity overshoot and the ballistic transport (LUNDSTROM, 2002). The TCAD tool used should be able to take those into account properly.

Figure 3.4 shows a hierarchical map of the methods used to solve the corresponding transport equation in device simulators. Shown on the left side are the semi-classical methods which include the drift-diffusion model, the hydrodynamic model and the Monte Carlo method for direct solution of the Boltzmann transport equation. For completeness, on the right side are shown quantum-mechanical methods for the analysis of ultra-nanoscale devices. It is important to highlight that corrections for quantum effects may be included in the semi-classical approaches. These include quantum mechanical tunneling and quantum-mechanical space quantization effect. When using a TCAD tool which includes a correction model, it is important to remember that in the correction the effect is not described by its physical nature, but emulated by a mathematical model fitted for that given case.

In order to be able to cope with stochastic nature of the trap impact, any analysis demands several samples to generate a reliable result. When choosing the method to solve the transport equation, it is important to weight not only the limitations of each method, but also their computational cost so that a sufficiently large number of simulations can be run. The drift-diffusion and the hydrodynamic models, despite of having a small computational cost, are not capable of taking into account the atomistic nature of the charged carriers and the impurities or the ballistic transport. Methods based on quantum mechanics, on the other hand, are for the most part unnecessarily complex. The Ensemble Monte Carlo device simulation method, in spite of having a high computational cost, is the most accurate among the semi-classical

methods and can be used for 3D device simulations with atomistic dopants on real positions that are randomly defined. The Ensemble Monte Carlo device simulation method is the method chosen for the analysis in this thesis work.

Figure 3.4- Semi-classical and quantum-mechanical transport approaches.



The Ensemble Monte Carlo (EMC) device simulation method nowadays is a well established tool for simulating transport in semiconductor devices. The Monte Carlo method is attributed to Fermi, Von Neumann, and Ulam, during their research in Los Alamos for the construction of the atomic bomb, even though there are indications of previous uses of the technique dating back to the beginning of the century (JACOBONI, 2002). In large semiconductor devices, the high computational cost and the large size of transistors, however, limited its use mostly to analyzing bulk material properties as alternative methods like drift-diffusion were at that time adequate and much faster. For the state-of-the-art devices the need for the EMC device simulation method was created; however, there are no commercial TCAD tools that implement the EMC device simulation method that the author is aware of. The data regarding the impact of traps in devices presented in the literature were generated by several groups that developed their own code to simulate these effects and in all cases considered they referred to n-channel devices. It is important to highlight the material and the carrier type being considered (electrons or holes), because in a simulation code based on the EMC method, the simulation depends strongly on the band characteristics. In the case of electrons, we refer to the properties of the conduction bands and in case of holes to the properties of the valence bands. The energy bands in which the carrier of interest is located affect the simulation both when defining the scattering rates of the carriers as well as when defining the relationship

between the energy and the wave vector of each carrier. To the best of the knowledge of the author, there is not another work that has developed a code with the required complexity needed to simulate the impact of traps in pMOSFETs. Hence, there is a deficiency of studies on the impact of traps at device level. To target this deficiency, in this work a device simulator based on the EMC method for p-type silicon devices was developed.

In the rest of this chapter, the device simulator based on the EMC method is discussed in details. Initially, the basic algorithm for Monte Carlo simulations in a bulk material is described. At this point the equations derived from the band structure, like the ones that evaluate the scattering rates and the energy-wave vector relationship, are presented. Comparisons between the results obtained with the developed algorithm and data from the literature is given that shows the validity of the approach. Afterwards, the bulk EMC code is extended into a 3D particle-based device simulator. Once again comparisons between the simulation results obtained with the code and data from the literature and other TCAD tools were made, thus showing the correctness of the tool.

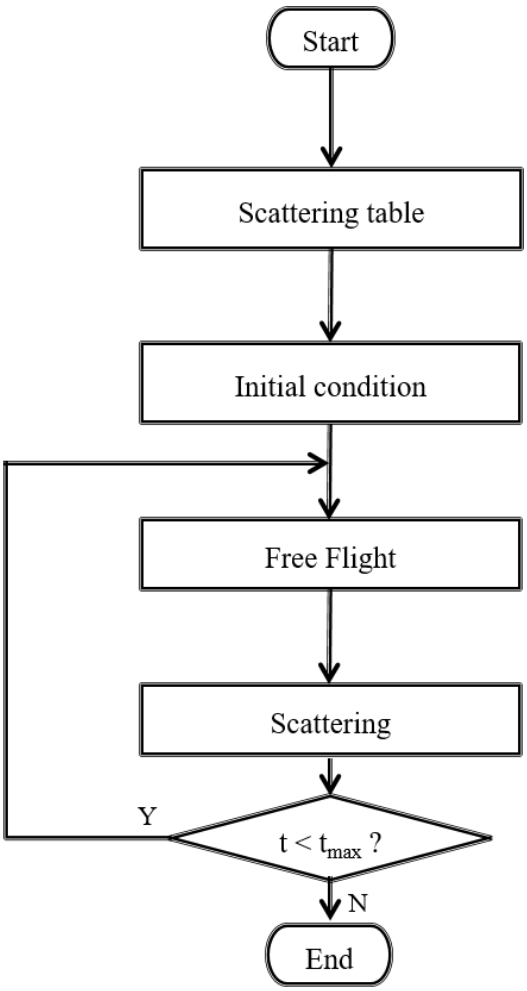
3.1 Bulk Monte Carlo

A simulation method based on the Monte Carlo algorithm consists of tracking one or more carriers for a given amount of time in which the carrier(s) go through a series of free-flight periods (where they accelerate due to the electrical field) that are followed by scattering events. For a bulk material simulation, one can track a single carrier for long enough time and obtain the characteristics of interest for that material regarding the carrier's transport (this is called single particle Monte Carlo algorithm). When a large number of carriers is tracked, the simulation may be run for a shorter time and transient effects, such as the velocity overshoot, can be observed. When several carriers are tracked simultaneously, the simulation is called Ensemble Monte Carlo (EMC). Both the single particle and the EMC simulation consist of a free-flight-scattering loop which repeats itself until the simulation time is reached. Statistics is gathered at the so-called observation times at which the ensemble is frozen and average particle velocity and average particle energy are being calculated. There is no specific physical limit on the length of the observation times for bulk simulations. Typical values are on the order of 10s of femtoseconds to get good resolution in the time domain. The duration of the simulation depends on when the system reaches steady-state and is on the order of 5-10 picoseconds. Hence, there are many observation points during the length of the simulation.

The time between scattering events, where the carrier accelerates according to the applied electric field (Newton's law), is called free-flight and has its length defined by the

maximum cumulative scattering rate of all the scatter mechanisms present in the given band. The scattering rates, in turn, are a function of the band structure, the presence of other possible bands and the energy of the carrier. In a given band or valley, for each mechanism and for different carrier energies the scattering rate is calculated and a table is generated. The process is repeated for all possible bands and valleys for that kind of carrier in that specific material thus generating a set of tables. Those are called scattering tables.

Figure 3.5- Bulk Monte Carlo simulation flowchart.



Source: (VASILESKA, 2010)

A Monte Carlo code for bulk materials has a flowchart presented in figure 3.5. In the case of a single carrier simulation the carrier is initialized with a given energy and wave vector. The initialization of the carrier is not critical as the result extracted by the simulation will be the one obtained after a long period of time. Thus, transient characteristics cannot be calculated. For an EMC bulk simulation, the flowchart becomes slightly different with the addition of an extra loop where the number of carriers is swept. As in the case of an EMC simulation one

might be interested in the transient results, such as the velocity overshoot effect, it is important to initialize the carriers properly.

The subsequent subsections detail the initial condition calculation, the scattering rates calculation for holes in silicon and the scattering table generation steps. The detailed free-flight-scattering routine will be explained and a comparison between the results obtained in this work with experimental data and other simulation work presented in the literature will be given.

3.1.1 Initial Condition

The initial condition of each carrier is defined assuming Boltzmann statistics, where the initial kinetic energy of the carrier is calculated by equation 3.17

$$E = -\tfrac{3}{2} k_B T \ln(rand1) \tag{3.2}$$

where*rand1* is a random number uniformly distributed between 0 and 1.

The direction of crystal momentum is determined by two additional random numbers, *rand*2 and *rand*3, which specify the azimuthal

$$\varphi = 2\pi \cdot rand2 \tag{3.3}$$

and the polar angle:

$$\cos\theta = 1 - 2 \cdot rand3 \tag{3.4}$$

$$sin\theta = \sqrt{1 - \cos^2\theta} \tag{3.5}$$

Having defined the energy and the azimuthal and polar angles, the absolute value and the k_x, k_y and k_z components of the carrier's wavevector can be calculated. The equation that relates the absolute value of the wave vector with the energy is dependent on the band in which the carrier is located. The equations used in these step will be presented in section 3.1.2. It is important to highlight that in the code developed here it is assumed that all the carriers are initially in the heavy-hole band.

Having calculated the absolute value and the angles of the wave vector, the *x*, *y*, and *z*-components of the wave vector are then obtained using a Spherical to Cartesian coordinate transformation.

$$\begin{cases} k_x = k\sin\theta\cos\varphi \\ k_y = k\sin\theta\sin\varphi \\ k_z = k\cos\theta \end{cases}. \tag{3.6}$$

3.1.2 Scattering Rates

The scattering theory is based on the Fermi's Golden Rule (3.3) which is obtained by solving the time-dependent Schrödinger equation and using first-order time-dependent perturbation theory. Fermi's Golden Rule defines the scattering rate (transition probability) from initial state k to final state k'. In other words,

$$S(\boldsymbol{k},\boldsymbol{k'}) = \frac{2\pi}{\hbar}|M(\boldsymbol{k},\boldsymbol{k'})|^2\delta(E_{k'} - E_{k'} \mp \hbar\omega) \tag{3.7}$$

Where S(**k**,**k'**) is the transition probability for scattering between some initial state **k** to some final state **k'**. In equation (3.7), E_k and $E_{k'}$ is the corresponding initial and final kinetic energy, and $\hbar\omega$ is the phonon energy. The matrix element $|M(k,k')|^2$ is characteristic of the scattering process which is being evaluated and contains the momentum conservation, while $\delta(E_{k'} - E_k \mp \hbar\omega)$ describes the conservation of energy during the scattering process.

The total scattering rate out of state **k** for a given process is obtained by summing over all k' states. The equation (3.4) is used to calculate the scattering rate out of state k as a function of the absolute value of the carrier's wave vector (XIAOJIANG, 2000).

$$W(k) = \frac{\Omega}{(2\pi)^3}\int_0^{2\pi} d\phi \int_0^{\pi} sin\theta d\theta \int_0^{\infty} S(k,k')k'^2 dk' \tag{3.8}$$

In Eq. (3.4), Ω is the crystal volume, θ is the polar angle and ϕ is the azimuthal angle.

The scattering processes are stochastic events and have their rates calculated quantum mechanically. Even though carriers do not scatter due to the purely periodic potential associated with the array of ions building a crystal, they may scatter due to vibrations on this array causing disturbances in the crystal potential. The carrier scattering caused by the vibrations in the crystalline lattice is called phonon scattering. There are two types of phonon scattering, the acoustic and the optical, and are defined by the types of displacement of neighboring atoms in the lattice. When two atoms are displaced in the same direction (only the amount of displacement differs), we call this phonon mode acoustic mode. When two neighboring atoms are displaced in the opposite direction we talk about optical modes of vibration.

The matrix elements for the acoustic phonon and non-polar optical phonon scattering mechanisms are given by equations (3.5) and (3.6) respectively. By substituting these equations

into (3.3), and then integrating as shown in (3.4), the scattering rate out of state k for the given scattering mechanism is calculated (RAGHURAJ, 2011).

$$|M(\boldsymbol{k},\boldsymbol{k'})|^2 = \frac{{E_{ac}}^2 k_B T}{\rho\Omega {v_s}^2}\delta(\boldsymbol{k}-\boldsymbol{k'}\pm\boldsymbol{q}) \tag{3.9}$$

Where E_{ac} is the acoustic deformation potential constant which is experimentally determined, ρ is the density, $\mathbf{q} = \mathbf{k} - \mathbf{k'}$ is the phonon wave vector and v_s is the sound velocity in the crystal. In deriving this result, it is assumed that acoustic phonon scattering process is elastic. Also, at room temperature, one can invoke equipartition approximation and combine absorption and emission processes into a single scattering process.

