

POLITECNICO DI MILANO  
SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE  
LAUREA MAGISTRALE IN INGEGNERIA MATEMATICA



# Development of a 3D-Solver Oriented to Numerical Simulations for Semiconductor Devices

Relatore: Prof. Riccardo SACCO  
Correlatore: Dott. Aurelio MAURI

Tesi di Laurea di:  
Andrea BORTOLOSSI  
Matr. n. 783023

Anno Accademico 2013–2014



# Contents

<b>Introduction</b>	<b>ix</b>
<b>1 Semiconductor model</b>	<b>1</b>
1.1 Basic Device Physics . . . . .	1
1.1.1 Intrinsic semiconductor . . . . .	1
1.1.2 Extrinsic semiconductor . . . . .	4
1.1.3 Densities at nonequilibrium condition . . . . .	6
1.1.4 Carrier transport in semiconductor . . . . .	7
1.2 Drift Diffusion Model for semiconductor . . . . .	9
1.2.1 Drift Diffusion formulation . . . . .	10
1.2.2 Generation and Recombination phenomenon . . . . .	11
1.2.3 Mobility models . . . . .	14
<b>2 Resolution of the system</b>	<b>19</b>
2.1 Geometry and boundary conditions . . . . .	19
2.2 Iteration algorithms . . . . .	22
2.2.1 Newton's method . . . . .	23
2.2.2 Fully coupled Newton's method . . . . .	24
2.2.3 Gummel map algorithm . . . . .	25
<b>3 Finite element discretization</b>	<b>29</b>
3.1 Non Linear Poisson Equation: weak form . . . . .	29
3.2 Continuity Equations: weak form . . . . .	31
3.3 Numerical approximation . . . . .	32
3.3.1 Geometrical discretization . . . . .	33
3.3.2 Linearized Non Linear Poisson equation . . . . .	34
3.3.3 Continuity equations . . . . .	37
<b>4 Simulation results</b>	<b>43</b>
4.1 Test cases . . . . .	43
4.1.1 p-n junction . . . . .	43

4.1.2	p-n junction in oxide . . . . .	50
4.1.3	MOSFET n-channel . . . . .	56
4.2	Calculation of the current at contacts . . . . .	64
4.2.1	Simulation results . . . . .	67
<b>Bibliografia</b>		<b>79</b>
<b>Ringraziamenti</b>		<b>83</b>

# List of Figures

1.1	Two typical examples of state density occupation ( $g(E)$ ) and probability distribution ( $f(E)$ ). . . . .	2
1.2	Construction of the band diagram. . . . .	4
1.3	Band diagrams of extrinsic silicon for (1.16) and (1.17). . . . .	6
2.1	(a) MOS device with net dopant concentration distributed according to a gaussian profile and $\Gamma_D$ colored in black. The oxide layer is colored in light blue. (b) Outline of the MOS device with $\Gamma_{int}$ in light gray. . . . .	20
2.2	Gummel map algorithm . . . . .	27
3.1	(a) Number of iteration against residual for different voltages in a diode test case. (b) Magnitude of the damping parameter $t_k$ . . . . .	37
3.2	Evaluation of the Zikatanov condition over a simple partition. Red elements doesn't satisfy condition (3.48) over four edges while blue elements fully accomplished the criterion. . . . .	42
4.1	p-n junction. . . . .	44
4.2	1D plots of the solutions and the quasi fermi potential levels along the line parallel to the Z-axis and placed at the center of the device. On the left is presented the test case at $V_A = 0.3[V]$ while on the right at $V_A = 1.0[V]$ . . . . .	45
4.3	p-n junction 0.3[V] - Electrostatic Potential. . . . .	46
4.4	p-n junction 0.3[V] - Electron density. . . . .	46
4.5	p-n junction 0.3[V] - Hole density. . . . .	46
4.6	p-n junction 1.0[V] - Electrostatic Potential. . . . .	47
4.7	p-n junction 1.0[V] - Electron density. . . . .	47
4.8	p-n junction 1.0[V] - Hole density. . . . .	47
4.9	Total time Gummel Map. . . . .	48
4.10	Time NLP and DD, iteration GM. . . . .	49

4.11	Initial step for different bias coupled with the solution. . . . .	49
4.12	Test case p-n junction in oxide. . . . .	50
4.13	1D plots of the solutions and the quasi fermi potential levels along the line parallel to the Z-axis and placed at the center of the device. On the left test case at $V_A = 0.3[V]$ is reported while on the right at $V_A = 1.0[V]$ . . . . .	51
4.14	p-n junction in oxide 0.3[V] - ElectrostaticPotential. . . . .	52
4.15	p-n junction in oxide 0.3[V] - Electron density. . . . .	52
4.16	p-n junction in oxide 0.3[V] - Hole density. . . . .	52
4.17	p-n junction in oxide 1.0[V] - Electrostatic Potential. . . . .	53
4.18	p-n junction in oxide 1.0[V] - Electron density. . . . .	53
4.19	p-n junction in oxide 1.0[V] - Hole density. . . . .	53
4.20	Test case dide p-n in oxide 0.3[V] - Electric field. . . . .	54
4.21	$E_y$ along a line parallel to Y-axis, $z = 0.22[\mu m]$ and $x = 0.1[\mu m]$ . . . . .	55
4.22	Geometry of MOS n-channel. . . . .	56
4.23	Energy band levels for nMOSFET along channel. . . . .	57
4.24	Channel of the nMOSFET. . . . .	58
4.25	Electric field density - $V_G = 2.0[V]$ . . . . .	58
4.26	Electrostatic potential - $V_G = 0.0[V]$ . . . . .	59
4.27	Electron density - $V_G = 0.0[V]$ . . . . .	59
4.28	Hole density - $V_G = 0.0[V]$ . . . . .	59
4.29	Electrostatic potential - $V_G = 2.0[V]$ . . . . .	60
4.30	Electron density - $V_G = 2.0[V]$ . . . . .	60
4.31	Hole density - $V_G = 2.0[V]$ . . . . .	60
4.32	Negative carriers spots for the electron density solution. . . . .	62
4.33	Electron density with finer mesh. . . . .	62
4.34	Zikatanov condition. . . . .	62
4.35	Electrostatic potential - Inverse polarization. . . . .	63
4.36	Hole density - Inverse polarization. . . . .	63
4.37	Diode characteristic. . . . .	67
4.38	SRH and Auger RG contribution. . . . .	68
4.39	p-n junction in oxide current at contact - Direct polarization. . . . .	69
4.40	pMOSFET. . . . .	70
4.41	$I_D - V_G$ nMOSFET characteristic - mobility models. . . . .	71
4.42	$I_S - V_G$ pMOSFET characteristic - mobility models. . . . .	72
4.43	$I_D - V_G$ nMOSFET - several drain voltages. . . . .	74
4.44	$I_S - V_G$ pMOSFET - several source voltages. . . . .	74
4.45	pMOSFET mesh. . . . .	76
4.46	Contribution of the impact ionization with the Van Over- straeten - de Man model inside the nMOSFET $V_D = 0.5[V]$ . . . . .	77

4.47 Contribution of the impact ionization with the Van Over- straeten - de Man model inside the pMOSFET $V_S = 0.5[V]$ . . .	77
4.48 Contribution of the impact ionization with the Van Over- straeten - de Man model inside the pMOSFET $V_S = 1.2[V]$ . . .	77
4.49 Inverse polarization of the nMOSFET. . . . .	78
4.50 Inverse polarization of the pMOSFET. . . . .	78





# List of Tables

1.1	List of parameters for the electron and hole Shockley-Read-Hall generation/recombination model. . . . .	13
1.2	List of parameters for the electron and hole in Auger generation/recombination model. . . . .	14
1.3	List of parameters of the electron and hole in van Overstraeten-de Man model . . . . .	15
1.4	List of parameters for the electron and hole mobility models including scattering from lattice thermal vibrations . . . . .	15
1.5	List of parameters in the electron and hole mobility models including scattering from ionized dopant impurities. . . . .	16
1.6	List of parameters for the electron and hole mobility models including scattering from velocity saturation. . . . .	17
4.1	p-n junction - list of settings, parameters and models. . . . .	44
4.2	p-n junction in oxide - list of settings, parameters and models. . . . .	50
4.3	nMOSFET - list of settings, parameters and models. . . . .	57
4.4	nMOSFET (inverse) - list of settings, parameters and models. . . . .	61
4.5	p-n junction (characteristic) - list of settings, parameters and models. . . . .	68
4.6	p-n junction in oxide - list of settings, parameters and models. . . . .	69
4.7	nMOSFET (low drain bias characteristic) - list of settings, parameters and models. . . . .	71
4.8	pMOSFET (low drain bias characteristic) - list of settings, parameters and models. . . . .	72
4.9	nMOSFET (different drain bias) - list of settings, parameters and models. . . . .	73
4.10	pMOSFET (different drain bias) - list of settings, parameters and models. . . . .	73
4.11	nMOSFET (off-state characteristic) - list of settings, parameters and models. . . . .	76

4.12 pMOSFET (off-state characteristic) - list of settings, parameters and models. . . . .	76
--	----

# Introduction

In 1947 John Bardeen, William Shockley and Walter Brattain (three scientists of Bell Telephone Labs) invented the bipolar transistor and since that crucial point there has been a growth of the semiconductor industry never known before, with serious impact on the way people work and live today.

Before reach the functionality and the miniaturization of modern devices, some fundamental steps has been made. In 1958 the first intagrated circuits (IC) was produced, followed by the introduction of the first MOSFET(1960) and CMOS(1963). Into these inventions the first micro-processor(1971) sank his roots and since that time until present, an ever-increasing progress has continued, according to the indication of *Moore's Law* (formulated by Gordon Moore in 1965).

These events led microelectronic industry at the doors of the VLSI era (Very-Large-Scale-Integration). Indeed in the last thirty years the benefits of miniaturization have been the key in the evolutionary progress leading to today's computers, wireless units, and comunication systems that offer superior performance, dramatically reduced cost per function, and much reduced physical size.

The large worldwide investment in VLSI technology constitutes a formidable driving force that guarantee the continued progress in IC integration density and speed, for as long as physical principles will allow.

From this point we want start and remark that the aim of numerical simulations is the comprehension of the physical phenomenon which lies behind the function of modern device.

Even if many commercial software are able to resolve different physic situations, they are often specialized on particular physic branch: obviously this strategic choice guarantees more efficiency but it implies a lost in generality. The consequence is that the work of device engineer becomes harder when he had to analyze electrical response coming from different phenomena.

Let us consider, as example, the functionality of a new device, which its electric behaviour is strong influeced by its mechanical response. Basically you are interested to the resolution of Maxwell's law (which is well

performed by SDEVICE simulator) and the Navier-Lamè equations (which is well performed by COMSOL simulator). Now the question is: how to put in communication the different outputs? Because it's not possible known precisely how the above programs resolves the equations, a relevant risk occurs when you decide to combine the solutions. In other words the development of an own code is at least desirable and possibly helpful: the main advantage is the control on simulation procedure and the possibility of fully customization, the major drawback is that this requires time and human resources, which in many cases are not available.

The FEMOS project (*Finite Element Method Oriented Solver*) tries to overcome the above limitations. FEMOS is designed for the treatment of chemical, mechanical, thermal and fluid phenomena. In this project we can't avoid the treatment of the classical semiconductor devices, this thesis found its origin in the development of this achievement.

In the first chapter we briefly recall the semiconductor material properties, physical behaviours and relations between the fundamental quantity (e.g. electrostatic potential, electric field, carrier densities and current densities). The classical Drift-Diffusion model has been detailed discussed; with the needed models for carrier mobility and generation/recombination phenomena.

The second chapter consists of two main sections. The first one presents the geometry framework and introduce some useful notations. The second one illustrates the algorithms used in order to treat the equations of the first chapter (decoupled Gummel map approach).

The third chapter the well-posedness analysis and the numerical approximation of the equations has been detailed discussed.

The fourth chapter contains numerical results of the simulations of typical test case (diode, MOS). The tests are compared with the results of a commercial code (SDEVICE). At the end the calculation of the current at contacts have been performed, extending the *residual method* [GS74] in the 3D case.

The fifth chapter **capitolo ancora in fase di costruzione.**

# Chapter 1

## Semiconductor model

In this chapter we present the basic physical properties of semiconductor material accordingly with the quantum mechanics theory [YT09] and the Drift-Diffusion model.

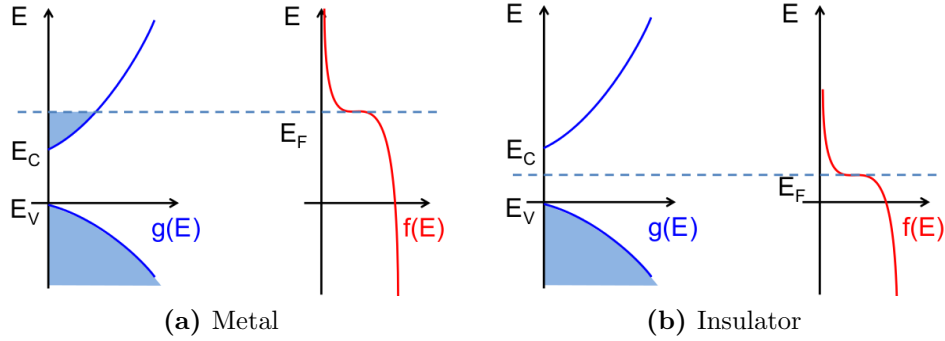
### 1.1 Basic Device Physics

As the most used material in the fabrication of VLSI devices is silicon, the description of physics is based on this material choice.

#### 1.1.1 Intrinsic semiconductor

In a silicon crystal each atom has four valence electrons to share with its four nearest neighboring atoms. The valence electrons are shared in a paired configuration called covalent bond. The description of electrons in a solid, made by the quantum mechanics, is that the allowed energy levels of electrons are grouped into bands separated by regions of not allowed energy: forbidden gaps. The highest energy band completely filled by electron at 0[K] is called the *valence band* ( $E_V$ ), the next energy band levels are called *conduction band* ( $E_C$ ).

Because in silicon the band gap is 1.11 [eV] [BGB00] also at room temperature a small fraction of the electrons are excited into the conduction band, leaving behind vacancies (called *holes*) in the valence band. In contrast, an insulator has a much larger forbidden gap making room-temperature conduction virtually impossible, while metals have partially filled conduction bands even at absolute zero temperature, this make them good conductors at any temperature.