For non-polar optical phonon scattering:

$$|M(\boldsymbol{k},\boldsymbol{k'})|^2 = \frac{\hbar {D_{iv}}^2 q^2}{2\rho\Omega\omega_0}\left(n_0 + \frac{1}{2} \mp \frac{1}{2}\right)\delta(\boldsymbol{k}-\boldsymbol{k'}\pm\boldsymbol{q}) \tag{3.10}$$

Where D_{iv} is the optical deformation potential constant which is experimentally determined and n_0 is the number of phonons in a given mode. The minus/plus sign corresponds to the phonon absorption/emission process.

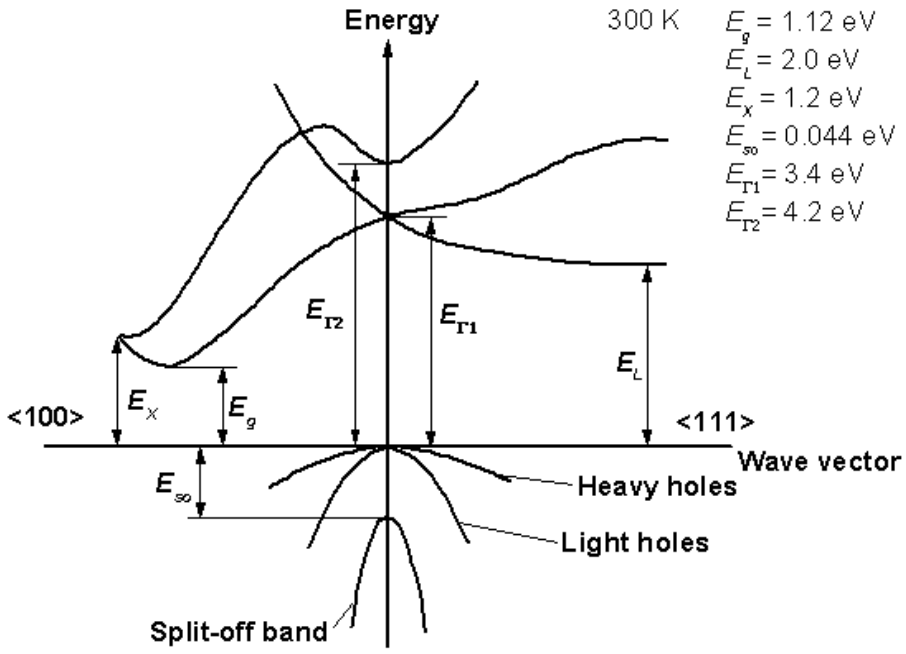
When performing the integration one has to consider the proper relation between the wavevector and the energy. This relation is a unique characteristic of each band. In the Monte Carlo simulation the transport of the holes is modeled using the three band model: heavy, light and split-off bands. For silicon the split off band is 44mV below the top of the valence band. The highest energy point is the Γ point as shown in figure 3.6. Thus, there is just one valley for each band.

In the case of developing a code for electrons in silicon, the band used should be the conduction band, also shown in figure 3.6. In the case of the conduction band for silicon, its minimum point is not in the Γ point, but in a different point close to the X point, which corresponds to the <100> direction. Since there are 6 equivalent <100> directions, there are six equivalent valleys and the electrons can scatter among them. There is, however, just one band that must be considered since the minimum of the second band is 880meV higher than the minimum of the lowest conduction band. Rarely there will be an electron with enough energy to populate that band, in particular for ultra nano-scale devices where the supply voltages are on the order of 0.9 V.

A particle in an energy band can be described via the wave vector dependence on the energy (dispersion relation). Depending upon the energy, the band can be considered to be parabolic or non-parabolic. For many first order studies, the bands can be considered to be parabolic, however for higher energies one must consider the non-parabolicity effects.

The energy band can also be described by the iso-surfaces of the wave vector which have the same energy (also known as constant energy surfaces). By this characteristic a band can be spherical, elliptical or warped. It all depends upon the mass of the particle in different crystallographic directions. Only for the case when the mass is isotropic we have spherical bands.

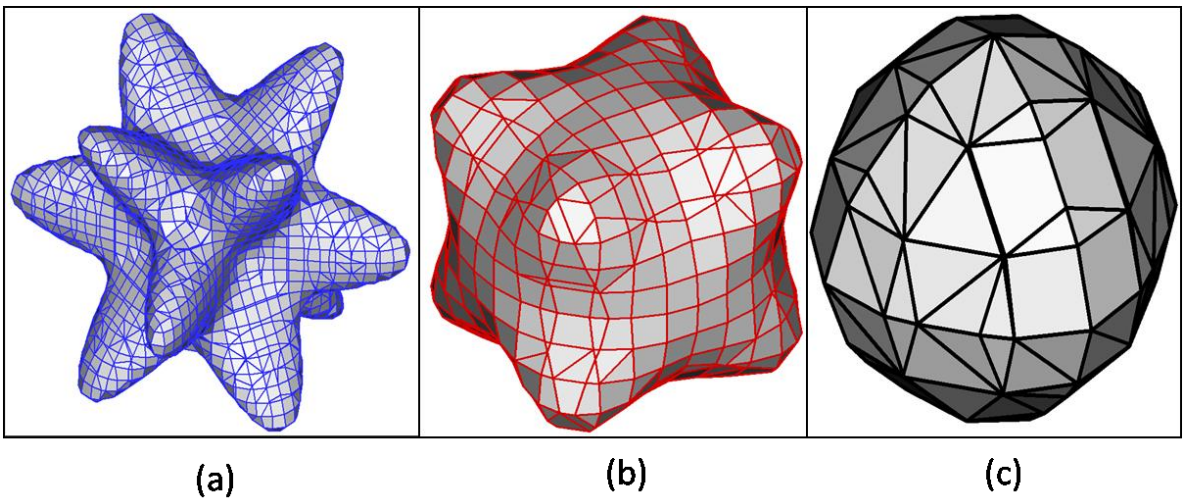
Figure 3.6 Silicon band structure including the Valence bands showing the Heavy (HH), light (LH) and split-off (SO) bands.



Source: (Ioffe.ru)

The split-off band can be considered to be spherical and parabolic, however the heavy and light hole bands are warped and non-parabolicity effects must be included. The warping of the bands is depicted in figure 3.7. In the heavy and light hole bands the warping is included by $g(\theta, \phi)$ described in equation (3.7) where A= -4.22, B= -0.78, and C=4.80 (GAGLIANI,1975).

Figure 3.7 – Iso-surfaces of energy for the (a) Heavy-hole band, (b) Light-hole band and (c) split-off band.



Source: (KRISHNAN, 2005)

$$g(\theta,\phi) = \left[\left(\frac{B}{A}\right)^2 + \left(\frac{C}{A}\right)^2 (sin^2\theta\, cos^2\theta + sin^4\theta cos^2\phi sin^2\phi)\right]^{1/2} \quad (3.11)$$

The non-parabolicity of the bands is considered by the $(1-\beta)$ factor presented by Dewey (1993) in equations (3.9) and (3.10) for the heavy and light hole bands, respectively. Including both the warping and the non-parabolicity of the bands, the energy-wave vector relationship is given by equation (3.11).

$$1-\beta = a_2E^2 + a_1E + a_0 \quad (3.12)$$

For heavy holes the a_0, a_1 and a_2 constants are given by:

$$1-\beta = \begin{cases} 1 - 43E + 1200E^2 & 0 \le E < 0.015eV \\ 0.8006 - 14.3632E + 177.3117E^2 & 0.015 \le E < 0.040eV \\ 0.6434 - 4.3521E + 25.2716E^2 & 0.040 \le E < 0.100eV \\ 0.46087 & E \ge 0.100eV \end{cases} \quad (3.13)$$

For light holes the a_0, a_1 and a_2 constants are given by:

$$1-\beta = \begin{cases} 1 - 14.419E + 106.7E^2 & 0 \le E < 0.050eV \\ 0.7946 - 6.296E + 26.37E^2 & 0.050 \le E < 0.100eV \\ 0.4286 & E \ge 0.100eV \end{cases} \quad (3.14)$$

$$E(k) = (1-\beta)\left(\frac{|A|\hbar^2 k^2}{2m_0}\right)\left(1 \pm g(\theta,\phi)\right) \tag{3.15}$$

The +, - signs in (3.11) are for phonon heavy and light hole bands respectively. Note that for the split-off band, the non-parabolicity is not considered (β=0) nor is the warping effect ($g(\theta,\phi) = 0$). Hence, the energy relation with the wave vector is quadratic.

The energy equation regarding each band is substituted into the equation 3.4. The integral over k is solved using the relation between dE and dk defined in the equation 3.12, but due to its warping the integral over θ and ϕ cannot be analytically solved. The scattering rates are calculated by the equations (3.13) for the acoustic phonon scattering, and (3.14) for the non-polar optical phonon scattering. These equations are solved numerically in the code to generate the scattering tables. The results obtained with those are presented in Figure 3.8, 3.9, 3.10 and 3.11.

$$2kdk = \frac{2m_0}{|A|\hbar^2}\frac{1}{1\pm g(\theta,\phi)}\frac{dE}{(1-\beta)}\left(1 + \frac{2a_2E + a_1}{(1-\beta)}\right) \tag{3.16}$$

$$W(E) = \frac{E_{ac}^2 k_B T}{(2\pi)^2\hbar 4\rho v_s}\left(\frac{2m_0}{|A|\hbar^2}\right)^{3/2}\frac{\sqrt{E\prime}}{(1-\beta(E'))}\left(1 + \frac{2a_2E\prime + a_1}{(1-\beta(E'))}\right)\int_0^{2\pi}\int_0^{\pi}\frac{sin\theta d\theta d\phi}{\left(1\pm g(\theta,\phi)\right)^{3/2}} \tag{3.17}$$

$$W(E) = \frac{D_0^2(2m_0)^{\frac{3}{2}}N_{OP}}{(2\pi)^2 2\rho\omega_0\hbar^3 A^{\frac{3}{2}}}\frac{\sqrt{E\prime}}{(1-\beta(E'))^{3/2}}\left(1 + \frac{2a_2E\prime + a_1}{(1-\beta(E'))}\right)\int_0^{2\pi}\int_0^{\pi}\frac{sin\theta d\theta d\phi}{\left(1\pm g(\theta,\phi)\right)^{3/2}} \tag{3.18}$$

Where:

$$E' = E - \hbar\omega - \Delta E_{ij} \tag{3.19}$$

And

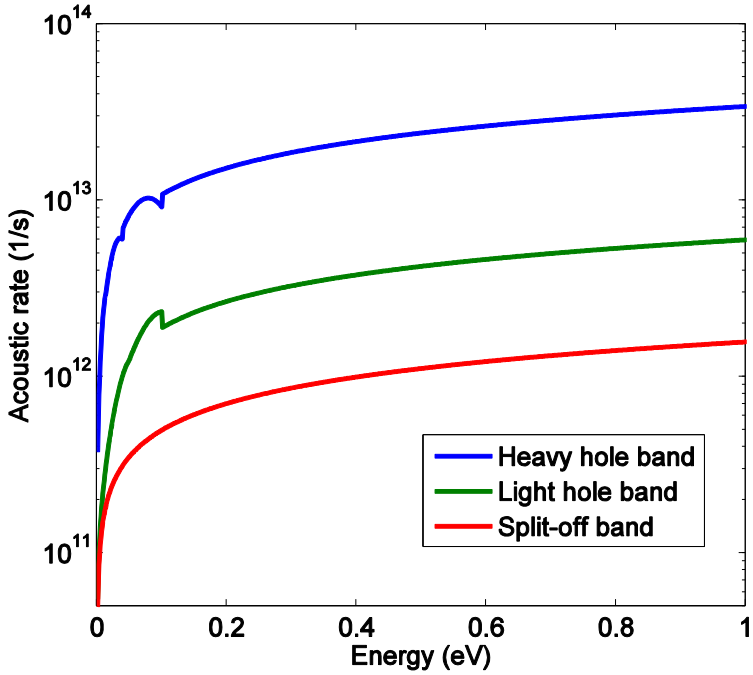
$$N_{OP} = \left(n_0 + \frac{1}{2} \pm \frac{1}{2}\right) \tag{3.20}$$

The +, - signs in Eq. (3.16) are for phonon emission and phonon absorption process, respectively.