**Figure 1.1:** Two typical examples of state density occupation ( $g(E)$ ) and probability distribution ( $f(E)$ ).

A suitable formulation of the electron concentration is given (for holes is symmetric) by the follow integral:

$$n = \int_{E_c}^{\infty} g(E)f(E) dE \quad (1.1)$$

With  $g(E)dE$  we indicate the number of electronic states per unit volume with an energy between  $E$  and  $E + dE$  in the conduction band and  $f(E)$  is the *Fermi-Dirac distribution function*, which gives the probability that an electronic state at energy  $E$  is occupied by an electron,

$$f_D(E) = \frac{1}{1 + \exp\left(\frac{E - E_f}{k_B T}\right)} \quad (1.2)$$

here  $k_B = 1.38 \times 10^{-23} [J/K]$  is Boltzmann's constant,  $T$  is the absolute temperature and  $E_f$  is the *Fermi level*. In general (1.1) is a Fermi integral of the order 1/2 and must be evaluated numerically.

**Definition 1.1.** The Fermi level ( $E_f$ ) is the energy at which the probability of occupation of an energy state by an electron is exactly one-half.

In most cases when the energy is at least several  $k_B T$  above or below the Fermi level (case of non degenerate semiconductor), equation (1.2) can be approximated by the Maxwell-Boltzmann statistics for classical particles, which reads as follows:

$$f_D(E) \simeq f_{MB}(E) = \begin{cases} \exp\left(-\frac{E - E_f}{KT}\right) & E \gg E_f \\ 1 - \exp\left(-\frac{E_f - E}{KT}\right) & E \ll E_f \end{cases} \quad (1.3)$$

Fermi level plays an essential role in characterizing the equilibrium state of a system, it is important to keep in mind the follow observation.

*Observation 1.1.* When two systems are in thermal equilibrium with no current flow between them, their Fermi levels must be equal: in other words for a continuous region (of metals or semiconductors in contact), the Fermi level at thermal equilibrium is flat (spatially constant throughout the region).

Considering (1.1) with (1.3) the analytical results of the integration is

$$n = N_c \exp\left(-\frac{E_c - E_f}{KT}\right) \quad (1.4)$$

$$p = N_v \exp\left(-\frac{E_f - E_v}{KT}\right) \quad (1.5)$$

where  $N_c$  and  $N_v$  are the *effective density of states*. In intrinsic semiconductor  $n = p$  and the *intrinsic Fermi level*  $E_i$  can be calculated using equations (1.4) and (1.5) as:

$$E_i = E_f = \frac{E_c + E_v}{2} - \frac{KT}{2} \ln\left(\frac{N_c}{N_v}\right) \quad (1.6)$$

By replacing (1.6) in (1.4) we have the expression of the intrinsic carrier concentration  $n_i = n = p$ :

$$n_i = \sqrt{N_c N_v} \exp\left(-\frac{E_g}{2KT}\right) \quad (1.7)$$

*Observation 1.2.* Since the thermal energy,  $k_B T$  is much smaller than the usual semiconductor bandgap  $E_g$ , the intrinsic Fermi level is very close to the midpoint between the conduction band and the valence band.

Equations (1.4) and (1.5) can be rewritten in terms of the intrinsic carrier density ( $n_i$ ) and energy ( $E_i$ ) :

$$n = n_i \exp\left(\frac{E_f - E_i}{KT}\right) \quad (1.8)$$

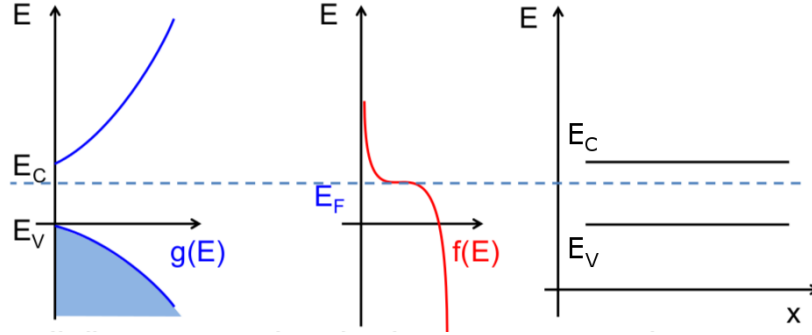
$$p = n_i \exp\left(\frac{E_i - E_f}{KT}\right) \quad (1.9)$$

Finally we remark a fundamental and useful relation holds at the thermal equilibrium

$$np = n_i^2 \quad (1.10)$$

this relation is usually note as *mass action law*.

The analysis of the work principles of devices can be effective done by the band diagram (Fig.1.2), which summarizes the informations presented above.



**Figure 1.2:** Construction of the band diagram.

### 1.1.2 Extrinsic semiconductor

At room temperature intrinsic semiconductor has an extremely low free-carrier concentration, therefore, its resistivity is very high. In order to make semiconductor a better conductor impurities atoms are added which introduce additional energy levels in the forbidden gap: these impurities are easily ionized adding either electrons to the conduction band or holes to the valence band. Here the electrical conductivity is dominated by the type and concentration of the impurity atoms.

In the case of silicon two are the types of impurities which are electrically active: those from column V such as arsenic or phosphorus, and those from column III such as boron or indium.

In most cases, the thermal energy at room temperature is sufficient to ionize the impurities and free the extra electron to the conduction band (column V) or accepting an electron from valence band (column III). Column V impurities are called *donors*; they become positively charged when ionized. Silicon material doped with column-V impurities or donors is called *n-type* silicon.

Column III impurities are called *acceptors*: they become negatively charged when ionized. Silicon material doped with column-III impurities or acceptors is called *p-type* silicon.

A p-type or an n-type is named as *extrinsic* silicon. In terms of the energy-band diagrams, donors add allowed electron states in the bandgap



close to the conduction-band edge, while acceptors add allowed states just above the valence-band edge.

The Fermi level in n-type silicon moves up towards the conduction band while in p-type silicon moves down towards the valence band. The exact position of the Fermi level depends on both the ionization energy and the concentration of dopants. For the sake of simplicity we consider that at room temperature all impurities are ionized ( $N_d = N_d^+$  and  $N_a = N_a^-$ ). For an n-type material with a donor impurity concentration,  $N_d$ , the charge neutrality condition requires that

$$n = N_d^+ + p \quad (1.11)$$

where  $N_d^+$  is the density of ionized donors. Similarly for a p-type material with acceptor impurity concentration  $N_a$  we have

$$p = N_a^- + n \quad (1.12)$$

Because the magnitude of impurities is between  $10^{16} \div 10^{20}[cm^{-3}]$ , and intrinsic carrier concentration in the order of  $10^{10}[cm^{-3}]$ , in typical case we can approximate the (1.10) as:

$$n \simeq N_d \quad p \simeq \frac{n_i^2}{N_d} \quad (1.13)$$

$$p \simeq N_a \quad n \simeq \frac{n_i^2}{N_a}$$

Replacing (1.13) in (1.4) and (1.5) in (1.11) and (1.12) and solving the algebraic equation we have

$$E_c - E_f = KT \ln \left( \frac{N_c}{N_d} \right) \quad (1.14)$$

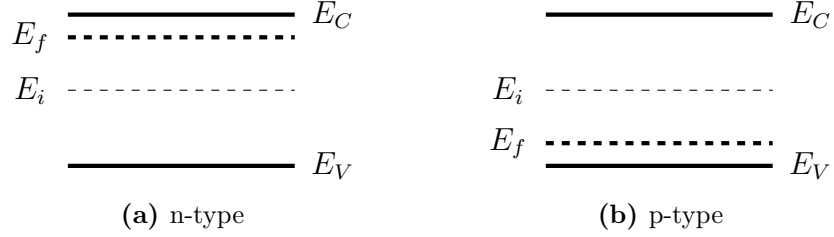
$$E_f - E_v = KT \ln \left( \frac{N_v}{N_a} \right) \quad (1.15)$$

Equation (1.14) and (1.15) can be written in a more useful form using (1.7) and (1.6) (for  $n_i$  and  $E_i$ ):

$$E_f - E_i = KT \ln \left( \frac{N_d}{n_i} \right) \quad (1.16)$$

$$E_i - E_f = KT \ln \left( \frac{N_a}{n_i} \right) \quad (1.17)$$

*Observation 1.3.* The distance between the Fermi level and the intrinsic Fermi level (near the midgap) is a logarithmic function of doping concentration.



**Figure 1.3:** Band diagrams of extrinsic silicon for (1.16) and (1.17).

### 1.1.3 Densities at nonequilibrium condition

In VLSI device a nonequilibrium condition is often possible where the densities of one or both types of carriers depart from their equilibrium values given by (1.8) and (1.9). In particular, the minority carrier concentration can be easily overwhelmed by injection from neighboring regions. Under these circumstances, while the electrons and holes are in local equilibrium with themselves, they are not in equilibrium with each other. In order to extend the relationship between Fermi level and densities discussed above, we can introduce different Fermi levels for electrons and holes. They are called *quasi Fermi levels* defined as follow

$$E_{fn} = E_i + k_B T \ln \left( \frac{n}{n_i} \right) \quad (1.18)$$

$$E_{fp} = E_i - k_B T \ln \left( \frac{p}{n_i} \right) \quad (1.19)$$

Considering the well known relation between electrostatic potential and energy (1.18) and (1.19) can be written as

$$n = n_i \exp \left( \frac{\varphi_i - \varphi_n}{k_B T / q} \right) \quad (1.20)$$

$$p = n_i \exp \left( \frac{\varphi_p - \varphi_i}{k_B T / q} \right) \quad (1.21)$$

where  $\varphi_n$  and  $\varphi_p$  are the quasi fermi electrostatic potential levels and  $\varphi_i$  is the midgap electrostatic potential level.

*Observation 1.4.* In non equilibrium condition quasi Fermi levels have the same physical interpretation in terms of the state occupancy as the Fermi level, therefore the electron (hole) density in the conduction band can be calculated using  $E_{fn}$  ( $E_{fp}$ ).

### 1.1.4 Carrier transport in semiconductor

Carrier transport or current flow in silicon is driven by two different mechanisms:

- the **drift** of carriers, which is caused by the presence of an electric field;
- the **diffusion** of carriers, which is caused by an electron or hole concentration gradient.

#### Drift current - Ohm's law

When an electric field is applied to a silicon device, the free carriers are accelerated and acquire a drift velocity superimposed upon their random thermal motion.

*Observation 1.5.* The drift velocity of holes ( $h$ ) is in the direction of the applied field, and the drift velocity of electrons ( $e$ ) is opposite to the field.

The velocity of the carriers does not increase indefinitely under field acceleration, since they are scattered frequently and lose their acquired momentum after each collision. During their motion throughout the lattice structure, carriers travel at an average speed defined by

$$\mathbf{v}_d^e = -\frac{q\mathbf{E}\tau_e}{m_e} \quad \mathbf{v}_d^h = +\frac{q\mathbf{E}\tau_h}{m_h} \quad (1.22)$$

where  $q = 1.602e^{-19}[C]$  is the elementary charge,  $\mathbf{E}$  the electric field,  $\tau_e$ ,  $\tau_h$  the average time between two consecutive scattering events and  $m_e$ ,  $m_h$  are the effective mass. The coefficient  $q\tau_e/m_e$  ( $q\tau_p/m_p$ ) characterizes how quickly a carrier can move through the lattice and it's well known as carrier mobility [ $m^2V^{-1}s^{-1}$ ]. In general, to include different scattering mechanism to the mobility the *Matthiessen's rule* is used:

$$\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_I} + \dots \quad (1.23)$$

where  $\mu_L$  and  $\mu_I$  correspond to the lattice and impurity scattering (for a more detailed description of mobility models see [YT09]).

Therefore the drift electron (hole) current density, reads as follows:

$$\mathbf{J}_n = -qn\mathbf{v}_d^n = qn\mu_n\mathbf{E} = \sigma_n\mathbf{E} \quad (1.24)$$

$$\mathbf{J}_p = +qp\mathbf{v}_d^p = qp\mu_p\mathbf{E} = \sigma_p\mathbf{E} \quad (1.25)$$

The scalar coefficient  $qn\mu_n$  ( $qp\mu_p$ ) is called electron (hole) conductivity  $\sigma_n$  ( $\sigma_p$ ).

Relations (1.24) and (1.25) expresses the well known *Ohm's law* stating that the current density is directly proportional to the applied electric field.

### Diffusion current - Fick's law

In semiconductor devices it's usual have different profiles of dopant in order to allow particular behaviors, this implies a not uniform concentration of carriers which they also diffuse as a result of the concentration gradient. This leads to an additional current contribution accordingly to the *Fick's law*:

$$\mathbf{J}_n = -D_n(-q\nabla n) \quad (1.26)$$

$$\mathbf{J}_p = -D_p(+q\nabla p) \quad (1.27)$$

The constants  $D_n$  and  $D_p$  are called electron and hole diffusion coefficients and have units of [ $cm^2s^{-1}$ ]. Drift and diffusion are closely associated with the random thermal motion of carriers and their collisions with the silicon lattice in thermal equilibrium. The *Einstein relation* (1.28) expresses the relation between diffusivity and mobility

$$D_\eta = \frac{k_B T}{q} \mu_\eta. \quad (1.28)$$

### Drift-Diffusion transport equations

By considering (1.24), (1.25), (1.26) and (1.27), the total electron and hole current densities become:

$$\mathbf{J}_n = qn\mu_n\mathbf{E} + qD_n\nabla n \quad (1.29)$$

$$\mathbf{J}_p = qp\mu_p\mathbf{E} - qD_p\nabla p \quad (1.30)$$

The total conduction current density is  $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$ .

Equations (1.29) and (1.30) are called constitutive laws and they can be rewritten in two other ways highlighting different physical explanations of the same phenomenon. These reinterpretations give also different start points for the discrete solver algorithm.

Considering that electric field is related to the scalar potential as:

$$\mathbf{E} = -\nabla\varphi \quad (1.31)$$

using (1.28) the current densities can be written as:

$$\mathbf{J}_n = -qn\mu_n \left( \nabla\varphi - \frac{k_B T}{qn} \nabla n \right) \quad (1.32)$$

$$\mathbf{J}_p = -qp\mu_p \left( \nabla\varphi + \frac{k_B T}{qp} \nabla p \right) \quad (1.33)$$

Considering equations (1.20) and (1.21) the above can be written as:

$$\mathbf{J}_n = -qn\mu_n \nabla\varphi_n \quad (1.34)$$

$$\mathbf{J}_p = -qp\mu_p \nabla\varphi_p \quad (1.35)$$

With these equations we underlying an important aspect which occur in semiconductor material:

*Observation 1.6.* The current density is proportional to the gradient of the quasi Fermi potential.

The third way to represent the current density is based on *Slotboom variables* which are particularly suited for the semiconductor equations.

$$u_n = n_i \exp \left( -\frac{\varphi_n}{V_{th}} \right) \quad (1.36)$$

$$u_p = n_i \exp \left( \frac{\varphi_p}{V_{th}} \right) \quad (1.37)$$

where  $V_{th} = k_B T/q$ . Using the above equations into (1.29) and (1.30) we obtain:

$$\mathbf{J}_n = qD_n \exp \left( \frac{\varphi}{V_{th}} \right) \nabla u_n \quad (1.38)$$

$$\mathbf{J}_p = -qD_p \exp \left( -\frac{\varphi}{V_{th}} \right) \nabla u_p \quad (1.39)$$

*Observation 1.7.* The drift-diffusion current density in a semiconductor, is a totally diffusive flux of a new kind of carrier with a proper diffusion coefficient.

## 1.2 Drift Diffusion Model for semiconductor

Simulations on integrated devices works on several different scale, the *Drift Diffusion model* (DD) is the most widely used mathematical tool for industrial simulation of semiconductor devices. In this section we'll show how is possible to deduce the DD model.

### 1.2.1 Drift Diffusion formulation

The system of Maxwell equations describes the propagation of electromagnetic signal in a medium:

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad (1.40)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (1.41)$$

$$\nabla \cdot \mathbf{D} = \rho \quad (1.42)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.43)$$

with the following set of constitutive laws that characterize the electromagnetic properties of the medium:

$$\begin{aligned} \mathbf{D} &= \epsilon \mathbf{E} \\ \mathbf{B} &= \mu_m \mathbf{H} \end{aligned} \quad (1.44)$$

where  $\epsilon$  is the material dielectric permittivity [ $Fcm^{-1}$ ] and  $\mu_m$  is the magnetic permeability [ $Hcm^{-1}$ ]. Since  $\nabla \cdot (\nabla \times \mathbf{A}) = 0$  for any vector  $\mathbf{A}$ , (1.43) is satisfied by introducing a vector potential  $\mathbf{A}$  such that  $\mathbf{B} = \nabla \times \mathbf{A}$ . We replace it in (1.41) obtains

$$\nabla \times \left( \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0. \quad (1.45)$$

From this we can state that exist a scalar potential  $\varphi$  such that

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \varphi \quad (1.46)$$

Applying the divergence operator and we obtain using (1.31), (1.44) and (1.42); (1.46) becomes

$$\rho + \frac{\partial \rho}{\partial t} = -\nabla \cdot (\epsilon \nabla \varphi) \quad (1.47)$$

We now assume that  $\frac{\partial \rho}{\partial t} = 0$  (quasi static approximation) and we have the *Poisson Equation*

$$\nabla \cdot (\epsilon \nabla \varphi) = \rho. \quad (1.48)$$

Applying the divergence operator on the equation (1.40) and we get the *Continuity Equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \quad (1.49)$$

To close the above system ((1.48) and (1.49) ), we need to specify the mathematical form of the electric charge density ( $\rho$ ) and the electric conduction current density ( $\mathbf{J}$ ). Considering (1.13)  $\rho$  can be expressed by (1.50)

$$\rho = \underbrace{q(p - n)}_{\rho_{free}} + \underbrace{q(N_D - N_A)}_{\rho_{fixed}} \quad (1.50)$$

- free charge ( $\rho_{free}$ ) (free electron and holes carriers),
- fixed charge ( $\rho_{fixed}$ ) (ionized dopant impurities).

Notice that we assume  $N_D$  and  $N_A$  time invariant ( $\partial N_D / \partial t = \partial N_A / \partial t = 0$ ).

Splitting the continuity equation for holes and electrons Drift Diffusion (DD) model formulation looks as follows:

$$\left\{ \begin{array}{l} \nabla \cdot (-\epsilon \nabla \varphi) = q(p - n + N_D^+ - N_A^-) \\ -q \frac{\partial n}{\partial t} + \nabla \cdot (-q \mu_n n \nabla \varphi + q D_n \nabla n) = qR \\ q \frac{\partial p}{\partial t} + \nabla \cdot (-q \mu_p p \nabla \varphi - q D_p \nabla p) = -qR \end{array} \right. \quad (1.51)$$

where  $R(\mathbf{x}, t)$  can be considered as the net rate of generation and recombination. The system is an incompletely parabolic initial value/boundary problem in three scalar unknown dependent variables  $\varphi(\mathbf{x}, t)$ ,  $n(\mathbf{x}, t)$  and  $p(\mathbf{x}, t)$ : the presence of the drift terms ( $n \nabla \varphi$  and  $p \nabla \varphi$ ) makes (1.51) a nonlinear coupled system of PDE's.

From Maxwell equations we are able to guarantee only that  $\mathbf{J}$  is a solenoidal field, we can't say nothing about the properties of  $\mathbf{J}_n$  and  $\mathbf{J}_p$ .

The stationary form can be easily derivate from (1.51) neglecting the temporal derivative.

### 1.2.2 Generation and Recombination phenomenon

The modelling of  $R(\mathbf{x}, t)$  is fundamental for device modeling due to its important role in determining the current-voltage characteristic.

It's important to keep in mind that electrons and holes are in continuous fluctuation due to their thermal energy, but the macroscopic result is that the net recombination rate at equilibrium is identically zero. Our interest is to analyze the deviations from this condition.

While generation events are usually due to thermal agitation or an external input source the recombination events happen in order to neutralize an excess of charge.

The phenomenological model for the net recombination rate  $R$  is often given by

$$R(n, p) = (pn - n_i^2)F(n, p) \quad (1.52)$$

where  $F$  is a function modelling for specific recombination/generation (R/G) event. In the following we present the classical theory about three kind of contributions.

### Shockley-Read-Hall recombination (SRH)

Electron and hole generation and recombination can take place directly between the valence band and the conduction band, or mediated via trap centers in the energy gap. Shockley-Read-Hall phenomena is a two-particle process which mathematically expresses the probability that:

- an electron in the conduction band neutralizes a hole at the valence band through the mediation of an unoccupied trapping level located in the energy gap ( $R_{SRH}$ ),
- an electron is emitted from the valence band to the conduction band, through the mediation of an unoccupied trapping level located in the energy gap ( $G_{SRH}$ ).

The function modeling  $F$  is

$$F_{SRH}(n, p) = \frac{1}{\tau_n \left( p + n_i \cosh \left( \frac{E_T}{k_B T} \right) \right) + \tau_p \left( n + n_i \cosh \left( \frac{E_T}{k_B T} \right) \right)} \quad (1.53)$$

where  $E_T$  is the energy level where traps live,  $\tau_n$  and  $\tau_p$  are called *carrier lifetimes* and are physically defined as the reciprocals of the capture rates. The typical order of magnitude lies in the range of  $10^{-3} \mu s \div 1 \mu s$  (see [VOT83] and [GH92]).



Parameter	Unit	Electrons	Holes
$\tau$	s	$1.0 \times 10^{-5}$	$3.0 \times 10^{-6}$
$E_T$	eV	0.0	0.0

**Table 1.1:** List of parameters for the electron and hole Shockley-Read-Hall generation/recombination model.

### Auger recombination (AU)

Auger R/G is a three-particle process and take place directly between the valence band and the conduction band. We distinguish four cases which depend to the kind of carriers involved in the phenomena:

$R_{AU}^{2n,1p}$  a high-energy electron in the conduction band moves to the valence band where it neutralizes a hole, transmitting the excess energy to another electron in the conduction band;

$G_{AU}^{2n,1p}$  an electron in the valence band moves to the conduction band by taking the energy from a high energy electron in the conduction band and leaves a hole in the valence band;

$R_{AU}^{2p,1n}$  an electron in the conduction band moves to the valence band where it neutralizes a hole, transmitting the excess energy to another hole in the valence band;

$G_{AU}^{2p,1n}$  an electron in the valence band moves to the conduction band by taking the energy from a high energy hole in the valence band and leaves a hole in the valence band.

The function modeling  $F$  becomes

$$F_{AU}(n, p) = C_n n + C_p p \quad (1.54)$$

where the quantities  $C_n$  and  $C_p$  are the so called Auger capture coefficients typically in the order of magnitude of  $10^{-25}[cm^6 s^{-1}]$  [LH80]. Note that Auger R/G is relevant only when both carrier densities attain high values.

### Impact ionization (II)

The impact ionization mechanism is a three-particle phenomena where carrier generation is triggered by the presence of high electric field: due to this field an electron could gains enough energy to excite an electron-hole pair out of a silicon lattice bond. Then the process can be repeated until an avalanche

Parameter	Unit	Magnitude
$C_n$	$cm^6 s^{-1}$	$2.9 \times 10^{-31}$
$C_p$	$cm^6 s^{-1}$	$1.028 \times 10^{-31}$

**Table 1.2:** List of parameters for the electron and hole in Auger generation/recombination model.

of generated carriers is produced within the region: this process can't be described with a relation like (1.52).

Among the different models for the impact ionization generation we chose the van Overstraeten - de Man model [vOdM70], based on the Chynoweth law [Chy58]:

$$G_{II}(n, p) = \alpha_n n |\mathbf{v}_n| + \alpha_p p |\mathbf{v}_p| \quad (1.55)$$

where:

$$\alpha(E_{ava}) = \gamma \exp\left(-\frac{\gamma b}{E_{ava}}\right) \quad (1.56)$$

$$\gamma = \frac{\tanh\left(\frac{\hbar\omega_{op}}{2k_B T_0}\right)}{\tanh\left(\frac{\hbar\omega_{op}}{2k_B T}\right)} \quad (1.57)$$

where  $\hbar\omega_{op}$  is the phonon energy and  $\gamma$  expresses the temperature dependence of the phonon gas against which carriers are accelerated,  $E_{ava}$  is the driving force that can be computed as

- the component of the electrostatic field in the direction of the current flows

$$E_{ava}^{n,p} = \frac{\mathbf{E} \cdot \mathbf{J}_{n,p}}{||\mathbf{J}_{n,p}||} \quad (1.58)$$

- the module of the quasi fermi gradient

$$E_{ava}^{n,p} = |\nabla \varphi_{n,p}| \quad (1.59)$$

### 1.2.3 Mobility models

In the following section we illustrate the most common phenomenological models to describe carrier mobilities. The main underlain physical phenomena for a mobility reduction are:

Parameter	Unit	Electrons	Holes	Valid range of electric field
$E_0$	$V\ cm^{-1}$	$4.0 \times 10^5$	$4.0 \times 10^5$	
$a_{high}$	1	$7.03 \times 10^5$	$6.71 \times 10^5$	$E_0$ to $6.0 \times 10^5$
$a_{low}$	1	$7.03 \times 10^5$	$1.582 \times 10^6$	$1.75 \times 10^5$ to $E_0$
$b_{high}$	1	$1.231 \times 10^6$	$1.693 \times 10^6$	$E_0$ to $6.0 \times 10^5$
$b_{low}$	1	$1.231 \times 10^6$	$2.036 \times 10^6$	$1.75 \times 10^5$ to $E_0$
$\hbar\omega_{op}$	eV	0.063	0.063	

**Table 1.3:** List of parameters of the electron and hole in van Overstraeten-de Man model

- interaction with the silicon atoms (due to thermal vibrations);
- interaction with ionized dopant impurities in the crystal.

### Scattering with lattice

Carrier mobility is a decreasing function of temperature, as we expect collisions to become more and more frequent as T gets higher (see [Lom88]). This can be represented as follows

$$\mu_\nu^L = \mu_\nu^0 \left( \frac{T}{T_0} \right)^{-\beta_\nu} \quad \nu = n, p \quad (1.60)$$

where  $\mu_\nu^0$  is the low-field mobility,  $\beta_\nu$  are positive numbers and  $T_0$  is a reference temperature  $T_0 = 300[K]$ .

Parameter	Unit	Electrons	Holes
$\mu^0$	$cm^2V^{-1}s^{-1}$	1417.0	470.5
$\beta$	1	2.5	2.2

**Table 1.4:** List of parameters for the electron and hole mobility models including scattering from lattice thermal vibrations

### Scattering from Ionized Impurities

Dopant ionized impurities represent local perturbations of the periodic silicon lattice, they strongly influence the carrier motion through electrostatic interaction, reducing the mobility. To take into account this Masetti has proposed the following model [MS83]

$$\mu = \mu_{min1} \exp\left(-\frac{P_c}{N_{tot}}\right) + \frac{\mu^L - \mu_{min2}}{1 + \left(\frac{N_{tot}}{C_r}\right)^\alpha} - \frac{\mu_1}{1 + \left(\frac{C_s}{N_{tot}}\right)^\beta} \quad (1.61)$$

where  $N_{tot} = N_D + N_A$ ,  $\mu_\nu^L$  is given by (1.60),  $\mu_{min1}$  and  $\mu_{min2}$  are the minimum value of  $\mu$ ;  $P_c$ ,  $C_r$  and  $C_s$  are reference doping values.

Parameter	Unit	Electrons	Holes
$\mu_{min1}$	$cm^2V^{-1}s^{-1}$	52.2	44.9
$\mu_{min2}$	$cm^2V^{-1}s^{-1}$	52.2	0
$\mu_1$	$cm^2V^{-1}s^{-1}$	43.4	29.0
$P_c$	$cm^{-3}$	0	$9.23 \times 10^{16}$
$C_r$	$cm^{-3}$	$9.68 \times 10^{16}$	$2.23 \times 10^{17}$
$C_s$	$cm^{-3}$	$3.43 \times 10^{20}$	$6.10 \times 10^{20}$
$\alpha$	1	0.680	0.719
$\beta$	1	2.0	2.0

**Table 1.5:** List of parameters in the electron and hole mobility models including scattering from ionized dopant impurities.

### Veclocity saturation at high electric field

Under the assumption of low electric field, mobilities are reasonably constant and the carrier drift velocity is proportional to the electric field. As the applied field strength increases, the above assumption predicts an unbounded carrier velocity as  $|\mathbf{E}| \rightarrow \infty$ . Although at high fields the average carrier energy increases and carriers lose their energy by optical-phonon emission nearly as fast as they gain it from the field.