The scattering rates of each process depend on the band in which the carrier will be found after the scattering process. In the acoustic phonon process, the carrier does not change bands and the process is considered to be elastic. Figure 3.8 shows the scattering rates as a

function of the energy for acoustic phonon scattering. There are four discontinuity points in the scattering rates to the heavy hole band and three to the light hole band. These are present due to the non-parabolicity model used in equations 3.9 and 3.10. As the process is considered elastic there is not a restriction on the energy for the process to happen.

Figure 3.8- Acoustic phonon scattering rates for holes residing in the heavy, light and split-off bands. This scattering process is intra-band (the initial and the final bands are the same).



Figures 3.9, 3.10 and 3.11 present the results of the scattering rates for the non-polar optical phonon process when the carrier is residing in the heavy-hole band, the light-hole band and the split-off band, respectively. In the non-polar optical phonon scattering the carrier may change its band and need to either emit or to absorb a phonon. Focusing on the curves where the carrier scatters from the heavy hole to either the heavy hole or the light hole band in figure 3.9, one can notice that for an emission process the carrier needs to have at least the phonon energy to be able to scatter as this energy will be transferred from the carrier to the phonon bath during the scattering process. When scattering to the split-off band, the minimum energy to an emission process to happen is even higher as not only the phonon energy is needed, but also the energy difference between the heavy hole band and the split-off band has to be overcome. The same behavior can be seen in figure 3.11 when the carriers initial band is the split-off band. It is also interesting to notice that the scattering rate of the emission process is close to ten times higher than for the absorption process.

Figure 3.9 - Non-polar optical phonon rates for holes scattering from the heavy hole band.

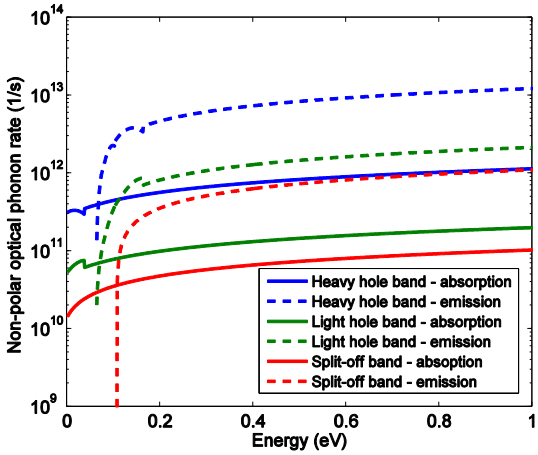


Figure 3.10 - Non-polar optical phonon rates for holes scattering from the light hole band.

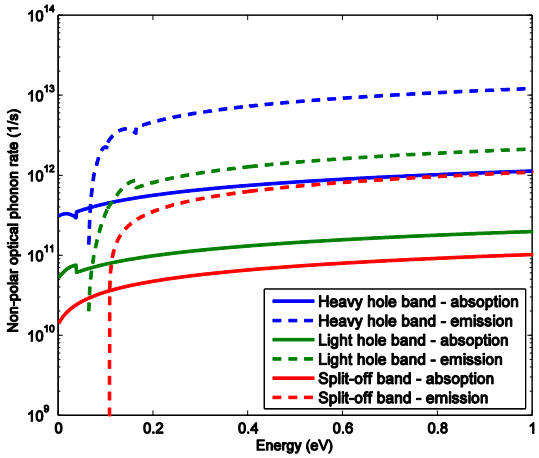
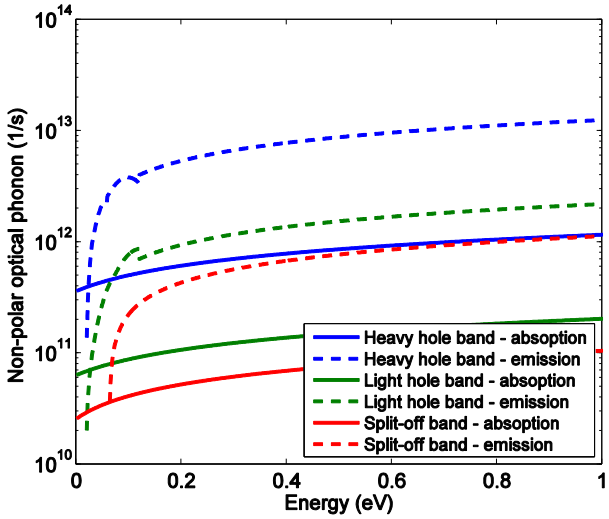


Figure 3.11 - Non-polar optical phonon rates for holes scattering from the split-off band.



3.1.3 Scattering Tables

As shown in the flowchart in figure 3.5, after each free-flight time of the carrier a scattering process will take place and the next free-flight time will be determined based on the scattering rates of the band in which the carrier is located. The equations 3.13 and 3.14 have a high computational cost to be solved as they include a numerical integral, thus becoming prohibitive to be solved for each choice of scattering mechanism. In order to avoid solving the same equations each time a scattering mechanism happens, a scattering table for each band is created where the information of the relative importance of the scattering rates from each process is included. The table is created for a wide range of energy so that no carrier reaches energy than what the table covers.

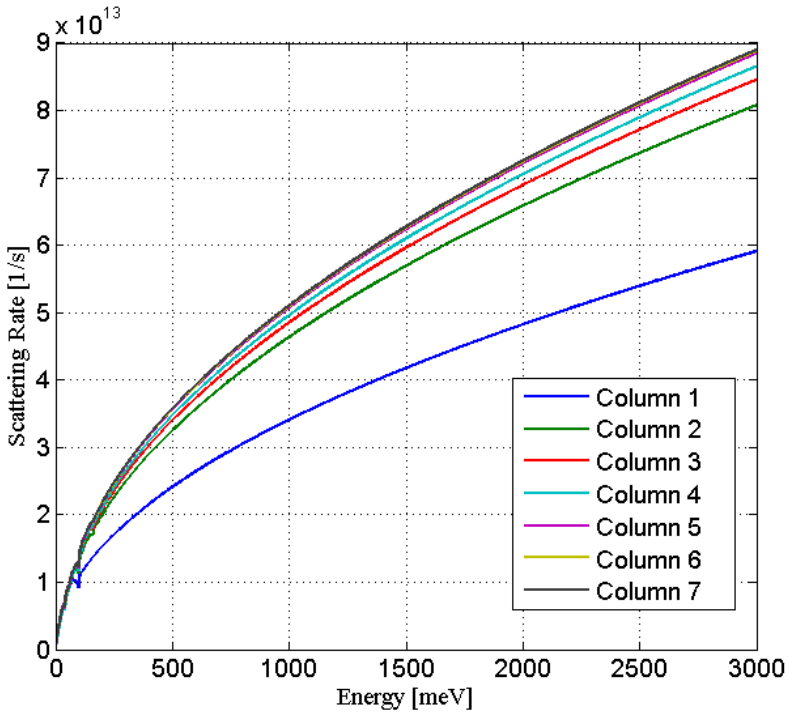
A different scattering table is created for each band. In each scattering table, the scattering regarding each process should be included for each energy value. In the first column the energy level is defined. The scattering processes are numerated from 1 to N, where N is the total number of processes included in the model for a given band. The second column receives the value of the scattering rate of the first scattering process, the third column receives the value of the sum scattering rates of the second process and the first one, the third column receives the value of the sum of the rates of the scattering processes from 1 to 3, and so on. The last column will contain the sum of the scattering rates for all processes for each given energy. Finally, the table is renormalized by its largest value, which is called Γ_0.

An example of a scattering table is shown in table 3.1 and in a graphical way in figure 3.12. Notice that for some energy values one process might be dominant while for others a different process dominates. As seen in the figures 3.8, 3.9, 3.10 and 3.11, the scattering rates between processes can reach several orders of magnitude for the energy range of interest. Including all possible scattering mechanisms, even though is not wrong, will add unnecessary computational cost to the simulation. The scattering mechanisms that have a relevant impact on the transport of carriers in a given energy range vary according to the material and the carrier type. For holes in silicon with energies lower than the bandgap, the acoustic phonon, the non-polar optical phonon and the Coulomb scattering must be considered, while the impact ionization effect should be considered for high electric fields (KUNIKIYO,1994). The omission of impact ionization process is justified with the fact that in the smallest devices the biases are on the order of 0.9 V, which means that the maximum carrier energy is 0.9 eV. This energy is not enough to generate electron-hole pairs due to impact ionization process in silicon. As the Coulomb scattering will be considered in real space, only acoustic and non-polar optical phonon scattering need to be considered in k-space.

Table 3.1- Example of a scattering table for Heavy-hole Band

Energy (position)	Column 2	Column3	Column4	Column5	Column6	Column7	Column8
...	...	...	...	...	...	...	...
400	2,159E+19	2,894E+19	3,022E+19	3,086E+19	3,161E+19	3,174E+19	3,181E+19
500	2,414E+19	3,250E+19	3,397E+19	3,470E+19	3,553E+19	3,568E+19	3,575E+19
600	2,644E+19	3,572E+19	3,734E+19	3,816E+19	3,906E+19	3,922E+19	3,930E+19
700	2,856E+19	3,866E+19	4,043E+19	4,133E+19	4,230E+19	4,247E+19	4,255E+19
800	3,054E+19	4,140E+19	4,330E+19	4,427E+19	4,530E+19	4,548E+19	4,557E+19
...	...	...	...	...	...	...	...

Figure 3.12 - Example of a scattering table for the Heavy-hole band for energies smaller than 3eV.



3.1.4 Free flight

Once the scattering tables are generated and the initial conditions of the carriers are defined, the simulation enters in the loop where the carrier transport is evaluated. The movement of the carriers is defined by two processes, the free-flight and the scattering. The duration of the free-flight is randomly defined by equation 3.21:

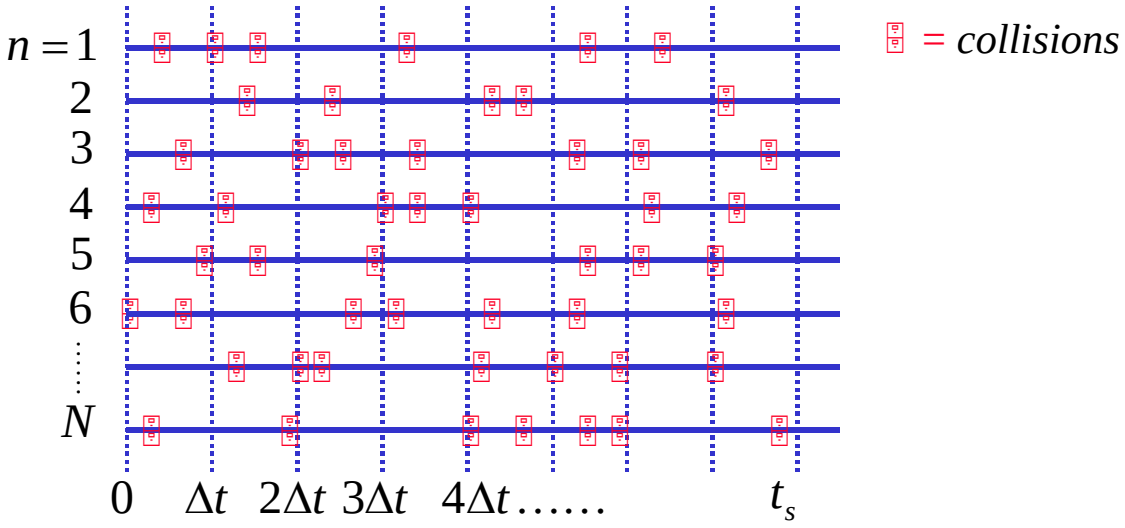
$$\tau = -\frac{1}{\Gamma_0}\ln(rand4) \ . \qquad (3.21)$$

As previously defined, the free-flight duration is the time between two scattering events. The scattering rates, as seen in the previous subsection, are highly dependent on the energy. It is impractical and has a high computational cost to generate a free-flight duration for each energy, so by randomly picking the free-flight duration based on Γ_0 instead of basing it on the actual scattering rates for that given energy the computational cost is reduced, but the rest of the code must be coherent with this factor. When using Γ_0, we are selecting a free-flight duration based on the worst case scenario, when the sum of the scattering rates is the largest for the energy range to which the simulation is limited. This will generate a smaller free-flight duration than it should be for every case but one. The self-scattering concept is created to fix this issue. The self-scattering has a probability to happened after each τ equal to the one minus the sum of the scattering rates of each process for that given energy normalized by Γ_0. When the self-scattering process is selected, nothing happens with the carrier and it simply resumes its free-flight.