This results in a decrease of the carrier velocity according to the following mathematical expression

$$\lim_{|\vec{E}| \rightarrow \infty} \mu|\mathbf{E}| = v_{sat} \quad (1.62)$$

A common adopted formula is the *Canali model* [Can75] with temperature dependent parameters

$$\mu = \frac{\mu_L}{\left[1 + \left(\frac{\mu_L |\mathbf{E}|}{v_{sat}}\right)^\beta\right]^{1/\beta}} \quad (1.63)$$

where  $\mu_L$  is (1.60) while  $v_{sat}$  and  $\beta$  are given by

$$v_{sat} = v_0 \exp\left(\frac{300}{T}\right)^{v_{exp}} \quad \beta = \beta_0 \left(\frac{T}{300}\right)^{\beta_{exp}}. \quad (1.64)$$

where  $v_0$  and  $\beta_{exp}$  are fitting parameters.

Parameter	Unit	Electrons	Holes
$v_0$	$cm\ s^{-1}$	$1.07 \times 10^7$	$8.37 \times 10^6$
$v_{exp}$	1	0.87	0.52
$\beta_0$	1	1.109	1.213
$\beta_{exp}$	1	0.66	0.17

**Table 1.6:** List of parameters for the electron and hole mobility models including scattering from velocity saturation.



# Chapter 2

## Resolution of the system

In this chapter we introduce geometry and boundary conditions for the stationary form of system (1.51) and we discuss the iteration algorithms used to decouple the same system.

### 2.1 Geometry and boundary conditions

In order to close the *Poisson equation* and the *Drift Diffusion equation* for electrons and holes of the stationary form of problem (1.51), suitable boundary conditions must be considered.

Let us consider the device domain as the union of two open disjoint subsets,  $\Omega_{Si}$  (doped silicon part), and  $\Omega_{ox}$  (oxide part), such that their intersection  $\partial\Omega_{Si} \cap \partial\Omega_{ox} = \Gamma_{int}$  is the interface. The oxide region  $\Omega_{ox}$  is assumed to be a perfect insulator so that:

$$\begin{aligned} n &= p = 0 \\ \mathbf{J}_n &= \mathbf{J}_p = \mathbf{0} \end{aligned} \tag{2.1}$$

The device boundary  $\partial\Omega$  is divided into two disjoint subsets:  $\partial\Omega_c$  and  $\partial\Omega_a$ . The subset  $\partial\Omega_c$  includes the so called *ohmic contacts* (with ohmic contacts we define every electrical terminal of the device on which the external input voltages are applied). Ohmic contacts are assumed to be *ideal*, they are equipotential surfaces and no voltage drop occurs at the interface between the contact and the neighbouring domain. This is well performed by suitable Dirichlet boundary conditions, therefore in the follow we indicate  $\partial\Omega_c = \Gamma_D$ .

$$\begin{aligned} \varphi &= \varphi_D \\ n &= n_D \\ p &= p_D \end{aligned} \quad \text{on } \Gamma_D. \tag{2.2}$$

we point that in the case of a perfect insulator domain, (2.2) reduces to the only condition on the electrostatical potential.

Artificial boundaries ( $\partial\Omega_a$ ) are needed in order to obtain a self-contained simulation domain. On these boundaries no electric and current flux is exchanged with the surrounding environment, this fact is well performed by homogeneous Neumann boundary condition ( $\partial\Omega_a = \Gamma_N$ )

$$\begin{aligned} \mathbf{D} \cdot \mathbf{n} &= 0 \\ \mathbf{J}_n \cdot \mathbf{n} &= 0 \\ \mathbf{J}_p \cdot \mathbf{n} &= 0 \end{aligned} \quad \text{on } \Gamma_N \quad (2.3)$$

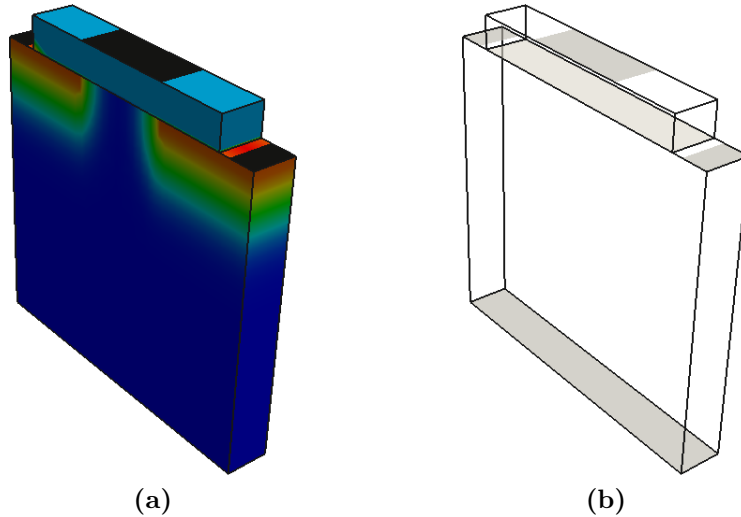
where  $\mathbf{n}$  is the outward unit normal vector defined over  $\partial\Omega$ . As we noted before on  $\partial\Omega_{ox} \cap \Gamma_N$  condition (2.3) is reduced to the first equation.

When oxide is present the silicon boundaries for continuity equations become

$$\begin{aligned} \Gamma_{D,Si} &= \Gamma_D \cap \partial\Omega_{Si} \\ \Gamma_{N,Si} &= \Gamma_N \cap \partial\Omega_{Si} \cup \Gamma_{int}. \end{aligned} \quad (2.4)$$

Fig.2.1 shows an example of boundary setting for a MOS device: in Fig.2.1a contacts are colored in black and in Fig.2.1b with light gray we indicate the interface between oxide and silicon.

Thermodynamical equilibrium and charge neutrality are the condition of an ideal contact. These conditions correspond to the follow algebraic system



**Figure 2.1:** (a) MOS device with net dopant concentration distributed accordingly to a gaussian profile and  $\Gamma_D$  colored in black. The oxide layer is colored in light blue. (b) Outline of the MOS device with  $\Gamma_{int}$  in light gray.



for  $n_D$  and  $p_D$

$$\begin{cases} p_D n_D & = n_i^2 \\ p_D - n_D + N_D - N_A & = 0 \end{cases} \quad (2.5)$$

Solving (2.5) on  $\Gamma_{D,Si}$  we have:

$$n_D = \frac{D + \sqrt{D^2 + 4n_i^2}}{2} \quad (2.6)$$

$$p_D = \frac{-D + \sqrt{D^2 + 4n_i^2}}{2} \quad (2.7)$$

where  $D := N_D - N_A$  is the net doping concentration. Furthermore at contact quasi fermi potential levels of silicon at contact are aligned with the external applied voltage. As a consequence we can easily calculate potential condition on  $\Gamma_{D,Si}$  using (1.8) and (1.9)

$$\varphi_D = \varphi_f + V_{th} \ln \left( \frac{n_D}{n_i} \right) = \varphi_f - V_{th} \ln \left( \frac{p_D}{n_i} \right) \quad (2.8)$$

where  $\varphi_f = -E_f/q$  is the unique quasi fermi potential level defined on contacts. When  $\Omega_{ox} \neq \emptyset$  we set  $\varphi_D$  equal to the external applied voltage on  $\Gamma_D/\Gamma_{D,Si}$ .

The stationary form of (1.51) can be now written in a closed form as:

$$\begin{aligned} -\Delta \epsilon \varphi - q(p - n) &= qD & \text{in } \Omega = \Omega_{ox} \cup \Omega_{Si} \\ \varphi &= \varphi_D & \text{on } \Gamma_D \\ \nabla \varphi \cdot \mathbf{n} &= 0 & \text{on } \Gamma_N \\ \\ \nabla \cdot (q\mu_n n \nabla \varphi - qD_n \nabla n) &= -qR & \text{in } \Omega_{Si} \\ n &= n_D & \text{on } \Gamma_{D,Si} \\ \nabla n \cdot \mathbf{n} &= 0 & \text{on } \Gamma_{N,Si} \\ \\ \nabla \cdot (-q\mu_p p \nabla \varphi - qD_p \nabla p) &= -qR & \text{in } \Omega_{Si} \\ p &= p_D & \text{on } \Gamma_{D,Si} \\ \nabla p \cdot \mathbf{n} &= 0 & \text{on } \Gamma_{N,Si} \end{aligned} \quad (2.9)$$

The high coupled nonlinear nature of system (2.9) makes an analytical treatment very difficult, if not even impossible. For this reason, numerical schemes must be used to compute an approximate solution.

## 2.2 Iteration algorithms

The most used algorithms are *the fully coupled Newton's method* and *the decoupled Gummel map*. System (2.9) can be written in a compact form as:

$$\mathbf{F}(\mathbf{U}) = \mathbf{0} \quad (2.10)$$

where:

$$\mathbf{U} := [\varphi, n, p]^T \quad \mathbf{F}(\mathbf{U}) := \begin{bmatrix} F_1(\mathbf{U}) \\ F_2(\mathbf{U}) \\ F_3(\mathbf{U}) \end{bmatrix} \quad (2.11)$$

and having set:

$$\begin{aligned} F_1(\mathbf{U}) &= \nabla \cdot (-\epsilon \nabla \varphi) - q(p - n + D) \\ F_2(\mathbf{U}) &= \nabla \cdot (q\mu_n n \nabla \varphi - qD_n \nabla n) + qR \\ F_3(\mathbf{U}) &= \nabla \cdot (-q\mu_p p \nabla \varphi - qD_p \nabla p) + qR \end{aligned}$$

Problem (2.10) is the generalization of the zero search for a real function  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Because the vector function  $\mathbf{F}$  is a nonlinear differential operator, the associated problem which we intend to resolve is: given a functional space  $V$  and the operator  $\mathbf{F} : V \rightarrow V$ , find  $\mathbf{U} \in V$  such that (2.10) is satisfied.

In our application, the function space  $V$  is typically a subset of the Sobolev space  $[H^1(\Omega)]^d$  (where  $d$  is the number of component of  $\mathbf{F}$ ). The general form of a Sobolev space for an integer  $m \geq 0$  is

$$H^m(\Omega) := \{v : D^\alpha v \in L^2(\Omega), \forall |\alpha| \leq m\}. \quad (2.12)$$

where  $L^2(\Omega)$  is the space of square integrable functions on  $\Omega$

$$L^2(\Omega) := \left\{ v : \int_{\Omega} |v|^2 d\Omega = \|v\|_{L^2(\Omega)}^2 < +\infty \right\}. \quad (2.13)$$

On these space, we shall use the semi-norm

$$|v|_{m,\Omega}^2 = \sum_{|\alpha|=m} \|D^\alpha v\|_{L^2(\Omega)}^2 \quad (2.14)$$

and the norm

$$\|v\|_{m,\Omega}^2 = \sum_{k \leq m} |D^k v|_{k,\Omega}^2 \quad (2.15)$$

We shall also need to consider functions that vanish on either the entire or a part of the boundary

$$H_0^1 := \{v : v \in H^1(\Omega), v|_{\partial\Omega} = 0\} \quad (2.16)$$

$$H_{0,\Gamma_D}^1 := \{v : v \in H^1(\Omega), v|_{\Gamma_D} = 0\} \quad (2.17)$$

For  $v \in H_0^1(\Omega)$  we have the *Poincaré inequality*

$$|v|_{0,\Omega} \leq C(\Omega)|v|_{1,\Omega} \quad (2.18)$$

and the seminorm  $|\cdot|_\Omega$  is therefore a norm in  $H^1(\Omega)$ , equivalent to  $\|\cdot\|_{1,\Omega}$ .

The above function spaces are used widely in the proceeding of this work, especially during the well-posedness analysis as reported in chapter 3.

### 2.2.1 Newton's method

**Definition 2.1** (Frechét differentiable). Let be  $X$  and  $Y$  two vector spaces. Given  $f, g \in X$  and a functional  $F : X \rightarrow Y$ , the functional  $F$  is Frechét differentiable if it exists a linear bounded operator  $A_f : X \rightarrow Y$  such that:

$$\lim_{\|g\| \rightarrow 0} \frac{\|F(f+g) - F(f) - A_f(g)\|_Y}{\|g\|_X} = 0 \quad (2.19)$$

If the limit exists, we write  $DF(f) = A_f$  and call it the Frechét derivative of  $F$  at  $f$ .

Considering the functional operator (2.11) we can easily compute the relative *Jacobian matrix*  $\mathbf{F}'$ , whose  $(i, j)$ -th entry represents the Frechét derivative of the  $i$ -th row of the non linear operator with respect to the  $j$ -th variable.

$$\mathbf{F}'_{ij}(\mathbf{U})[\mathbf{V}]_j := \lim_{\eta \rightarrow 0} \frac{F_i(\mathbf{U} + \eta[\mathbf{V}]_j) - F_i(\mathbf{U})}{\eta} \quad \mathbf{V} \in V \quad (2.20)$$

where  $[\mathbf{V}]_j \in V$  is the projection of  $\mathbf{V}$  in the  $j$ -th direction.

$\mathbf{F}'_{ij}(\cdot)$  is a linear operator from  $V$  into  $L(V, V)$ , while  $\mathbf{F}'_{ij}(\mathbf{U})$  is the Frechét derivative of the functional  $F_i$  respect the variable  $[\mathbf{U}]_j$ .

Accordingly with the above definitions the Newton's method reads as follows:

**Newton's method**

Let be  $X, Y$  two vector spaces and  $\mathbf{F} : X \rightarrow Y$  a function operator Frechèt differentiable, given an initial step  $\mathbf{U}^0 \in X$  and  $toll > 0$ , for all  $k \geq 0$  solve the following linear problem

$$\begin{aligned} \mathbf{F}'(\mathbf{U}^k)\delta\mathbf{U}^k &= -\mathbf{F}(\mathbf{U}^k) \\ \mathbf{U}^{k+1} &= \mathbf{U}^k + \delta\mathbf{U}^k \end{aligned} \quad (2.21)$$

until  $\|\mathbf{F}(\mathbf{U}^{k+1})\|_Y < toll$ , where  $\|\cdot\|_Y$  is a suitable norm on the space  $Y$ .

The application of Newton's method has transformed the original problem (2.10) into the *fixed-point problem* of finding  $\mathbf{U} \in V$  such that

$$\mathbf{U} = T_{\mathbf{F}}(\mathbf{U}) \quad (2.22)$$

where

$$T_{\mathbf{F}}(\mathbf{U}) = \mathbf{F}'(\mathbf{U})^{-1}(\mathbf{F}'(\mathbf{U})\mathbf{U} - \mathbf{F}(\mathbf{U})) \quad (2.23)$$

is the *iteration function* associated with the Newton method. The main result about the convergence of this method.