During the free flight, the carrier is accelerated by the electric field. The change of carrier momentum in the electric field direction is proportional to the product of the force applied on the electron and the free-flight time. The new kinetic energy and the drift velocity are calculated using equation 3.11 using the momentum after the free-flight. As the equation 3.11 accounts for different behaviors according to the energy range it operates, there is a verification after each calculation to check if the range used was correct.

When dealing with several carriers, as in the Ensemble Monte Carlo case, to calculate temporal variation of the average drift velocity, average carrier energy, valley population, etc. an observation window is created and defines the amount of time the code tracks on carrier within one step. For a bulk simulation when the electrical field is constant this observation window can be arbitrary. In figure 3.13, the length of this window is defined as Δt. If the carrier free-flight time is larger than Δt, the carriers accelerate until the observation time is reached. The new momentum and energy are calculated at the end of each Δt. The remainder of the free-flight is recorded and used when the next observation window is executed. A schematic of a synchronous ensemble Monte Carlo scheme, that illustrates the free-flight duration times and the observation times, is shown in Figure 3.13.

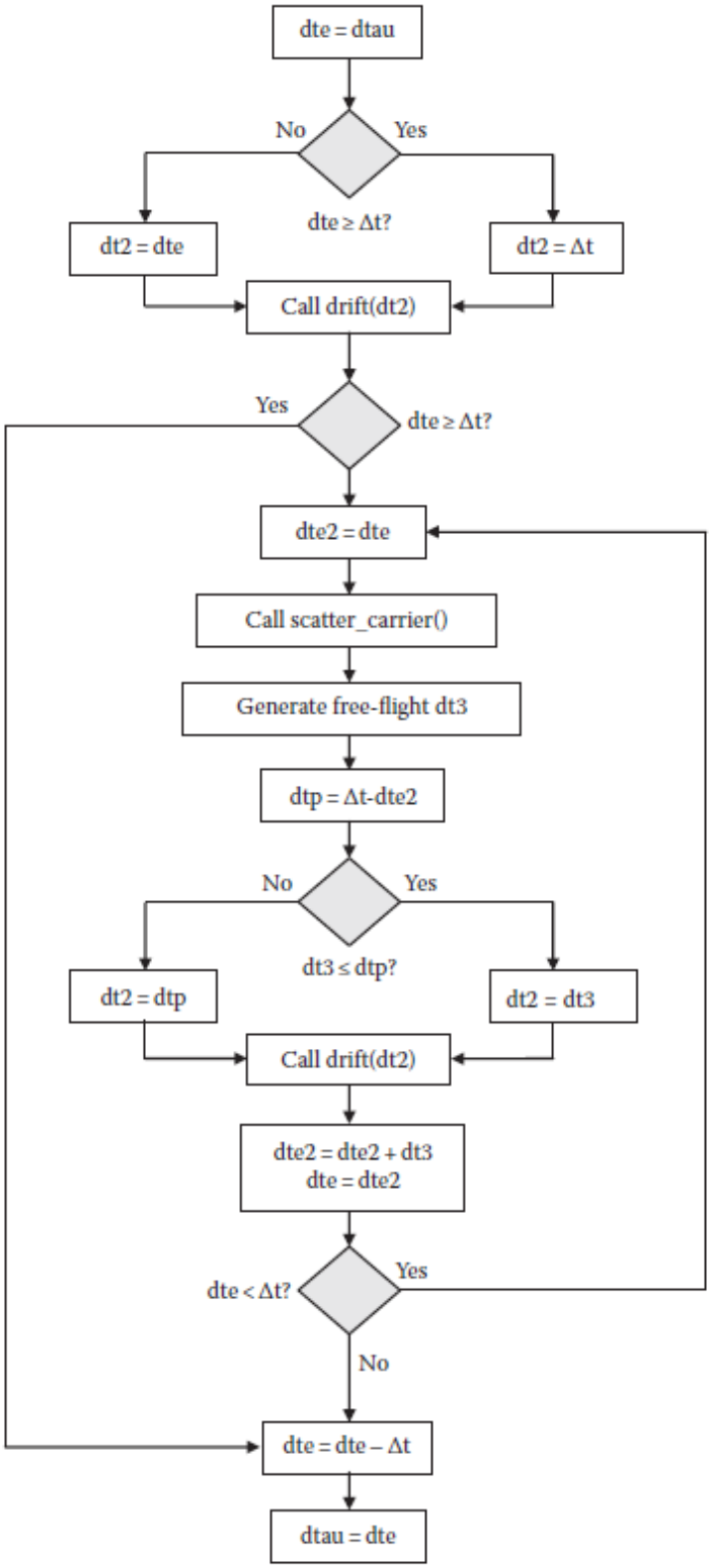
Figure 3.13 - Ensemble Monte Carlo simulation in which an observation time, Δt, is introduced over which the motion of particles is synchronized. The squares represent random scattering events.



Source: (VASILESKA, 2010)

A flowchart explaining each step of the way along the free-flight routine is presented in Figure 3.14. As in figure 3.13, Δt is the observation time and dte is the time until the next scattering event. The code begins checking if the time until the next scattering event is smaller than the observation time. If dte is larger than Δt, than no scattering will occur in that Δt and so the new momentum is calculated for the end of this Δt period. On the other hand, if dte is smaller than Δt, it means that a scatter event will occur. As the scattering rates are a function of the energy, the energy of the carrier should be calculated when the scattering happens, at the dte time. The scattering happens and a new free-flight time is assigned to that carrier. If the new scattering time is smaller than the time that is left in the observation period, a new scattering will happen and before that a new momentum will be calculated. The process is repeated until the observation time is reached.

Figure 3.14 - Free-flight-scatter flowchart.



Source: (VASILESKA, 2010)

3.1.5 Scattering Events

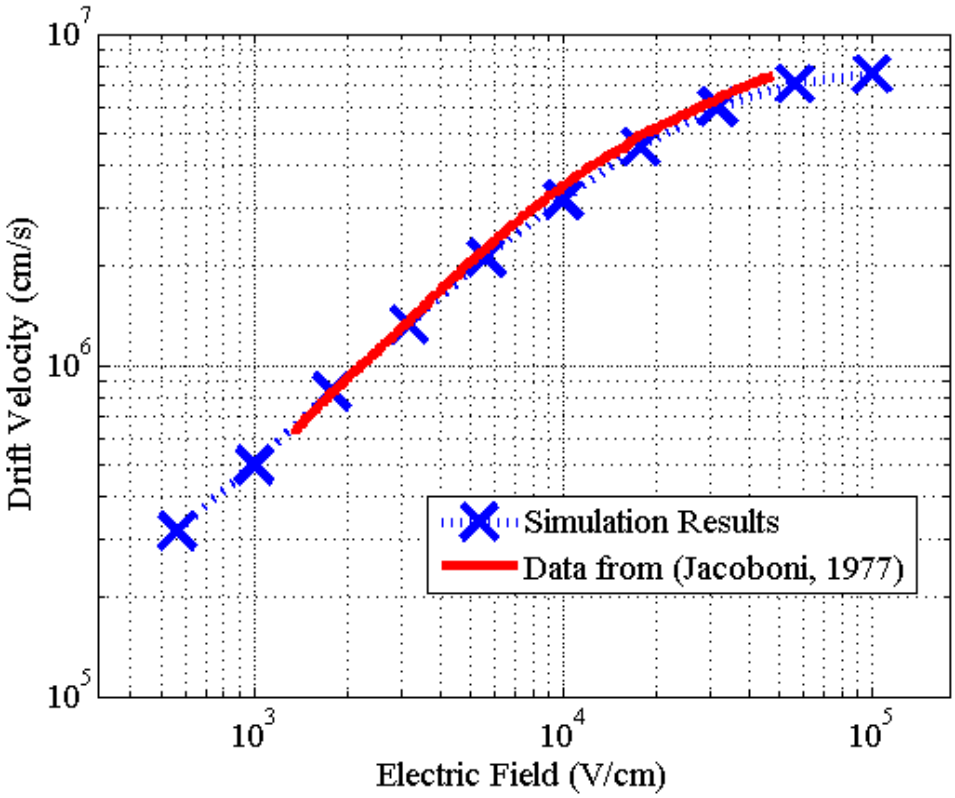
After a time τ calculated according to equation 3.21, a scattering event occurs. The scattering event will impact the carrier differently depending on the scattering process that happened, so first the scattering process is chosen randomly. When defining the carriers scattering process, a uniformly distributed random number between 0 and 1, *rand5*, is assigned to it. The rand5 number is compared to the scattering table for the carrier's given energy. If *rand5* is smaller than the value on column 2, than the scattering mechanism chosen is the mechanism 1, if it is larger the value *rand5* is tested against column 3, and so on. As the table was normalized by its maximum value, the highest value in the latest column will probably still be smaller than 1 and so there is a possibility that none of the real scattering mechanisms is selected. When no mechanism is selected, self-scattering event happens, which means the carrier will continue its free-flight for the next τ unaltered. In the code, the self-scattering is always checked first due to its highest probability of occurrence for low-energy electrons.

After the scattering process is selected, the energy and wave-vector are changed correspondingly. Due to separate absorption and emission phenomena for non-polar optical phonon scattering the energy of the carriers will change by the phonon energy. The acoustic phonon scattering mechanism is considered as elastic process and the energy remains the same. From the carrier energy, a new wave vector is chosen. Even though it is random (both non-polar optical phonon scattering and acoustic phonon scattering are isotropic), the new wave vector should respect the relation with the energy defined in equation 3.11 for the heavy and light hole bands. In order to assure this, the rejection technique presented by Nintunze in (1995) is used. In the case of scattering mechanisms to the split-off band, the scattering is considered isotropic as the band is considered spherical.

3.1.6 Bulk Simulation Results

The steady state and transient drift velocities of holes in silicon were calculated using the Monte Carlo model presented at 300K. 20000 carriers were considered in the simulation using the Ensemble Monte Carlo method. Figure 3.15 presents a comparison between the steady state drift velocities obtained by the Monte Carlo model compared to the experimental data from Jacoboni (1977). The model is in agreement with the data for the entire range of electric fields used.

Figure 3.15- Bulk Monte Carlo steady state drift velocity of holes compared with experimental data.



In Figure 3.16, it is shown the probability of a carrier to be located in each band. It is important to highlight that the occupancy probability changes as the electric field increases. As expected, for higher electric fields, when the carriers have higher energies, the importance of taking into account the split-off band is enhanced as the number of carriers in that band increases and, thus, it becomes relevant.

Figure 3.18 shows the average energy of the carriers in a transient simulation. As the electrical field applied increases the time it takes for the system to reach its steady state increases, however, eventually the system reaches the steady state and the energy absorbed by the carriers from the electrical field is lost in the scattering processes. Energy might be lost in the scattering processes either during emission process, if a phonon is emitted, or during the scattering, if the momentum of the carrier changes its direction in opposition to the direction of acceleration by the electric field.

Figure 3.17 shows the velocity overshoot effect through the time evolution of the drift velocity plotted for applied electric fields of 20, 50, 75 and 100kV/cm. The velocity overshoot happens when high electrical fields are applied and the drift velocity of the carriers becomes larger than the saturation velocity of the material. This effect impacts sub-100nm MOSFETs increasing the current that flows through the device.

Figure 3.16 - Valley population in steady state condition vs. Electric field.

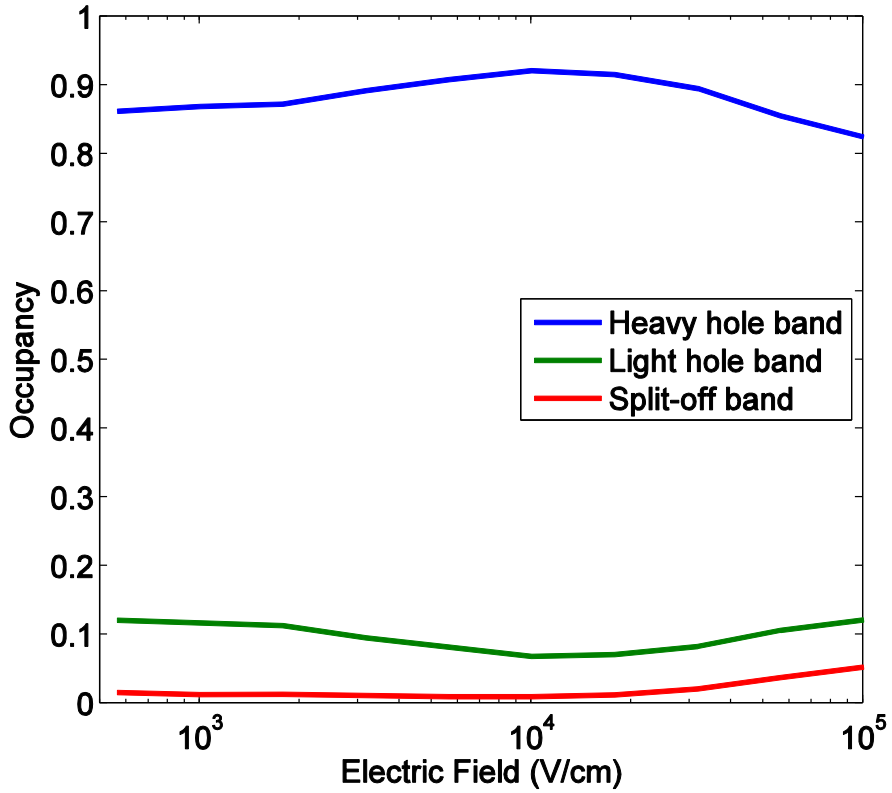


Figure 3.17 - Transient drift velocity of holes in silicon for applied fields of 15, 25, 50, 75 and 100kV/cm showing the velocity overshoot and the steady state.

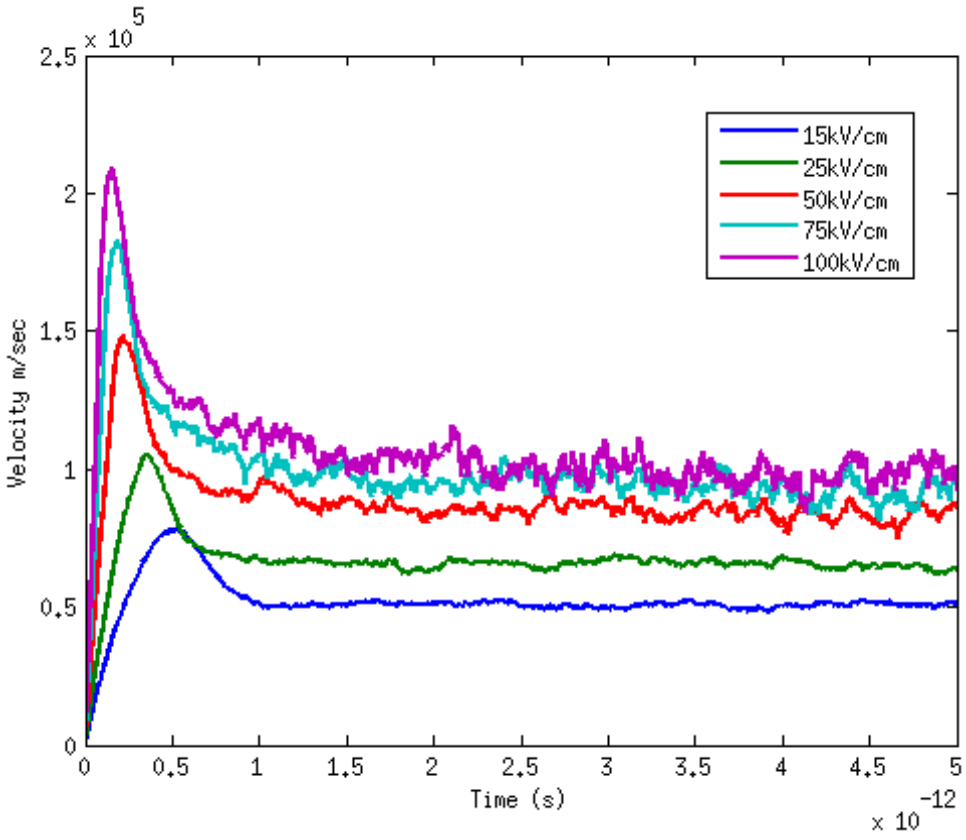
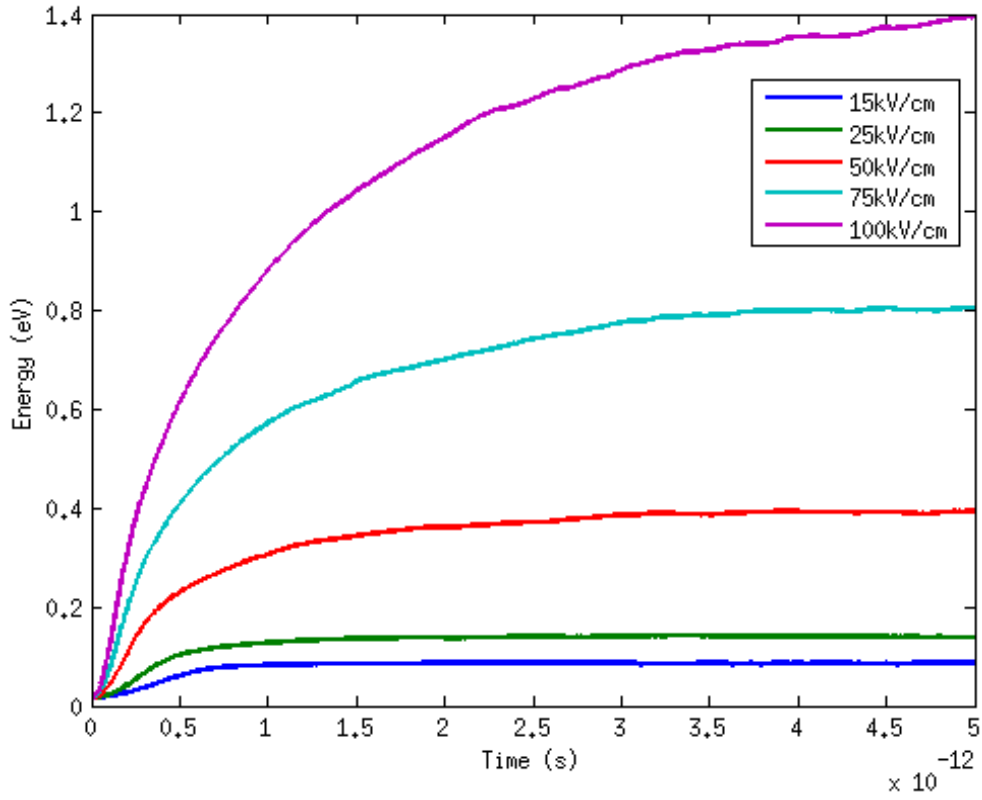


Figure 3.18 - Transient energy of hole in silicon for applied fields of 15, 25, 50, 75 and 100kV/cm.



3.2 Device Simulation

In the previous section a Monte Carlo simulator for the transport of holes in bulk silicon was presented. The simulation was based on a free-flight-scattering loop which is also the fundamental block in a device simulator. The device simulator couples the solution from the Boltzmann transport equation, here calculated using the Ensemble Monte Carlo method, and the Poisson equation. Around this basic setup there are several things that define the device.

The simulation framework of the device simulator consists of a loop where the free-flight occurs and the Poisson equation is solved for the new arrangement of the carriers in the device. The Poisson equation is then used to generate the electric potential profile in the device. The free-flight is once again evaluated with the new electrical field. On each step during the free-flight, the boundary conditions are checked to assure no carrier left the device unaccounted.

The simulation starts with the initialization of the material parameters. Here the code reads an input file where all the parameters of the material are defined. As presented in section 3.1, the scattering tables are generated based on the material parameters defined and using the same equations described in the previous section.

After the material initialization, the device structure is defined. First the 3D mesh is defined by the user. The mesh size is an important factor as it must be neither too large so that

the relevant information on the potential profile are lost, nor too small so that it causes trouble for the Poisson solver to converge. Important parameter that suggests the proper mesh size is the extrinsic Debye length which is inversely proportional to the square-root of the doping density.

In a 3D device simulator using the Ensemble Monte Carlo method, the computational cost of solving the Poisson equation is much larger than solving for the Monte Carlo loop. The mesh size is critical for the solution of the Poisson equation, so for a better computational efficiency, it is possible to generate a non-uniform mesh where the mesh cell size is smaller where a higher precision is needed, as in the channel, and larger where it is not needed, as in the body (substrate). In this code the mesh was defined to be uniform in the semiconductor, but smaller in the oxide region as one of the intentions of this code is to study the impact caused by charges in the dielectric. The mesh size can be changed by the user, so when simulating larger devices, the user might opt for a larger grid while for smaller devices the user needs to choose a smaller grid.

With the mesh in place, the next step is to define the boundary conditions. Here are defined the locations of the dialectic, the contacts and the regions where dopants will be added. It is in this step that the device structure is defined. Here the device can be defined as a Tri-gate, a simple planar MOSFET, etc. Also more design options can be included such as the addition of a high-k dielectric or a halo doping. In this thesis, all the simulations were run for a simple planar MOSFET.

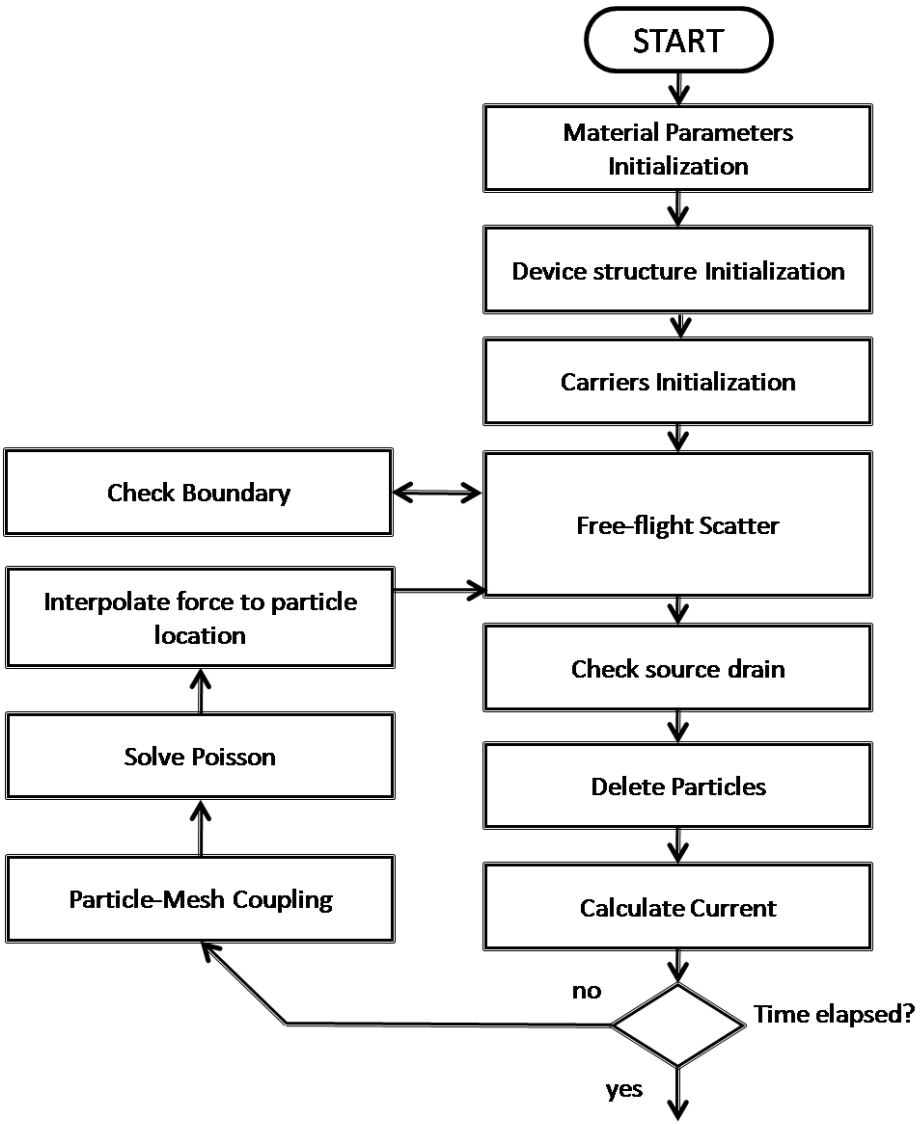
When defining the boundary conditions, the user should be aware on how these definitions might affect other regions in their vicinity. The contacts, for instance, affect the valence band of the semiconductor. In the gate, the band banding caused by the workfunction difference of the gate material and the semiconductor should also be considered.