**Theorem 2.1.** *Let  $\mathbf{U} \in V$  be a solution of problem (2.10). Assume that  $\mathbf{F}'$  is Lipschitz continuos in the ball  $\mathcal{B}(\mathbf{U}, \delta)$ , i.e., that there exists  $K > 0$  such that:*

$$\|\mathbf{F}'(\mathbf{v}) - \mathbf{F}'(\mathbf{z})\|_{L(V,V)} \leq K\|\mathbf{v} - \mathbf{z}\|_V \quad \forall \mathbf{v}, \mathbf{z} \in \mathcal{B}(\mathbf{U}, \delta), \mathbf{v} \neq \mathbf{z} \quad (2.24)$$

*Then there exists in correspondence  $\delta' > 0$ , with  $\delta' \leq \delta$ , such that for all  $\mathbf{U}^0 \in \mathcal{B}(\mathbf{U}, \delta')$  the sequence  $\{\mathbf{U}^k\}$  generated by (2.21) converges quadratically to  $\mathbf{U}$ , i.e., there exists  $C > 0$  such that, for a suitable  $k_0 \geq 0$  we have:*

$$\|\mathbf{U} - \mathbf{U}^{k+1}\|_V \leq C\|\mathbf{U} - \mathbf{U}^k\|_V^2 \quad \forall k \geq k_0 \quad (2.25)$$

### 2.2.2 Fully coupled Newton's method

If we consider the linearization of the entire system (2.9) the relative Jacobian matrix is a 3x3 matrix and the problem reads as

$$\begin{bmatrix} F_{1,\varphi} & F_{1,n} & F_{1,p} \\ F_{2,\varphi} & F_{2,n} & F_{2,p} \\ F_{3,\varphi} & F_{3,n} & F_{3,p} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta n \\ \delta p \end{bmatrix} = \begin{bmatrix} -F_1(\varphi, n, p) \\ -F_2(\varphi, n, p) \\ -F_3(\varphi, n, p) \end{bmatrix}. \quad (2.26)$$

Each row of the above matrix is a PDE's equation which we can discretize with suitable numerical proceedings (i.e. finite element method). If we spent for example  $N_{dof}$  degrees of freedom to represent  $\delta\varphi$ ,  $\delta n$  and  $\delta p$  we note that the structure of the relative discretized matrix is a 3x3 block matrix system where every block is a  $N_{dof} \times N_{dof}$  matrix:

$$\begin{bmatrix} \mathbf{K}_{1,\varphi} & \mathbf{K}_{1,n} & \mathbf{K}_{1,p} \\ \mathbf{K}_{2,\varphi} & \mathbf{K}_{2,n} & \mathbf{K}_{2,p} \\ \mathbf{K}_{3,\varphi} & \mathbf{K}_{3,n} & \mathbf{K}_{3,p} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta n \\ \delta p \end{bmatrix} = \begin{bmatrix} -\mathbf{F}_1(\varphi, n, p) \\ -\mathbf{F}_2(\varphi, n, p) \\ -\mathbf{F}_3(\varphi, n, p) \end{bmatrix}. \quad (2.27)$$

This implies that at every iteration step we have to solve a linear problem of  $3 \times N_{dof}$  variables.

Moreover to ensure convergence of the Newton iterative process, it is important to provide a very good initial guess vector  $[\varphi_0, n_0, p_0]$ . Because the variable in play have different order of magnitude and the jacobian matrix is often quite ill-conditioned, appropriate scaling and balancing are needed in order to avoid problems associated with round-off error.

This method is widely used in commercial software especially for the strong result of convergence.

### 2.2.3 Gummel map algorithm

In 1964 H. K. Gummel proposed an original and alternative to (2.26) in order to solve the system (2.9) in a semiconductor device in one spatial dimension [Gum64].

The main idea of the algorithm is to move the nonlinearity to the Poisson equation only, and once obtained the electric potential profile, both continuity equations are linearized. This is possible if we consider the Maxwell-Boltzmann approximation for electrons (1.8) and holes (1.9) obtaining

$$F_1(\varphi) = \nabla \cdot (-\epsilon \nabla \varphi) - q(n_i(e^{((\varphi_p - \varphi)/V_{th})} - e^{((\varphi - \varphi_n)/V_{th})}) + D). \quad (2.28)$$

The Gummel algorithm is

**Decoupled Gummel map.**

0. Give a suitable initial condition for  $\varphi^0$  and set a positive parameter  $toll_{GM} > 0$  (Gummel Map tollerance)
1. Fix a positive parameter  $toll_{NLP} > 0$  (Non Linear Poisson tollerance), solve the linearized Non Linear Poisson equation (NLP) in  $\Omega$  using the Newton's method untill  $\|F_1(\varphi^{k+1})\| > toll_{NLP}$ :

$$\begin{cases} \nabla \cdot (-\epsilon_{Si} \nabla \delta \varphi^k) + \frac{1}{V_{th}} \sigma_{Si}^k \delta \varphi^k = f_{Si}^k & \text{in } \Omega_{Si} \\ \nabla \cdot (-\epsilon_{ox} \nabla \delta \varphi^k) = f_{ox}^k & \text{in } \Omega_{ox} \\ \delta \varphi^k = 0 & \text{on } \Gamma_D \\ \nabla \delta \varphi^k \cdot \mathbf{n} = 0 & \text{on } \Gamma_N \\ \varphi^{k+1} = \varphi^k + \delta \varphi^k \end{cases} \quad (2.29)$$

having set,

$$\begin{aligned} \sigma_{Si}^k(\varphi^k) &= qn_i \left[ e^{((\varphi_p - \varphi^k)/V_{th})} - e^{((\varphi^k - \varphi_n)/V_{th})} \right] \\ f_{Si}^k(\varphi^k) &= \nabla \cdot (-\epsilon \nabla \varphi^k) + qn_i \left[ e^{((\varphi_p - \varphi^k)/V_{th})} - e^{((\varphi^k - \varphi_n)/V_{th})} + D \right] \\ f_{ox}^k(\varphi^k) &= \nabla \cdot (-\epsilon \nabla \varphi^k) \end{aligned}$$

computed accordingly to the definition of the Frechét derivate.

2. Solve the Linear Electron Continuity Equation (LEC):

$$\begin{cases} \nabla \cdot (q\mu_n n \nabla \varphi^i - qD_n \nabla n) = -qR(n^{i-1}, p^{i-1}) & \text{in } \Omega_{Si} \\ n = n_D & \text{on } \Gamma_{D,Si} \\ \nabla n \cdot \mathbf{n} = 0 & \text{on } \Gamma_{N,Si} \end{cases} \quad (2.30)$$

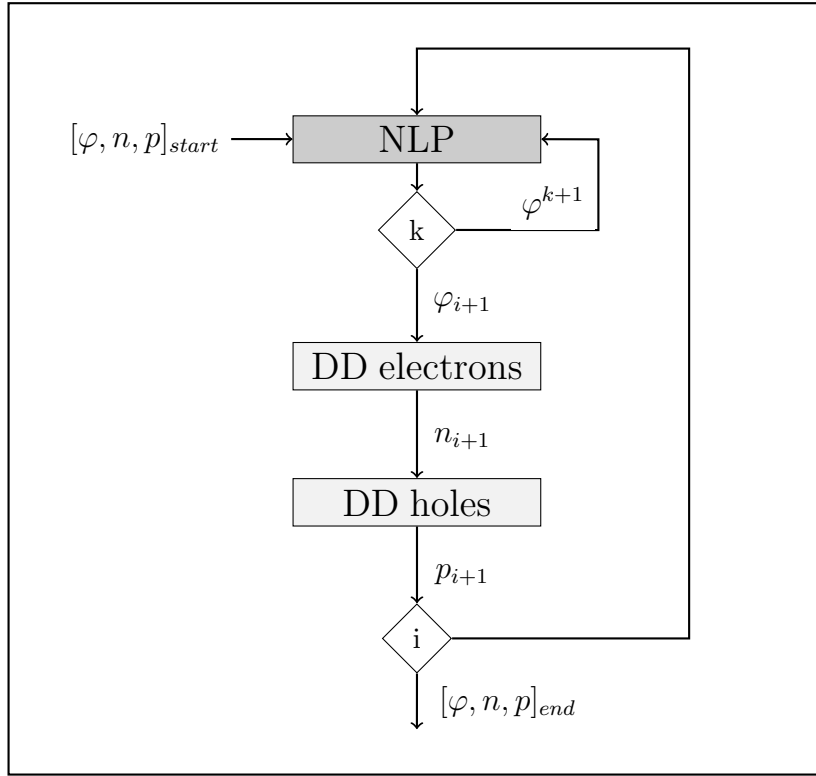
3. Solve the Linear Hole Continuity Equation (LHC):

$$\begin{cases} \nabla \cdot (-q\mu_p p \nabla \varphi^i - qD_p \nabla p) = -qR(n^{i-1}, p^{i-1}) & \text{in } \Omega_{Si} \\ p = p_D & \text{on } \Gamma_{D,Si} \\ \nabla p \cdot \mathbf{n} = 0 & \text{on } \Gamma_{N,Si} \end{cases} \quad (2.31)$$

4. If  $\max\{\|\varphi^i - \varphi^{i-1}\|_{L^\infty}, \|p^i - p^{i-1}\|_{L^\infty}, \|n^i - n^{i-1}\|_{L^\infty}\} > toll_{GM}$  restart from step (1).

We shall note with  $k$  the iteration step of the inner loop, while with  $i$  the iteration step of the Gummel Map: Fig.2.2 shows a concisely scheme. Unfortunately there isn't any convergence result for this method like (2.1), although there are several advantages which make Gummel map algorithm to the Fully Coupled Newton's Method. In fact simulations experience shows that the Gummel process is much more insensitive to the choice of the initial guess than Newton's method. This is particularly important in multidimensional problems where it is far from trivial to design a good starting point for initializing.

Another attractive feature is the reduced computational and memory cost: at each iteration step the Gummel algorithm requires the successive solution of three problems, each one of size equal to  $N_{dof} \times N_{dof}$ .



**Figure 2.2:** Gummel map algorithm

Let us discuss again on step 2-3 of *Decoupled Gummel map*. Accordingly with (1.52) the general R/G phenomenon can be separated in a reaction term and a force term (except for the II which is only a force term contribution). Considering

$$\begin{aligned} R_n^{i-1}(n) &= \sigma_n^{i-1}n - f^{i-1} \\ R_p^{i-1}(p) &= \sigma_p^{i-1}p - f^{i-1} \end{aligned} \quad (2.32)$$

where

$$\begin{aligned} \sigma_n &= \frac{p^{i-1}}{F(p^{i-1}, n^{i-1})} & \sigma_p &= \frac{n^{i-1}}{F(p^{i-1}, n^{i-1})} \\ f &= \frac{n_i^2}{F(p^{i-1}, n^{i-1})}. \end{aligned} \quad (2.33)$$

we can rewrite systems (2.30) and (2.31) as

$$\begin{cases} \nabla \cdot (q\mu_n n \nabla \varphi^i - qD_n \nabla n) + q\sigma_n^{i-1}n = qf^{i-1} & \text{in } \Omega_{Si} \\ n = n_D & \text{on } \Gamma_{D,Si} \\ \nabla n \cdot \mathbf{n} = 0 & \text{on } \Gamma_{N,Si} \end{cases} \quad (2.34)$$

$$\begin{cases} \nabla \cdot (-q\mu_p p \nabla \varphi^i - qD_p \nabla p) + q\sigma_p^{i-1}p = qf^{i-1} & \text{in } \Omega_{Si} \\ p = p_D & \text{on } \Gamma_{D,Si} \\ \nabla p \cdot \mathbf{n} = 0 & \text{on } \Gamma_{N,Si} \end{cases} \quad (2.35)$$

This splitting of R/G term is called *lagging approach* and corresponds to extending to the non-linear case the classical *Jacobi* method for the iterative solution of linear algebraic systems. Thus equations are sequentially solved, it's possible take advantage of the knowing a solution. Indeed an alternative approach would be use the solution of the first solved equation to compute the R/G contribute in the second equation. In such case the lagging method corresponds to extending to the nonlinear case the classical *Gauss-Seidel* method for the iterative solution in linear algebraic systems.



# Chapter 3

## Finite element discretization

In this section we present the classical variational formulation of problems (2.29), (2.34) and (2.35). For each kind of PDE problem we give a briefly presentation of the well-posedness analysis. Finally we describe the finite element discretization.

### 3.1 Non Linear Poisson Equation: weak form

Let us write problem (2.29) in a more compact form

$$\left\{ \begin{array}{ll} \nabla \cdot (-\epsilon \nabla \delta \varphi^k) + \sigma^k \delta \varphi^k &= f^k & \text{in } \Omega \\ \delta \varphi^k &= 0 & \text{on } \Gamma_D \\ \nabla \delta \varphi^k \cdot \mathbf{n} &= 0 & \text{on } \Gamma_N \\ \varphi^{k+1} &= \varphi^k + \delta \varphi^k \end{array} \right. \quad (3.1)$$

having set

$$\begin{aligned} \epsilon &= \epsilon_s \mathcal{I}_{\Omega_{Si}} + \epsilon_{ox} \mathcal{I}_{\Omega_{ox}} \\ f &= f_s \mathcal{I}_{\Omega_{Si}} + f_{ox} \mathcal{I}_{\Omega_{ox}} \\ \sigma &= \sigma_s \mathcal{I}_{\Omega_{Si}} \end{aligned}$$

where  $\mathcal{I}_A(\mathbf{x})$  is equal to 1 if  $\mathbf{x} \in A$  and 0 otherwise. System (3.1) is a classical Diffusion-Reaction (DR) problem in  $\Omega$ , with the respect to the variable  $\delta \varphi^k$ . Now we multiply the first equation in (3.1) with a test function  $v \in H_{\Gamma_D}^1$  and by integrating over all the domain we obtain

$$-\int_{\Omega} \epsilon \Delta \delta \varphi^k v \, d\Omega + \int_{\Omega} \sigma^k \delta \varphi^k v \, d\Omega = \int_{\Omega} f^k v \, d\Omega \quad \forall v \in H_{\Gamma_D}^1(\Omega). \quad (3.2)$$

Applying the Green-formula on (3.2) and then considering the boundary conditions, we get the weak formulation which reads as: find  $\delta\varphi^k \in H_{\Gamma_D}^1(\Omega)$  such that

$$\int_{\Omega} \epsilon \nabla \delta\varphi^k \nabla v \, d\Omega + \int_{\Omega} \sigma^k \delta\varphi^k v \, d\Omega = \int_{\Omega} f^k v \, d\Omega \quad \forall v \in H_{\Gamma_D}^1(\Omega) \quad (3.3)$$

We are able to define the following bilinear form

$$a : H_{\Gamma_D}^1(\Omega) \times H_{\Gamma_D}^1(\Omega) \rightarrow \mathbb{R}, \quad a(u, v) = \int_{\Omega} \epsilon \nabla u \nabla v \, d\Omega + \int_{\Omega} \sigma^k uv \, d\Omega \quad (3.4)$$

and the linear and bounded functional

$$F : H_{\Gamma_D}^1(\Omega) \rightarrow \mathbb{R}, \quad F(v) = \int_{\Omega} f^k v \, d\Omega \quad (3.5)$$

In order to prove the existence and uniqueness of the solution, we apply the *Lax-Millgram theorem* [Sal10] to the weak formulation (3.3). The well-posedness is ensured by several and physical hypotesis:

- $\epsilon \in L^\infty(\Omega)$  and  $\exists m$  s.t.  $0 < m \leq \epsilon$  (a.e.) in  $\Omega$ ;
- $\forall k \geq 0$   $\sigma^k \in L^\infty(\Omega)$  and  $\exists m$  s.t.  $0 < m \leq \sigma^k$  (a.e.) in  $\Omega_{S_i}$ .

We define some useful quantities:

$$\begin{aligned} \epsilon_M &= \max_{\Omega} \epsilon & \epsilon_m &= \min_{\Omega} \epsilon \\ \sigma_M &= \max_{\Omega} \sigma & \sigma_m &= \max_{\Omega} \sigma = 0 \end{aligned}$$

Take into account the above hypotesis it's possible to demonstrate:

- **Continuity of the bilinear form,**

$$\forall u, v \in H_{\Gamma_D}^1$$

$$\begin{aligned} \left| \int_{\Omega} \epsilon \nabla u \nabla v + \int_{\Omega} \sigma^k uv \right| &\leq \epsilon_M \|\nabla u\|_{L^2} \|\nabla v\|_{L^2} + \sigma_M \|u\|_{L^2} \|v\|_{L^2} \\ &\leq \max\{\epsilon_M, \sigma_M\} (\|\nabla u\|_{L^2} \|\nabla v\|_{L^2} + \|u\|_{L^2} \|v\|_{L^2}) \\ &\leq \max\{\epsilon_M, \sigma_M\} \|u\|_{H_{\Gamma_D}^1} \|v\|_{H_{\Gamma_D}^1} \end{aligned}$$

- **Coercivity of the bilinear form,**

$$\forall u \in H_{\Gamma_D}^1$$

$$\begin{aligned} \left| \int_{\Omega} \epsilon \nabla u \nabla u + \int_{\Omega} \sigma^k u^2 \right| &\geq \epsilon_m \|\nabla u\|_{L^2}^2 + \sigma_m \|u\|_{L^2}^2 \\ &= \epsilon_m \|\nabla u\|_{L^2}^2 \\ &= \epsilon_m \|\nabla u\|_{H_{\Gamma_D}^1}^2 \end{aligned}$$

- **Continuity of the functional,**

$$|\int_{\Omega} f^k v| \leq \|f^{(k)}\|_{L^2} \|v\|_{L^2} \quad \forall v \in H_{\Gamma_D}^1$$

Then we can state that  $\forall k \geq 0$  there exists a unique solution of the linearized Non Linear Poisson equation.

## 3.2 Continuity Equations: weak form

Without loss of generality we can consider only the electron continuity equation. System (2.34) is a classical diffusion-advection-reaction (DAR) problem written in conservative form. With a suitable change of variables we are able to treat these PDE's equations likewise the linearized Non Linear Poisson equation in the previous section. Consider the Slotboom variable (1.36), we can rewrite system (2.34) as:

$$\begin{cases} \nabla \cdot (-qD_n e^{(\varphi^i/V_{th})} \nabla u_n) + \sigma_n^{i-1} e^{(\varphi^i/V_{th})} u_n &= f^{i-1} & \text{in } \Omega_{Si} \\ u_n &= n_D e^{(-\varphi^i/V_{th})} & \text{on } \Gamma_{D,Si} \\ \nabla u_n \cdot \mathbf{n} &= 0 & \text{on } \Gamma_{N,Si} \end{cases} \quad (3.6)$$

We can easily obtain the weak formulation as section 3.1. Therefore the weak formulation of the Electron Continuity equation is: find  $u_n \in H_{\Gamma_{D,Si}}^1(\Omega)$  such that:

$$\int_{\Omega_{Si}} qD_n e^{(\varphi^i/V_{th})} \nabla u_n \nabla v \, d\Omega + \int_{\Omega_{Si}} \sigma_n^{i-1} e^{(\varphi^i/V_{th})} u_n v \, d\Omega = \int_{\Omega_{Si}} f^{i-1} v \, d\Omega \quad \forall v \in H_{\Gamma_{D,Si}}^1 \quad (3.7)$$

The existence and uniqueness of the unknown variable  $u_n$  ensures the same properties on  $n$ , thanks to the univocal relation between  $u_n$  and  $n$ . Further hypotheses on the coefficients  $\forall i \geq 0$ :

- $qD_n e^{(\varphi^i/V_{th})} \in L^\infty(\Omega_{Si})$  and  $\exists m$  s.t.  $0 < m \leq qD_n e^{(\varphi^i/V_{th})}$  (a.e.) in  $\Omega_{Si}$ ;
- $\sigma_n^{i-1} e^{(\varphi^i/V_{th})} \in L^\infty(\Omega_{Si})$  and  $\exists m$  s.t.  $0 < m \leq \sigma_n^{i-1} e^{(\varphi^i/V_{th})}$  (a.e.) in  $\Omega_{Si}$ .

We define the relative bilinear form

$$a(u, v) = \int_{\Omega_{Si}} qD_n e^{(\varphi^i/V_{th})} \nabla u_n \nabla v \, d\Omega + \int_{\Omega_{Si}} \sigma_n^{i-1} e^{(\varphi^i/V_{th})} u_n v \, d\Omega \quad (3.8)$$

and the linear and bounded functional

$$F(v) = \int_{\Omega_{Si}} f^{i-1} v \, d\Omega \quad (3.9)$$

Now the well-posedness of this problem is ensured just by following the procedure of section 3.1.

### 3.3 Numerical approximation

In this section we introduce the classical Galerkin's method to approximate the weak formulation on  $\Omega$ . Each weak formulation could be represented in a more compact and generic form as, find  $u \in V$  such that

$$a(u, v) = F(v) \quad \forall v \in V \quad (3.10)$$

where  $V$  is the space of admissible functions, e.g.  $H_{\Gamma_D}^1(\Omega)$ ,  $H_{\Gamma_D, Si}^1(\Omega_{Si})$ . Let us introduce  $V_h$  which is a family of finite-dimensional subspace of  $V$ , depending by a positive parameter  $h$ , such that

$$V_h \subset V, \quad \dim V_h < \infty \quad \forall h > 0 \quad (3.11)$$

The *Galerkin's problem* reads as, find  $u_h \in V_h$  such that:

$$a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h \quad (3.12)$$

Let be  $\mathcal{T}_h$ , a finite partition of  $\Omega$ , and  $K$  a generic element of  $\mathcal{T}_h$  such that  $\bar{\Omega} = \bigcup \bar{K}$ . In this case the parameter  $h$  refers to the characteristic dimension of the elements  $K$ . Let us introduce the general finite element spaces of the polynomial element-wise functions:

$$X_h^r(\Omega) := \{v_h \in C^0(\bar{\Omega}) : v_h|_K \in \mathbb{P}_r, \forall K \in \mathcal{T}_h\} \quad (3.13)$$

and the relative space where functions vanish on boundaries

$$X_{h, \Gamma_D}^r(\Omega) := \{v_h \in X_h^r : v_h|_{\Gamma_D} = 0\}. \quad (3.14)$$

If  $\Omega \in \mathbb{R}^3$  we have:

$$\dim \mathbb{P}_r := \frac{(r+1)^3}{2} + \frac{(r+1)^2}{2} + \frac{r(r+1)(2r+1)}{12} - \frac{r(r+1)^2}{2} - \frac{r(r+1)}{4} \quad (3.15)$$

More precisely we approximate  $H_{\Gamma_D}^1(\Omega)$  with  $X_{h, \Gamma_D}^1(\Omega)$  and  $H_{\Gamma_D, Si}^1(\Omega_{Si})$  with  $X_{h, \Gamma_D, Si}^1(\Omega_{Si})$ . Therefore accordingly with (3.15) we have

$$\begin{aligned}
\dim \mathbb{P}_1 &= 4 \\
\dim X_h^1 &= N_h \\
\dim X_{h,\Gamma_D}^1 &= N_h - N_g
\end{aligned}$$

where  $N_h$  is the number of vertices of the partition  $\mathcal{T}_h$  and  $N_g$  are the number of vertices lie on Dirichlet boundaries.

We denote by  $\{\psi_j\}_{j=1}^{N_h}$  the Lagrangian basis of the space  $X_h^1$ . Naturally as  $u_h \in X_h^1$  there are  $u_j \in \mathbb{R}$  with  $j = 1, \dots, N_h$  such that:

$$u_h = \sum_{j=1}^{N_h} u_j \psi_j \quad (3.16)$$

Because each functions of  $V_h$  is a linear combination of  $\psi_i$ , we can test equation (3.12) only for each basis function rather than  $\forall v_h \in V_h$ . The result of the complete discretization is find  $u_j$ , with  $j = 1, \dots, N_h$  such that:

$$\sum_{j=1}^{N_h} u_j a(\psi_j, \psi_i) = F(\psi_i) \quad \forall i = 1, \dots, N_h \quad (3.17)$$

In order to implement this routine it's useful make explicit the subdivision of the bilinear form on the element of the partition  $\mathcal{T}_h$ :

$$\sum_{j=1}^{N_h} u_j \sum_{K \in \mathcal{T}_h} a_K(\psi_j, \psi_i) = \sum_{K \in \mathcal{T}_h} F_K(\psi_i) \quad \forall i = 1, \dots, N_h \quad (3.18)$$

### 3.3.1 Geometrical discretization

Each element  $K \in \mathcal{T}_h$  is set as a tetrahedral of volume  $|K|$ ; Let  $\delta > 0$  be a constant such that:

$$\frac{h_K}{\rho_K} \leq \delta \quad \forall K \in \mathcal{T}_h \quad (3.19)$$

where  $h_k = \text{diam}(K) = \max_{x,y \in K} |x - y|$  and  $\rho_K$  is the diameter of the sphere inscribed in the tetrahedral  $K$ . Condition (3.19) is the so called *mesh regularity condition* [Qua08] and it ensures an istoropic partition. We denote with  $\mathcal{E}_h$ ,  $\mathcal{V}_h$  and  $\mathcal{F}_h$  the set of all the edges, vertices and faces of  $\mathcal{T}_h$  respectively, and for each  $K \in \mathcal{T}_h$  we denote by  $\partial K$  and  $\mathbf{n}_{\partial K}$  the boundary of the element and its outward unit normal.

We note that  $\mathcal{T}_h$  is built such that every  $K$  belongs to a single region, while it is possible that vertices belong to different regions.

### 3.3.2 Linearized Non Linear Poisson equation

As regards the linearized NLP equation we have:

$$a(\psi_j, \psi_i) = \int_{\Omega} \epsilon \nabla \psi_j \nabla \psi_i d\Omega + \int_{\Omega} \sigma^k \psi_j \psi_i d\Omega \quad (3.20)$$

and the relative restriction on element  $K$  is

$$a_K(\psi_j, \psi_i) = \int_K \epsilon \nabla \psi_j \nabla \psi_i dK + \int_K \sigma^k \psi_j \psi_i dK \quad (3.21)$$

Equation (3.21) it's formed by two distinct contributions, the former identifies the diffusive contribution and generates the so called *stiffness matrix*, while the latter refers to the reaction and generates the *mass matrix*.

The coefficient  $\epsilon$  is a piece-wise constant function, which changes on different material regions. Therefore  $\epsilon$  is constant over each elements and integral in (3.21) become easier.

As a consequence of choose the discrete space  $X_h^1$ , we can't expect a better priori estimation error on the solution, than the first order in  $\|\cdot\|_{1,\Omega}$  respect the characteristic discretization step  $h_K$  [Qua08]. This implies that is not necessary and useful the use of an high order quadrature, and the trapezoidal rule is enough accurate. The main consequence of using trapezoidal quadrature rule is that extra-diagonal elements of the mass-matrix disappear. This technique is well known as *lumping procedure* applied on the mass-matrix.

Finally the contributions of the local system matrix  $A_K^k$  is:

$$[A_K^k]_{ij} = \epsilon_K L_{ij} + \frac{|K|}{4} \sigma_i^k \quad (3.22)$$

having set

$$\begin{aligned} L_{ij} &\simeq \int_K \nabla \psi_i \nabla \psi_j d\Omega \\ \sigma_i^k &= \sigma^k(\mathbf{x}_i) \end{aligned} \quad (3.23)$$

The construction of the right hand side of (3.18) with trapezoidal rule is:

$$[F_K]_i^k = f_i^k |K|/4 \simeq \int_{\Omega} f^k \psi_i d\Omega \quad (3.24)$$

The local contributions of each element  $K$  must be assembled in the global matrix  $A$ : let  $I$  be the global index of a generic vertex belonging to the partition  $\mathcal{T}_h$ , we denote by  $\mathcal{J}_K : \mathcal{V}_{\mathcal{T}_h} \rightarrow \mathcal{V}_K$  the map which connects  $I$  to its corresponding local index  $i = 1, \dots, 4$  in the element  $K$ . Then we have

$$A_{IJ}^k = \sum_{\substack{\forall K \in \mathcal{T}_h \text{ s.t.} \\ \mathcal{J}_K(I), \mathcal{J}_K(J) \subset \mathcal{V}_K}} [A_K]_{ij}^k \quad (3.25)$$

analogously for the force term  $\mathbf{b}^{(k)}$ :

$$b_I^k = \sum_{\substack{\forall K \in \mathcal{T}_h \text{ s.t.} \\ \mathcal{J}_K(I) \subset \mathcal{V}_K}} [F_K]_i^k \quad (3.26)$$

Once we have built the global matrix  $A^k$  and the global vector  $\mathbf{b}^k$  we need to take into account the essential boundary conditions. In fact the displacement formulation is a primal formulation which forces Dirichlet boundary condition in a strong way. Therefore we have to modify the algebraic system. We choose the *diagonalization* technique which does not alter the matrix pattern nor introduce ill-conditioning for the system. Let  $i_D$  be the generic index of a Dirichlet node, we denote by  $[\delta\varphi_D]_i$  (which in this case is equal to zero) the known value of the solution  $\delta\varphi$  at the node. We consider the Dirichlet condition as an equation of the form  $a[\delta\varphi]_i = a[\delta\varphi_D]_i$ , where  $a \neq 0$  is a suitable coefficient. In order to avoid degrading of the global matrix condition number, we take  $a$  equal to the diagonal element of the matrix at the row  $i_D$ .