After defining all boundaries and including their impact on the operation of the semiconductor device, the dopants are randomly placed in the regions where they belong to build the device; in this case a planar PMOSFET. The dopants might be distributed according to different probability distributions. Based on the doping density for a given region and the region’s volume, the expected number of dopant atoms is calculated. This value serves as the mean of a Poisson distribution and the actual number of atoms in that given region is randomly selected from that distribution. For the sake of simplicity, in the simulations in this thesis, the position of each atom randomly selected along the region based on a uniform distribution.

The carriers’ initialization occurs in a similar way as in the bulk Monte Carlo code. The energy of the carriers is also defined randomly based on a Boltzmann distribution and a random

k-vector is generated to each carrier matching its energy condition. The main difference is that now the number of carriers is not arbitrary anymore, nor is its location. The number of carriers is chosen so that the device is initially charge neutral, balancing the ions originated from the dopants. In order to define their position an equilibrium solution for the Poisson equations is performed and the carrier positions are defined accordingly.

Figure 3.19 - Simulation Framework of the Ensemble Monte Carlo Device simulator.



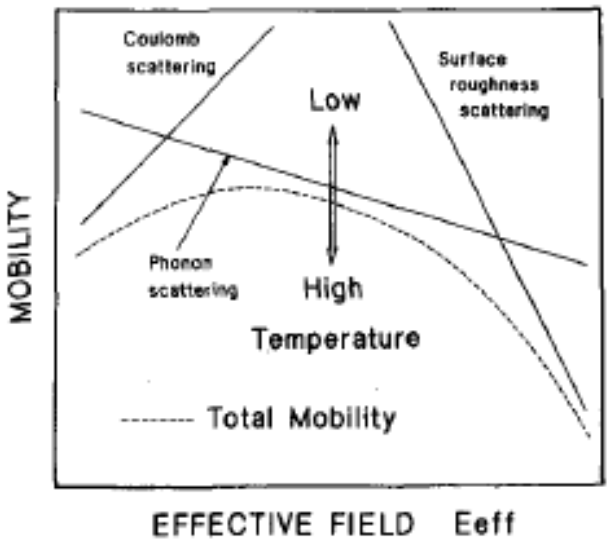
The Monte Carlo routine starts with the free-flight and scattering loop. This loop happens exactly in the same way as in the bulk code. Boundary checks are performed on each step to assure that the carriers don't leave the device through its borders or that it crosses the oxide. The device is limited by three kinds of boundaries, i) when a carrier crosses the border of the device in a contact, its exit is accounted and used to calculate the current (carriers that naturally came out), ii) when a carrier crosses a non-contact border, it is reflected into the device

once again, and iii) when the carrier crosses the oxide, it may be reflected specularly or scattered diffusively with a probability Ps_{urface} (TAKAGI, 1994).

The scattering caused by the short range interactions due to fluctuations of the silicon-oxide interface is called surface-roughness scattering. This is the scattering mechanism which generates the Ps_{urface} scattering probability. Traditionally the value attributed for the Ps_{urface} is 0.5 for electrons, however the author could not find any reference which includes the surface scattering probability of holes.

In a device large enough, where the mobility concept is still valid, different scattering mechanisms dominate for different conditions. For a small transversal electric field, the Coulomb scattering mechanism dominates. As the electric field increases then phonon scattering becomes the most dominant effect and for high transverse electric fields surface-roughness scattering dominates. This behavior is shown in figure 3.20. In order to obtain the value of $P_{Surface}$, a series of simulations with different values of $P_{Surface}$ were run for a large planar MOSFET with an applied transversal electrical field of 1MV/cm and a doping density of $6.6 \times 10^{17} cm^{-3}$. The results obtained for the mobility were compared to the experimental data from (Takagi, 1994). The best fit happened for $P_{Surface} = 0.6$.

Figure 3.20 - Mobility versus transversal electrical field in silicon.

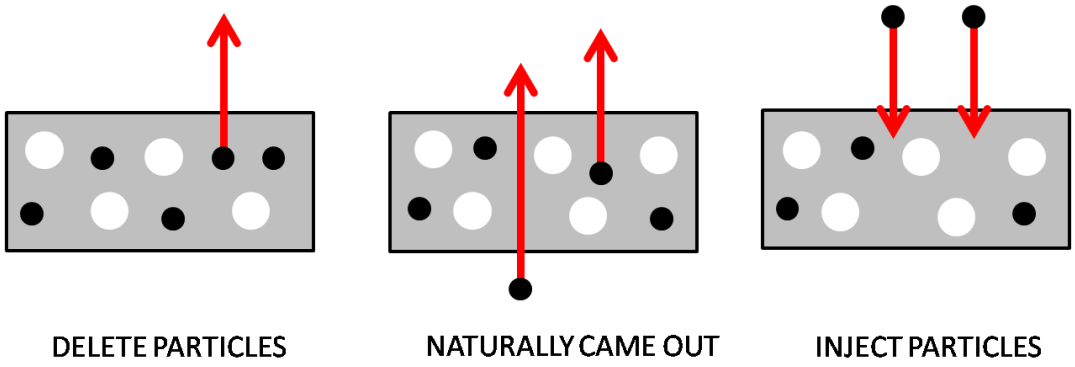


Source: (TAKAGI, 1994)

Another important step, which takes place after each Monte Carlo loop, is the balance of the source and drain contacts. In this step, the code checks if the charges are balanced in the mesh cells in the vicinity of the contacts. Three cases of activity may occur in the contacts, i) if the cell has an extra positive charge, a hole in the cell is deleted and is counted as if it left the

device (deleted carriers), ii) if a hole leaves the device through the contact during the free-flight stage, the carrier is also deleted and counted (carriers that naturally came out), iii) if the cell has an extra negative charge (a dopant), a carrier is added to the cell and is counted as if it entered the device (carriers injected). These cases are depicted in the figure 3.21 on the examples of different distribution of dopants and charge in one of the contacts (source or drain).

Figure 3.21 - Charge balance in the contacts.

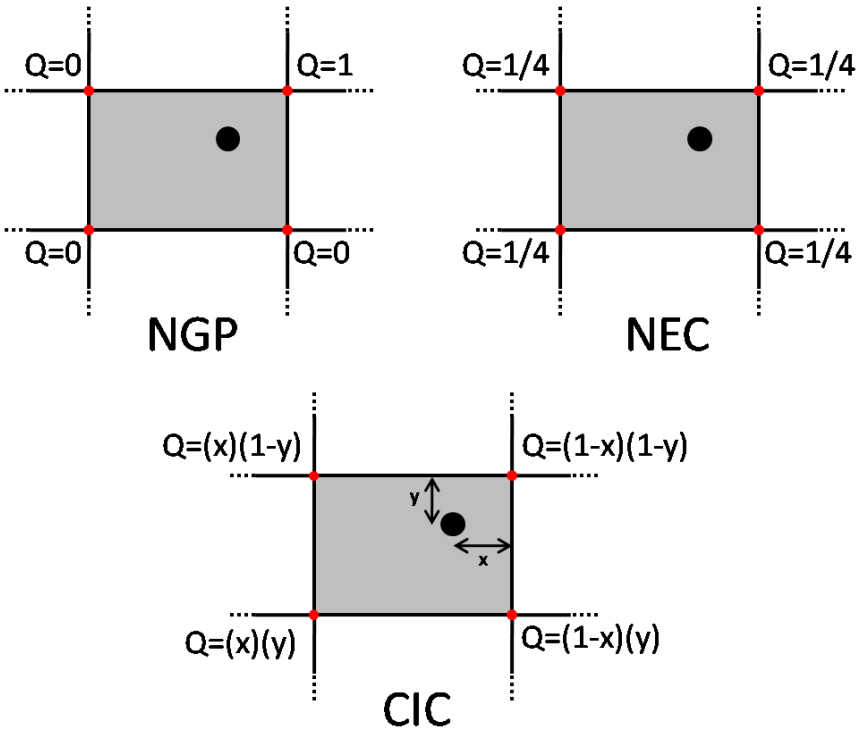


After each step a set of carriers are included and deleted. The stack of carriers is reorganized avoiding an overflow. The input and output flow of charge are used to calculate the current of the device as the derivative of the charge flow. The current leaving the device in one contact should be the same as the current entering the other contact due to charge conservation. The current in the bulk contacts is equal to zero as effects such as substrate leakage are not accounted for in this version of the code.

The Poisson equation is solved for the mesh cells and so it needs all the information regarding the charges of the device to be in grid, in a discrete space. The particles, due to Monte Carlo, have a real position and so for the Poisson solver all these charges must be redistributed to discrete points of space. The spread of the charges along the corners of the mesh cell is called particle-mesh coupling.

There are three methods for particle-mesh coupling. The NGP is the simplest and assumes that the charge of the particle is all in the nearest grid point. The NEC scheme spreads the charge equally among all corners and the CIC scheme relates the amount of charge that each corner is going to receive with its distance from the particle. The three methods are shown in figure 3.22, (LAUX, 1996). In this work the CIC method were used for the simulations, even though the code supports all three formats.

Figure 3.22 - Particle-mesh coupling methods.



With the charge densities along the device and the boundary conditions defined, the Poisson solver has all the information needed to calculate the potential along the device. The Strongly-Implicit Method (SIP) was used for the solution of the 3D Poisson equation (STONE, 1968).

3.2.1 Results

The code was implemented in Fortran and, just like for the bulk Monte Carlo, several tests were performed. The first test performed was aimed to test the contacts of the code. In order to do that, the charge conservation in the device was verified. The amount of carriers that passes through the contacts was measured along the simulation time. Assuming the current to be negligible in the gate and in the body contact once the simulation reaches steady state, the derivative of the cumulative charge in the source and drain contacts should have the same absolute value. This causes the curves in figure 3.23 to be parallel to each other once the steady state is reached. It is also interesting to look that while the steady state is not reached, there are more charges getting into the device than leaving it. This happens due to the formation of the channel.

The device used in the simulation presented in figures 3.23, 3.24 and 3.25 is a planar p-type MOSFET. The real channel length is equal to 60nm, the dielectric is entirely made of silicon oxide and is 1.2nm thick. The gate contact is made of poly-silicon while the source,

drain and bulk contacts are made of Aluminum. The source and drain regions depth are equal to 20nm. The doping density in the channel and in the substrate is $5\times10^{24}m^{-3}$, while in the source and drain regions the doping density is $1\times10^{25}m^{-3}$. No halo doping was included in the device.

Figure 3.23 - Simulation results for a 60nm PMOS presenting charge conservation in the device.

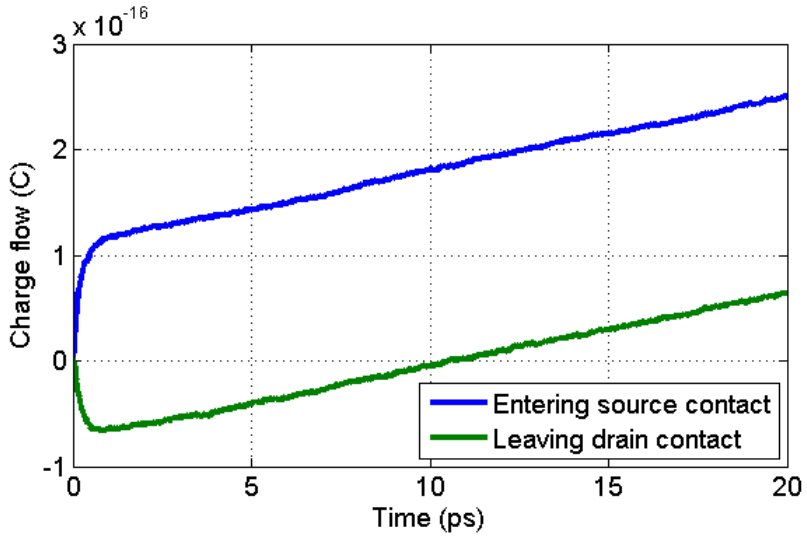
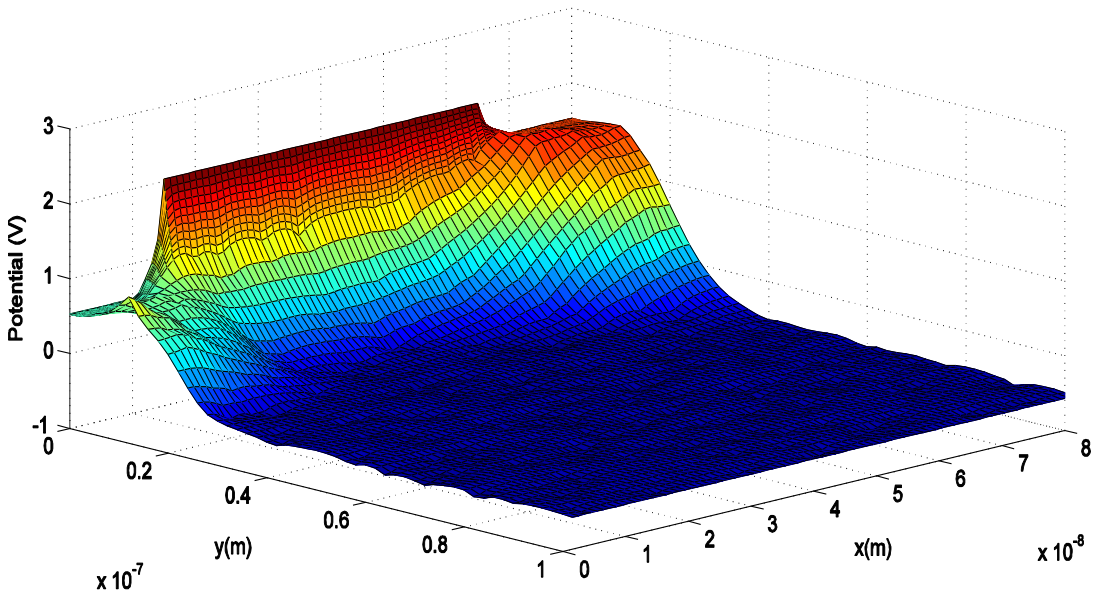