Finally we have completed the discretization of (Step 1), which reads as follows:

$$\begin{cases} A^k \delta\varphi^k &= \mathbf{b}^k \\ \varphi^{k+1} &= \varphi^k + \delta\varphi^k \end{cases} \quad (3.27)$$

As every iteration procedure, problem (3.27) needs a suitable convergence break criterion. A good method is based on checking the satisfaction of the fixed point equation (2.10) by the  $k$ -th solution. In this case the inner loop of the Gummel Map reads as: given a tolerance  $toll > 0$  solve problem (3.27) untill:

$$\|\mathbf{b}(\varphi^{k+1})\|_2 > toll \quad (3.28)$$

where  $\|\cdot\|_2$  is the usual Euclidean norm for a vector.

## Damping

Nevertheless the theorem (2.1), the system (3.27) may be affected by difficulties on the convergence velocity. The main problem associated with the classical Newton method is the tendency to overestimate the length of the present correction step. This phenomenon is frequently indicated as *overshoot*. In the case of the semiconductor equations this overshoot problem

has often been treated by simply limiting the size of the correction vector ( $\delta\varphi$ ) determined by Newton's method. The usual established modifications to avoid overshoot are given by the follow formulations:

$$\tilde{A}(\varphi_k) = \frac{1}{t_k} A(\varphi_k) \quad (3.29)$$

$t_k$  is a properly chosen positive parameter: for  $t_k = 1$  the modified Newton method reduces to the classical Newton method. We have now to deal with the question how to choose  $t_k$  such that the modified Newton method exhibits superior convergence properties compared to the classical Newton method. For the case (3.29) there's a simple criterion suggested by Deuffhard [Deu74].  $t_k$  is taken from the interval  $(0, 1]$  in such a manner that for any norm,

$$\|A(\varphi_k)^{-1} \mathbf{b}(\varphi_k - t_k A(\varphi_k)^{-1} \mathbf{b}(\varphi_k))\| < \|A(\varphi_k)^{-1} \mathbf{b}(\varphi_k)\| \quad (3.30)$$

Condition (3.30) guarantees that the correction of the  $k$ -th iterate is an improved approximation to the final solution, in other words the residual norm can only descends. This condition is hardly to be evaluated as the presence of the inverse system matrix. If the Jacobian matrix is factored into triangular matrices the evaluation of the argument of the norm on the left hand side of (3.30) is reduced to a forward and backward substitution and the evaluation of  $\mathbf{b}(\varphi)$ . Although we use an iterative method (BCG solver based on [PTVF07]) which implies serious difficulties to the application of the above criterion. Another valid possibility is to use the main diagonal of  $A(\varphi_k)$ , denoted as  $D(\varphi_k)$ :

$$\|D(\varphi_k)^{-1} \mathbf{b}(\varphi_k - t_k D(\varphi_k)^{-1} \mathbf{b}(\varphi_k))\| < \|D(\varphi_k)^{-1} \mathbf{b}(\varphi_k)\| \quad (3.31)$$

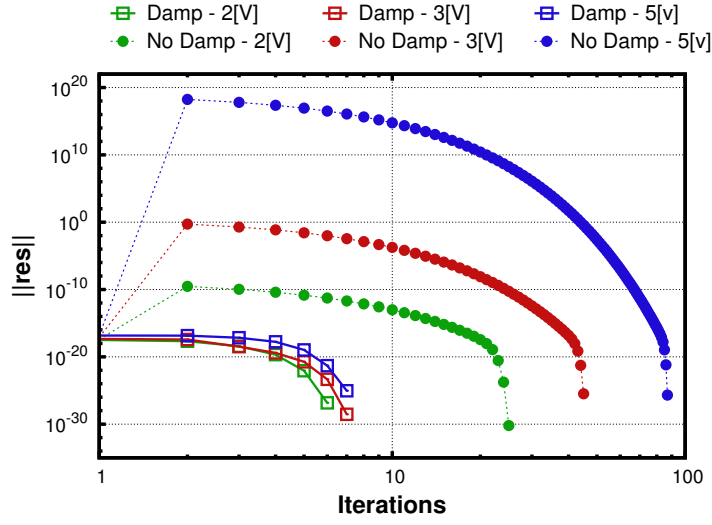
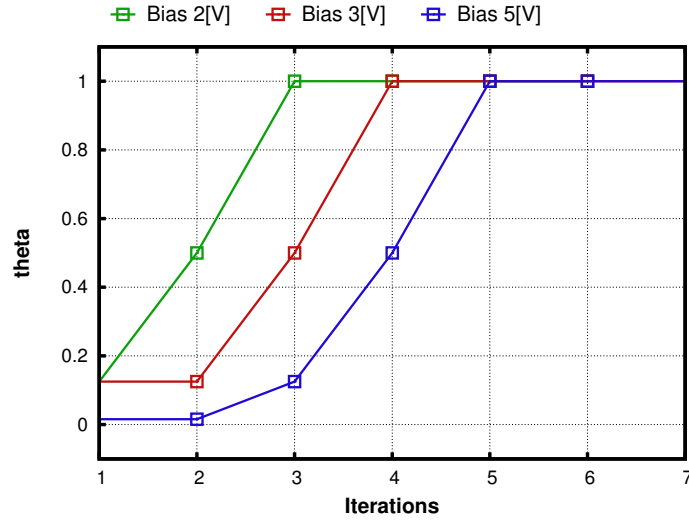
This criterion has been adopted in our code. However the value to use for  $t_k$  is a question of trial and error. Frequently the following sequence is used:

$$t_k = \frac{1}{2^i} \quad (3.32)$$

$$t_k = \frac{1}{\frac{i(i+1)}{2}} \quad (3.33)$$

where  $i$  is the subiterations of damping reached when (3.31) is satisfied. Close to the solution, (3.30) (and so (3.31)) will be satisfied with  $t_k = 1$  so that the convergence properties of the classical Newton method are recovered.



(a) *Non Linear Poisson residual: damping benefit prodedure.*(b)  *$t_k$  parameter.*

**Figure 3.1:** (a) Number of iteration against residual for different voltages in a diode test case. (b) Magnitude of the damping parameter  $t_k$ .

### 3.3.3 Continuity equations

As regards the (3.6) equation we can write the bilinear form as

$$a(u, v) = \int_{\Omega_{Si}} q D_n e^{(\varphi^i/V_{th})} \nabla \psi_j \nabla \psi_i d\Omega + \int_{\Omega_{Si}} \sigma_n^{i-1} e^{(\varphi^i/V_{th})} \psi_j \psi_i d\Omega \quad (3.34)$$

Even if this form guarantees an easily analysis of well-posedness, the choice of using Slotboom variables  $u_n$  and  $u_p$  causes the onset of overflow problems due to the evaluation of  $\exp(\varphi/V_{th})$ , which can be a rapidly varying function according to the behaviour of the potential  $\varphi$ .

Therefore special care has to be taken in the treatment of the diffusion coefficient. In view of further discussions of this issue, we introduce some useful notation. For each set  $S \subset \Omega$  having measure  $|S|$ , we introduce the following averages of a given function  $g$  that is integrable on  $S$ :

$$\mathcal{M}_S(g) = \frac{\int_S g dS}{|S|} \quad \mathcal{H}_S = (\mathcal{M}_S(g^{-1}))^{-1}$$

Notice that  $\mathcal{M}_S$  is the usual integral average, while  $\mathcal{H}_S$  is the *harmonic average*. It is well-known that the use of the harmonic average provides a superior approximation performance [IE83].

The weak form (3.34) is the result of a displacement approach which is the most classical way of setting these problems, although different variational formulations and therefore different finite element approximations may be used, like a primal mixed approach (PM) (for a more complete treatment see [ZL12] and [DF05]). First of all it's convenient to reformulate problem (2.34) by using the relation (1.36) and (1.38) in a more generic form. This yields the following equivalent form

$$\begin{cases} \nabla \cdot (\mathbf{J}_n(n)) + \sigma n = f & \text{in } \Omega_{Si} \\ \mathbf{J}_n = qD_n e^{(\varphi/V_{th})} \nabla (e^{(-\varphi/V_{th})} n) & \text{in } \Omega_{Si} \\ n = n_D & \text{on } \Gamma_{D,Si} \\ \mathbf{J}_n \cdot \mathbf{n} = 0 & \text{on } \Gamma_{N,Si} \end{cases} \quad (3.35)$$

We report here the weak formulation of (3.35) which is well investigated in [DF05], find  $\mathbf{J}_n \in [L^2(\Omega)]^d$  and  $n \in H_{\Gamma_{D,Si}}^1(\Omega)$  such that

$$-\int_{\Omega_{Si}} \mathbf{J}_n \cdot \nabla v d\Omega + \int_{\Omega_{Si}} \sigma n v d\Omega = \int_{\Omega_{Si}} f v d\Omega \quad \forall v \in H_{\Gamma_{D,Si}}^1(\Omega) \quad (3.36)$$

$$\int_{\Omega_{Si}} (qD_n e^{(\varphi/V_{th})})^{-1} \mathbf{J}_n \cdot \mathbf{q} d\Omega + \int_{\Omega_{Si}} \nabla (e^{(\varphi/V_{th})} n) \cdot \mathbf{q} d\Omega = 0 \quad \forall \mathbf{q} \in [L^2(\Omega)]^d \quad (3.37)$$

In order to approximate  $[L^2(\Omega)]^d$  we introduce a new discrete space

$$\Sigma_h := \{\mathbf{q}_h \in [L^2(\Omega)]^d : \mathbf{q}|_K \in [\mathbb{P}_0]^d \forall K \in \mathcal{T}_h\} \quad (3.38)$$

as usual  $d$  is the dimension of  $\Omega$  and if  $d = 3$ ,  $\mathbf{q}_h$  is characterized for every  $K \in \mathcal{T}_h$  by the triplet

$$\mathbf{q}_{1,2,3}^h = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \quad (3.39)$$

Therefore (3.36) and (3.37) can be restricted on a generic element  $K$  and the related bilinear form reads

$$\begin{cases} a_h^K(n_h, v_h) &= \int_K \mathbf{J}_{n,h}^K(n_h) \nabla v_h \, dK + \int_K \sigma n_h v_h \, dK \\ F(v_h)^K &= \int_K f v_h \, dK \\ \mathbf{J}_{n,h}^K &= D_K(qD_n e^{(\varphi/V_{th})}) \nabla(e^{(-\varphi/V_{th})} n_h) \end{cases} \quad (3.40)$$

where  $D_K \in \mathbb{R}^{3 \times 3}$  is the element stiffness matrix. Several treatments may be performed on this matrix

$$D_K(qD_n e^{(\varphi/V_{th})}) = \begin{cases} \mathcal{M}_K(qD_n e^{(\varphi/V_{th})}) \\ \mathcal{H}_K(qD_n e^{(\varphi/V_{th})}) \\ \frac{1}{|K|} \sum_{i=1}^6 \mathcal{H}_{e_i}(qD_n e^{(\varphi/V_{th})}) |e_i| \mathbf{s}_i \mathbf{t}_i \mathbf{t}_i^T \end{cases} \quad (3.41)$$

These different approaches in the computation of the average of the diffusion coefficient are responsible for the quite different numerical performance of the relative methods. We already presented the standard average and the harmonic average and we discussed briefly the advantages of them. The latter equation in (3.41) introduces an exponential treatment of the diffusion coefficient along each edge of the boundary  $\partial K$  of the subdomain  $K$ .

Considering that along the edges the approximate flux density can be written as a function of its tangential components. We have for each edge  $e_i$ , the tangential component of  $J_h^K(n_h)$

$$\begin{aligned} j_{e_i} &= \mathcal{H}_{e_i} \frac{\delta_i(e^{-\varphi/V_{th}} n_h)}{|e_i|} = \mathcal{H}_{e_i} \nabla(e^{-\varphi/V_{th}} n_h) \mathbf{t}_i \\ &= \mathcal{H}_{e_i}(qD_n e^{(\varphi/V_{th})}) \frac{\mathcal{B}(\delta_i(\varphi/V_{th})) n_{h,k} - \mathcal{B}(-\delta_i(\varphi/V_{th})) n_{h,j}}{|e_i|} \end{aligned}$$

where

$$\delta_i(\varphi/V_{th}) = \frac{\varphi_k - \varphi_j}{V_{th}} = 2 \frac{(\mathbf{E}_K \cdot \mathbf{t}_{e_i})|e_i|}{2\mathcal{H}_{e_i}(qD_n e^{(\varphi/V_{th})})} = 2\gamma_i \quad (3.42)$$

$$\mathcal{B}(z) = \begin{cases} \frac{z}{e^z - 1} & z \neq 0 \\ 1 & z = 0 \end{cases} \quad (3.43)$$

being  $\mathbf{E}_K$  the relative electric field on  $K$  and  $|\gamma_i|$  the Péclet number associated with the edge  $e_i$ . From (3.40) we immediatly obtain:

$$\mathbf{J}_h^K = \frac{1}{|K|} \sum_{i=1}^6 |e_i| s_i j_{e_i} \mathbf{t}_i \quad (3.44)$$

Furthermore having defined the flux vector over  $K$  in the form (3.44), it is possible to construct a family of Galerkin finite element approximations for the continuity equations by a proper choice of the quantities  $j_{e_i}$  (e.g. upwind techniques).