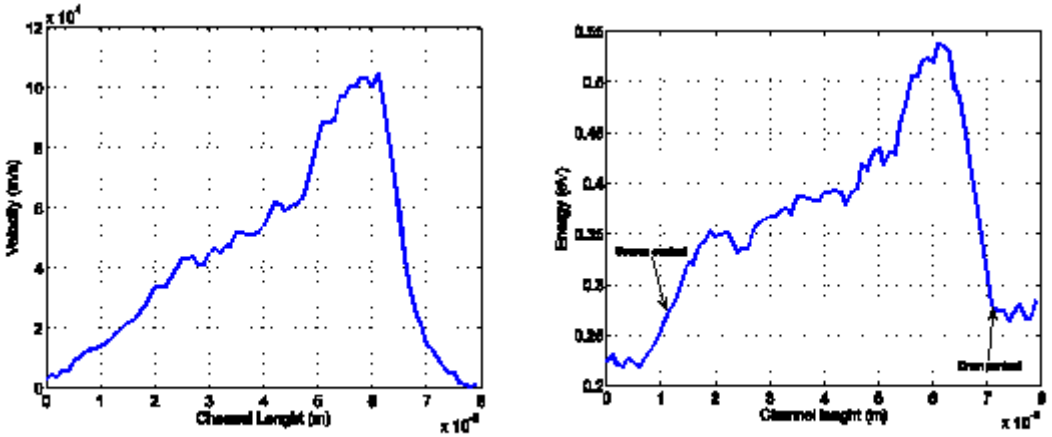
Figure 3.24 - Potential on a xy cross section of the device.



A 3D plot of the device allows us to check the boundary conditions and see the fluctuations caused by the random dopants. Figure 3.24 shows the potential in a xy cut-plane of the transistor, where the x is the axis in the channel direction and y goes from the oxide to the back contact. The band banding due to the gate is visible.

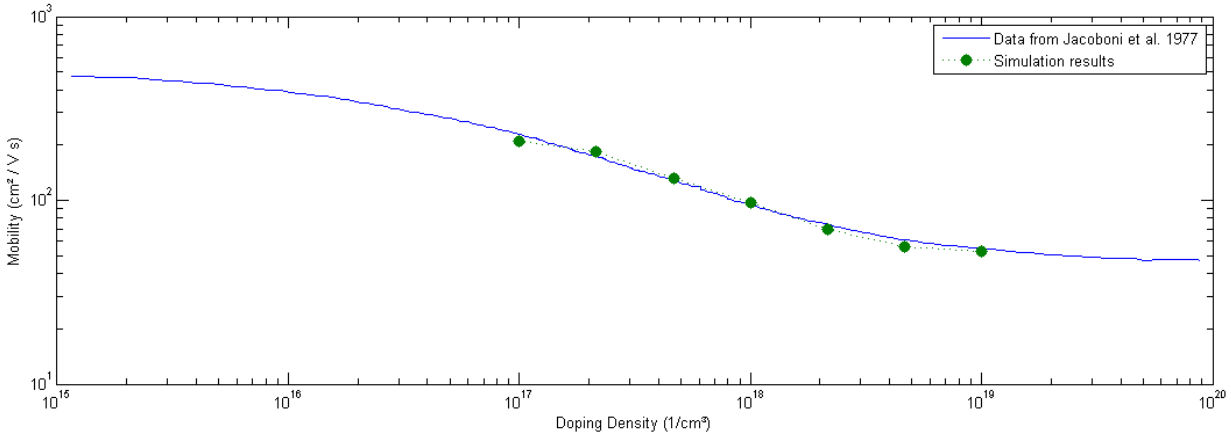
A typical figure of merit of device simulations is to plot the carriers' velocity and energy along the channel. It is important that the energy be sufficiently small so that semi classical approach remains valid and the effect of impact ionization can be neglected. One can also notice that the peak velocity of the carriers in some spots of the device is larger than the saturation velocity seen in the steady state study on bulk silicon (see figure 3.25). This observation illustrates that the velocity overshoot effect is happening in this device.

Figure 3.25- Behavior of the carrier's velocity (a) and energy (b) along the channel length.



After qualitatively checking the results produced by the code, some quantitative studies were performed to assure the correctness of the code. The main mechanisms responsible for the finite mobility of the material are the scattering mechanisms that are incorporated in the transport model. The phonon scattering mechanisms were already tested for the bulk properties, however, no tests were made for Coulomb scattering as this mechanism was only included in real space in the device simulations. Aiming to test the Coulomb scattering in the real space, we extracted the mobility of the silicon considering the real space position of the random dopants. The concept of mobility loses its significance for small transistors as the one which were used in the previous simulations because many carriers might not even scatter in such a small device. Hence, the boundary conditions were changed so that a resistor was built. A Monte Carlo simulation in the resistor was performed and the mobility was measured in the device. The resistor was sufficiently large (200nm × 5000nm^2) and the contacts were parallel and placed in the opposite sides so that the mobility could be properly measured. The mobility results were then compared to the ones obtained by Jacoboni (1977) and are shown in figure 3.26.

Figure 3.26 - Comparison of the mobility between a resistor simulated with the EMC device simulator and experimental data from Jacoboni (1977).



The next step was to compare the entire transistor simulation with the EMC device code developed. Comparisons with experimental data was not possible because to have a good comparison all the constructive information of the device should be available to be replicated in the code and this kind of information is usually restricted. The alternative was to compare the results of the EMC device code with the results generated by commercial TCAD tools. The TCAD tool chosen was the Atlas device simulator by Silvaco (SILVACO, 2010).

Figure 3.27 - Comparison of a PMOSFET transistor simulated with the EMC device simulator and the Drift-Diffusion tool from Silvaco.

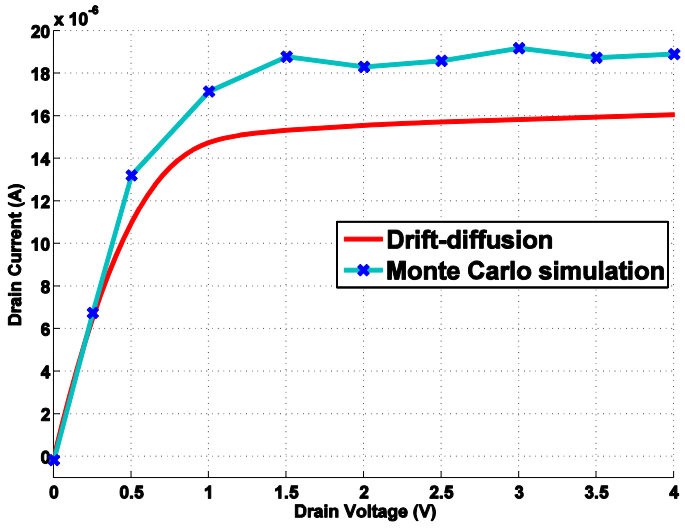


Figure 3.27 shows the I_d-V_d curve of a PMOSFET run with the EMC code and the Silvaco tool that utilizes the Drift-Diffusion method. The device simulated has the same characteristics as the one used in figures 3.23-3.25. In this figure the quantitative agreement is shown for low fields. For high drain voltages, where the carriers are subjected to higher electric

fields, the Monte Carlo simulation gives higher currents than the Drift-diffusion model. As shown previously on figure 3.27, in this device the velocity overshoot is present leading to a higher drift velocity which, in turn, leads to a higher current. This explains the higher current obtained with the Monte Carlo simulation.

3.4 Conclusions

In this chapter a device simulator was presented that is based on the Ensemble Monte Carlo method. The EMC method was detailed and a bulk simulator was first presented that reproduced steady-state velocity-field characteristics in agreement with available experimental data. For the first time it was implemented a bulk simulator for holes in silicon where both the non-parabolicity and the warping of the heavy and light hole bands were considered. The device simulator framework was described next based on the previous bulk Monte Carlo model. The device simulator for a PMOS considering the dopants in real position allows the effect of random dopants to be considered accurately and provide better results for high doping densities. Each relevant part of the code was discussed extensively. Results were also shown confirming that the code generates reliable results.

4 CONCLUSIONS

In this work two simulation tools were developed and extensively used for the purpose of achieving better understanding of the impact of traps in devices and circuits. Using the developed circuit simulation tool, case studies were performed bringing insights on the behavior of ring oscillators due to the trap activity. It was also shown that SSTA tools can well predict circuit behavior if properly used to predict BTI degradation. A device simulation tool based on the EMC method was built and tested. TCAD tools of the kind of the one developed in this thesis are not available commercially (there is no record of previous existence of a device simulator based on the EMC method for p-channel devices that simultaneously takes into account warping and non-parabolicity of the bands and the atomistic nature of the impurity atoms).

The device simulator presented in this thesis is capable of considering impact of discrete dopants and charges both in the interface and in the dielectric. The lack of studies on p-type transistors in the literature, combined with the capability of the simulation tool, allows a wide range of studies as future works to be developed by myself and other collaborators that have access to the simulation tool. The studies include the analysis of the impact of charged traps in the IV characteristics of the PMOS. Further development of the code to consider self-heating effects is already being implemented. Future works also include the characterization of the trap characteristics for a specific technology, aiming its use as input for the circuit simulator developed and presented in chapter 2.

REFERENCES

AADITHYA, K. V. et al., "Accurate Prediction of Random Telegraph Noise Effects in SRAMs and DRAMs," **Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on**, vol.32, no.1, p.73-86, Jan. 2013.

AGOSTINELLI, M et al. "Erratic fluctuations of sram cache vmin at the 90nm process technology node," In: Electron Devices Meeting, 2005**. IEDM Technical Digest. IEEE International**. [S.I]: IEEE, 2005. p.655,658.

AMOROSO, S.M. et al., "Impact of Cell Shape on Random Telegraph Noise in Decananometer Flash Memories," **Electron Devices, IEEE Transactions on**, vol.59, no.10, p.2774-2779, Oct. 2012

AMOROSO, S.M.; GERRER, L.; ASENOV, A. "3D TCAD statistical analysis of transient charging in BTI degradation of nanoscale MOSFETs," In: **International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)**. [S.I]: IEEE,2013. p. 5-8

AMOROSO, S.M. et al. "Inverse Scaling Trends for Charge-Trapping-Induced Degradation of FinFETs Performance," **Electron Devices, IEEE Transactions on**, vol.61, no.12, p.4014-4018, Dec. 2014

ASENOV, A. et al. "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," **Electron Devices, IEEE Transactions on**, vol.50, no.9, p. 1837- 1852, Sept. 2003

ASHRAF, Nabil et al. Accurate model for the threshold voltage fluctuation estimation in 45-nm channel length MOSFET devices in the presence of random traps and random dopants. **IEEE Electron Device Letters**, v. 32, n. 8, p. 1044-1046, 2011.

BINA, M. et al. "Modeling of hot carrier degradation using a spherical harmonics expansion of the bipolar Boltzmann transport equation," In: **Electron Devices Meeting (IEDM).** San Francisco: IEEE, 2012**.** p. 30.5.1-30.5.4.

BUKHORI, M.F.; ROY, S.; ASENOV, A., "Simulation of Statistical Aspects of Charge Trapping and Related Degradation in Bulk MOSFETs in the Presence of Random Discrete Dopants," **Electron Devices, IEEE Transactions on**, vol.57, no.4, p.795-803, April 2010

CATHIGNOL, A. et al. "Quantitative Evaluation of Statistical Variability Sources in a 45-nm Technological Node LP N-MOSFET," **Electron Device Letters, IEEE**, vol.29, no.6, p.609-611, June 2008

CHALKIADAKI, M. A. et al. “Why- and how- to integrate Verilog-A compact models in SPICE simulators.” **International Journal of Circuit Theory and Applications**, v. 41, n. 11, p. 1203-1211, 2013.

DEWEY, J., OSMAN, M. A. Monte Carlo study of hole transport in silicon. **Journal of applied physics**, v. 74, n. 5, p. 3219-3223, 1993.

FRANCO, J. et al. "Impact of Individual Charged Gate-Oxide Defects on the Entire ID–VG Characteristic of Nanoscaled FETs," **Electron Device Letters, IEEE**, vol.33, no.6, p.779,781, June 2012

GAGLIANI, G.; REGGIANI, L. “Nonparabolicity and intrinsic carrier concentration in Si and Ge.” **Il NuovoCimento B**, v. 30, n. 2, p. 207-216, 1975.

GERRER, L. et al. “Modelling RTN and BTI in nanoscale MOSFETs from device to circuit: A review,” **Microelectronics Reliability**, Volume 54, Issue 4, April 2014

GRASSER, T. et al. "The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability," In: **Reliability Physics Symposium (IRPS)**. Anaheim: IEEE, 2010, p. 16-25.

GRASSER, T. et al. "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction–Diffusion to Switching Oxide Traps," **Electron Devices, IEEE Transactions on**, vol.58, no.11, p.3652-3666, Nov. 2011

GRASSER, T.; "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," **Microelectronics Reliability**, Volume 52, Issue 1, p. 39-70 January 2012.

GRINSTEAD, C. M., SNEL, J. L., “Central Limit Theorem” in **Introduction to Probability**, Second Revised edition. C. M. Grinstead, J. L. Snel, United States: American Mathematical Society, 1997, p. 325

HORSTMANN, J.T.; HILLERINGMANN, U.; GOSER, K.; "Correlation Analysis of the Statistical Electrical Parameter Fluctuations in 50 nm MOS-Transistors," In: Solid-State Device Research Conference, 1998. **Proceeding of the 28th European**. Bordeaux: [s.n.], 1998. p. 512-515.

Ioffe.ru. In: <http://www.ioffe.ru/SVA/NSM/Semicond/Si/bandstr.html>. 2016.

JACOBONI, C., CANALI et al. “A review of some charge transport properties of silicon.” **Solid-State Electronics**, v. 20, n. 2, p. 77-89, 1977.

JACOBONI, C., LUGLI, P. **The Monte Carlo Method for Semiconductor Device Simulation**, Wien: Springer, 2002. Computational Microelectronics.

KACZER, B. et al. "Origin of NBTI variability in deeply scaled pFETs," In: **Reliability Physics Symposium (IRPS)**. [S.I.]:IEEE,2010. p.26-32.

KACZER, B. et al. "Atomistic approach to variability of bias-temperature instability in circuit simulations," In: **Reliability Physics Symposium (IRPS)**. [S.I]: IEEE,2011. pp.XT.3.1-XT.3.5.

KEYES, R. W. “The effect of randomness in the distribution of impurity atoms on FET thresholds,” **Appl. Phys.**, vol. 8, no. 3, p. 251–259, Nov. 1975.

KOLHATKAR, J.S.; “**Steady-state and cyclo-stationary RTS noise in MOSFETS**.” PhD thesis, The University of Twente, The Netherlands, 2005.

KRISHNAN, S**.;** "**Band-structure and detailed quantum effects on hole transport in p -channel MOSFETs",** PhD Thesis, Arizona State University, Tempe, USA, 2005.

KUMAR, S.V.; KIM, C.H.; SAPATNEKAR, S.S., "NBTI-Aware Synthesis of Digital Circuits," In: **Design Automation Conference, 44th**. [S.I.]: ACM/IEEE, 2007. p.370-375.

KUNIKIYO, T. et al. A Monte Carlo simulation of anisotropic electron transport in silicon including full band structure and anisotropic impact-ionization model. **Journal of Applied Physics**, v. 75, n. 1, p. 297-312, 1994.

LAUX, S. E. "On particle-mesh coupling in Monte Carlo semiconductor device simulation." **IEEE transactions on computer-aided design of integrated circuits and systems**, v. 15, n. 10, p. 1266-1277, 1996.

LEYRIS, C. et al. "Random telegraph signal noise SPICE modeling for circuit simulators," In: **Solid State Device Research Conference,** 37th. [S.I.: s.n.], 2007. p.187-190.

LUNDSTROM, M.; REN, Z.; "Essential physics of carrier transport in nanoscale MOSFETs." **IEEE Transactions on Electron Devices**, v. 49, n. 1, p. 133-141, 2002.

MACHLUP, S., "Noise in semiconductors: Spectrum of a two-parameter random signal," **J. Appl. Phys.**, vol. 25, no. 3, p. 341–343, Mar. 1954.

MARKOV, S. et al., "Statistical Interactions of Multiple Oxide Traps Under BTI Stress of Nanoscale MOSFETs," in **Electron Device Letters, IEEE** , vol.34, no.5, p.686-688, May 2013

NINTUNZE, N.; OSMAN, M. A. Hole drift velocity in the warped band model of GaAs. **Semiconductor science and technology**, v. 10, n. 1, p. 11, 1995.

PARRAT, L. G., "**Probability and Experimental Errors on Science**." New York, NY, USA: John Wiley and Sons Inc. 1961.

RABAEY, J. M., CHANDRAKASAN, A., NIKOLIC, B.. "**Implementation Strategies For Digital ICs**" in Digital Integrated Circuits, Second edition. Prentice Hall, 2003.

RAGHURAJ H., "**Generalized Monte Carlo Tool for Investigating Low-Field and High Field Properties of Materials Using Non-parabolic Band Structure Model**", MS Thesis, Arizona State University, June 2011

RAUCH, S.E.; "Review and Reexamination of Reliability Effects Related to NBTI-Induced Statistical Variations," **Device and Materials Reliability, IEEE Transactions on** , vol.7, no.4, p.524-530, Dec. 2007

REISINGER, H. et al. "Understanding and modeling AC BTI," In: **Reliability Physics Symposium (IRPS)**. [S.I.]: IEEE, 2011. p.6A.1.1-6A.1.8.

RODOPOULOS, D. et al. "Time and workload dependent device variability in circuit simulations," In: **IEEE International Conference on IC Design & Technology (ICICDT)**. [S.I.]: IEEE, 2011. p.1-4.

SANGWOO, H.; JUHO, K., "NBTI-aware statistical timing analysis framework," In: **IEEE International SOC Conference (SOCC)**. [S.I.]:IEEE, 2010. p.158-163.

SCHRODER, D. and BABCOCK, J. F., "Negative bias temperature instability Road to cross in deep submicron silicon semiconductor manufacturing," **J. Appl. Phys**., vol. 94, p. 1, 2003.

SILICON Integration Initiative, Inc. "**Statistical Methods For Semiconductor Chip Design.**". Austin, Texas, USA: Silicon Integration Initiative, Inc, 2007. ISBN: 1-882750-19-5

SILICON Integration Initiative, Inc. "**Si2 Effective Current Source Model (ECSM) Statistical Extensions Specification.**". Austin, Texas, USA: Silicon Integration Initiative, Inc., 2008. ISBN: 1-882750-35-7

SILVACO International Device simulator Atlas.Atlas User's Manual, 2010.

STONE, H. L., “Iterative solution of implicit approximations of multidimensional partial differential equations.” **SIAM Journal on Numerical Analysis**, v. 5, n. 3, p. 530-558, 1968.

TAKAGI, Shin-ichi et al. On the universality of inversion layer mobility in Si MOSFET's: Part I-effects of substrate impurity concentration. **IEEE Transactions on Electron Devices**, v. 41, n. 12, p. 2357-2362, 1994.

TAKEUCHI, K. et al. "Direct observation of RTN-induced SRAM failure by accelerated testing and its application to product reliability assessment," In: **Symposium on VLSI Technology (VLSIT)**. [S.I. : s.n.], 2010. p.189,190.

TAKEUCHI, K.; NAGUMO, T.; HASE, T., "Comprehensive SRAM design methodology for RTN reliability," In: **Symposium on VLSI Technology (VLSIT)**. [S.I. : s.n.], 2011, p.130,131.

TANG, T. B.; MURRAY, A.F.; ROY, S.; “Methodology of Statistical RTS Noise Analysis With Charge-Carrier Trapping Models.” in **Circuits and Systems I: Regular Papers, IEEE Transactions on**, V. 57, Issue 5, p. 1062-1070, 2010.

TOLEDANO-LUQUE, M. et al. "Response of a single trap to AC negative Bias Temperature stress," In: **Reliability Physics Symposium (IRPS)**. [S.I.]: IEEE, 2011. p.4A.2.1-4A.2.8.

TOLEDANO-LUQUE, M. et al. Temperature and voltage dependences of the capture and emission times of individual traps in high-k dielectrics, **Microelectronic Engineering**, Volume 88, Issue 7, Pages 1243-1246, July 2011-b.

TOLEDANO-LUQUE, M. et al. "Correlation of single trapping and detrapping effects in drain and gate currents of nanoscaled nFETs and pFETs," In: **Reliability Physics Symposium (IRPS)**. [S.I.]: IEEE,2012. p.XT.5.1-XT.5.6.

VAIDYANATHAN, B.; OATES, A.S., “Technology Scaling Effect on the Relative Impact of NBTI and Process Variation on the Reliability of Digital Circuits,” **Device and Materials Reliability, IEEE Trans. on** , vol.12, no.2, p.428-436, June 2012.

VAN DER WEL, AP. “Modeling random telegraph noise under switched bias conditions using cyclostationary RTS noise”; **Electron Devices, IEEE Transactions on**; v. 50, n. 5, p. 1378-1384, 2003.

VASILESKA, D., GOODNICK S. M. and KLIMECK, G.: **Computational Electronics: From Semi-Classical to Quantum Transport Modeling**, CRC Press, June 2010

WANG, W., WEI, Z., YANG, S., CAO, Y., "An efficient method to identify critical gates under circuit aging," In: **International Conference on Computer-Aided Design, 2007**. [S.I.]: IEEE/ACM, 2007. p.735-740.

WANG, et al. “The Impact of NBTI Effect on Combinational Circuit: Modeling, Simulation, and Analysis” **IEEE Transactions on Very Large Scale Integration (VLSI) Systems**, 2010

WANG, X. et al. "Statistical distribution of RTS amplitudes in 20nm SOI FinFETs," In: **Silicon Nanoelectronics Workshop (SNW)**. [S.I.]: IEEE, 2012. p.1-2.

WIRTH, G. et al. "Statistical model for MOSFET low-frequency noise under cyclo-stationary conditions," In: **IEEE International Electron Devices Meeting (IEDM), 2009**. [S.I.]: IEEE, 2009. p.1-4.

WIRTH, G. et al. “Compact modeling and simulation of Random Telegraph Noise under non-stationary conditions in the presence of random dopants” **Microelectronics and Reliability**, v. 52, p. 2955-2961, 2012.

XIAOJIANG, He, "**Two-Dimensional Monte Carlo Simulations of Ultra-Small MOSFETs**", MS Thesis, Arizona State University, May 2000

YAMAOKA, M. et al. "Evaluation methodology for random telegraph noise effects in SRAM arrays," In: **IEEE International Electron Devices Meeting (IEDM), 2011**. [S.I.]: IEEE, 2011. p.32.2.1-32.2.4.

Ye, Y.; Wang, C.C.; Cao, Y.; “Simulation of Random Telegraph Noise with 2-stage equivalent circuit.” In**: Proceedings of the IEEE/ACM International Conference on Computer Aided Design**. [S.I]: IEEE, 2010. p. 709-713.

ZHAO, W.; CAO, Y.; "New generation of Predictive Technology Model for sub-45nm early design exploration," **IEEE Transactions on Electron Devices**, vol. 53, no. 11, p. 2816-2823, November 2006.