### The discretization scheme

Given the choice for  $j_{e_i}$  and replacing the equation for  $\mathbf{J}_h^K$  in the bilinear form (3.40), we can compute the local system matrix as

$$\Phi_K = \begin{bmatrix} - \begin{pmatrix} a_{e12} \mathcal{B}_{12} L_{21}^K + \\ a_{e13} \mathcal{B}_{13} L_{31}^K + \\ a_{e14} \mathcal{B}_{14} L_{41}^K \end{pmatrix} & a_{e12} \mathcal{B}_{12} L_{21}^K & a_{e13} \mathcal{B}_{13} L_{31}^K & a_{e14} \mathcal{B}_{14} L_{41}^K \\ a_{e21} \mathcal{B}_{21} L_{12}^K & - \begin{pmatrix} a_{e21} \mathcal{B}_{21} L_{12}^K + \\ a_{e23} \mathcal{B}_{23} L_{32}^K + \\ a_{e24} \mathcal{B}_{24} L_{42}^K \end{pmatrix} & a_{e23} \mathcal{B}_{23} L_{32}^K & a_{e24} \mathcal{B}_{24} L_{42}^K \\ a_{e31} \mathcal{B}_{31} L_{31}^K & a_{e31} \mathcal{B}_{32} L_{32}^K & - \begin{pmatrix} a_{e31} \mathcal{B}_{31} L_{31}^K + \\ a_{e32} \mathcal{B}_{32} L_{32}^K + \\ a_{e34} \mathcal{B}_{34} L_{34}^K \end{pmatrix} & a_{e34} \mathcal{B}_{34} L_{34}^K \\ a_{e41} \mathcal{B}_{41} L_{41}^K & a_{e42} \mathcal{B}_{42} L_{42}^K & a_{e43} \mathcal{B}_{43} L_{43}^K & - \begin{pmatrix} a_{e41} \mathcal{B}_{41} L_{41}^K + \\ a_{e42} \mathcal{B}_{42} L_{42}^K + \\ a_{e43} \mathcal{B}_{43} L_{43}^K \end{pmatrix} \end{bmatrix} \quad (3.45)$$

$$A_K = \Phi_K + \frac{|K|}{4} \text{diag}(\sigma) \quad (3.46)$$

$$\mathbf{F}_K = \frac{|K|}{4} (f_1, f_2, f_3, f_4)^T \quad (3.47)$$

denoting by  $\mathcal{B}_{ij}$  the Bernoulli function applied to the potential difference between node  $j$  and node  $i$ .

The discretization scheme just presented is well known as *Edge Averaged Finite Elements* (EAFE) and it's particularly suitable for problems with highly variable coefficient. Furthermore this approach has several good

properties, e.i. in 2D simulation if  $\mathcal{T}_h$  is a Delaunay partition the system matrix is an *M-matrix* [BCC98]. The main consequence of this statement is that the solution could satisfy the *Discrete Maximum Principle*. This is a notable property which implies that no negative concentrations are admitted. Unfortunately this property is not anymore valid in 3D framework, indeed the Delaunay condition of the mesh is not sufficient to guarantee that the system matrix is an M-matrix. A more general condition is presented in [XZ99].

**Theorem 3.1** (Zikatanov condition). *The system matrix of the EAFE scheme is an M-matrix if and only if for any fixed edge  $E$  of the partition  $\mathcal{T}_h$  the following inequality holds:*

$$\omega_E = \frac{1}{n(n-1)} \sum_{T \supset E} |k_E^T| \cot \theta_E^T \geq 0, \quad (3.48)$$

where  $n$  is the dimension,  $\sum_{T \supset E}$  means summation over all simplexes  $T$  containing  $E$ ,  $\theta_E^T$  is the angle between the faces  $f_i, f_j \in \mathcal{T}_h$  such that  $f_i \cap f_j = E$  and  $k_E^T$  is the edge in  $T$  which doesn't share any vertices with  $E$ .

*Observation 3.1.* For  $n = 2$  the condition (3.48) means that the sum of the angles opposite to any edge is less than or equal to  $\pi$ , this condition implies that the partition is a Delaunay triangulation.

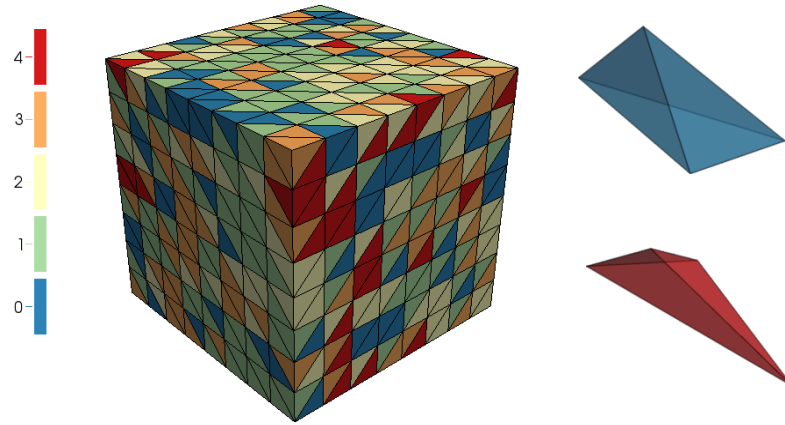
*Observation 3.2.* Condition (3.48) highlights that in order to satisfy the discrete maximum principle, a partition without obtuse angles is preferable.

We remark that presently meshing algorithm are oriented to care about the minimum angle of the elements, rather than the maximum, resulting that obtain a mesh which satisfied the condition (3.48) it's a really difficult task.

Fig.3.2 shows a simple partition of a cube performed with the Synopsis tool SNMESH. For every element we evaluated how many edges don't satisfy the condition (3.48). It's clear that there are a lot of edges which don't fulfill the condition and a precise pattern can't be individuated. When several bad edges belong to a single element we can indentify the presence of many obtuse angles.

In order to avoid this problem some alternative solutions are proposed in the literature, like the *Orthogonal Subdomain Collocation method* [PC98], but also this approach is not definitely.

Therefore in presence of negative concentration the most used technique during 3D numerical simulation is the increasing of the degree of freedom in the problematic regions, which often are the ones where the carrier density decrease.



**Figure 3.2:** Evaluation of the Zikatanov condition over a simple partition. Red elements doesn't satisfy condition (3.48) over four edges while blue elements fully accomplished the criterion.

# Chapter 4

## Simulation results

In this chapter we present the work done in order to validate the numerical implementation of the discretization method illustrated in chapter 3 compared with a reference simulation tool (SDEVICE commercialized by Synopsys [Sde13]). In particular we illustrate (and compare) the algorithm used to calculate current at the contacts.

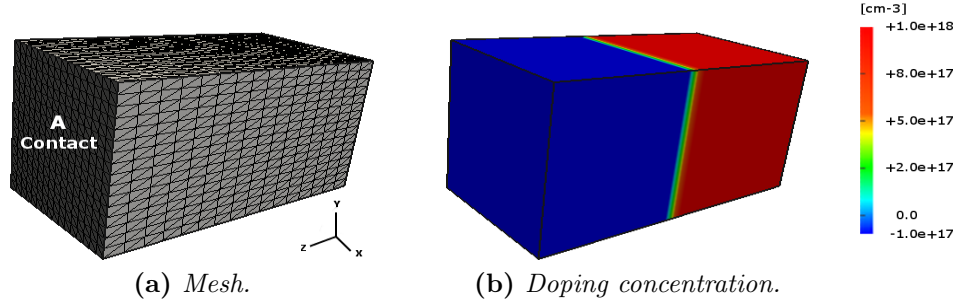
### 4.1 Test cases

We consider three kind of semiconductor devices:

- **p-n junction**
- **p-n junction in oxide**
- **MOSFET n-channel / p-channel**

#### 4.1.1 p-n junction

In this example we consider a simple p-n junction. Fig.4.1 presents the partition and the doping profile for this test. The section of the parallelepiped is a  $0.05 \times 0.05[\mu m^2]$  square while the device is  $0.1[\mu m]$  long. The number of vertices are 4933, while the elements are 24576. The doping concentration is obtained setting a constant profile of acceptors over all the domain ( $N_A = 1.0 \times 10^{17}$ ) overwhelmed by a doping profile of donors ( $N_D = 1.0 \times 10^{18}$ ) bounded on one side of the device resulting in an almost abrupt junction. Two contacts are defined: (A) contact is placed at  $Z = 0.1[\mu m]$  and (B) contact is placed at  $Z = 0.0[\mu m]$ .

**Figure 4.1:** p-n junction.

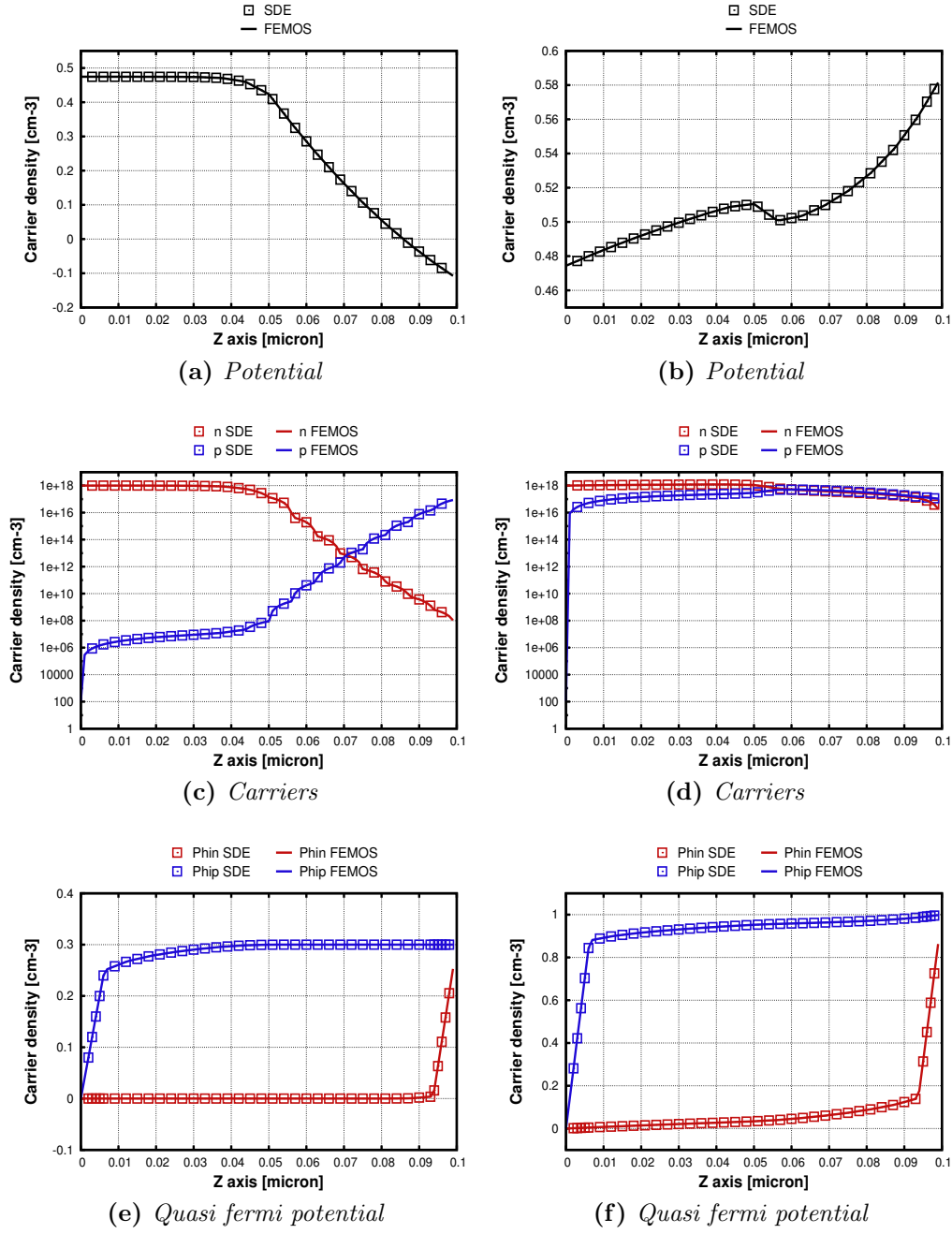
In order to analyze the operating function of the diode, two cases of direct polarization are performed: 0.3[V]-1.0[V]. The setting values and the parameters are summarized in Tab.4.1. Fig.4.2 reports the solutions for  $V_A = 0.3[V]$  and  $V_A = 1.0[V]$ , along a line parallel to the Z-axis and placed at the center of the device. Because the built-in voltage is around  $0.7 \div 0.8[V]$  the behaviour of the device is different when the applied bias is below or above this threshold. At 0.3[V] of polarization, potential drop is almost bounded around the junction, and due to the asymmetric doping, is major extended in the p-side. The carriers can't cross the potential barrier and this will cause a low current flux inside the device. At 1.0[V] the minority carriers density becomes almost ten order bigger, resulting in a large amount of current toward the contacts: the device turns from exponential to linear resistive. This is clear in Fig.4.2b where the potential shape becomes similar to a resistance voltage profile (linear potential profile). Comparing the quasi-fermi potential of Fig.4.2e and Fig.4.2f the boundary layers at contacts increase with the polarization. This effect is related to the ohmic contact hypothesis, and can be avoided with different boundary condition: this occurs also for the carrier concentration because at the contacts the charge neutrality and the thermodynamic equilibrium are imposed.

Fig.4.3÷4.5 shows the comparison between SDEVICE and FEMOS in 3D plots for electrostatic potential, electron and hole densities at 0.3[V], while Fig.4.6÷4.8 show the same comparison at 1.0[V]. In both of the condition the agreement is very good.

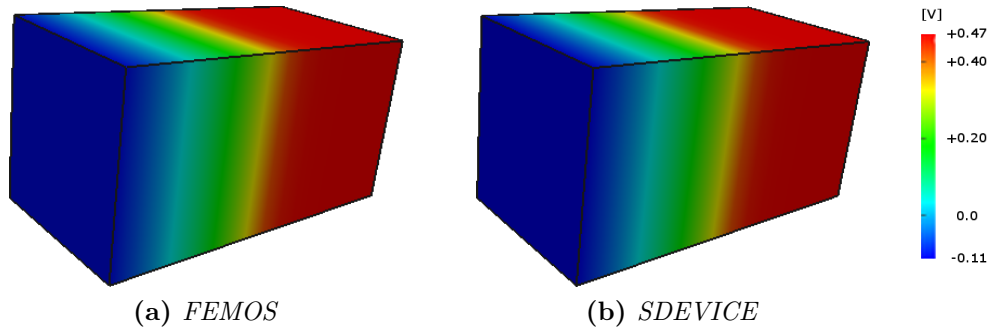
Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$
$V_A = 0.3$	$\mu_n = 1417, \mu_p = 470.5$	SRH, Auger	11.6
$V_A = 1.0$	$\mu_n = 1417, \mu_p = 470.5$	SRH, Auger	11.6

**Table 4.1:** p-n junction - list of settings, parameters and models.

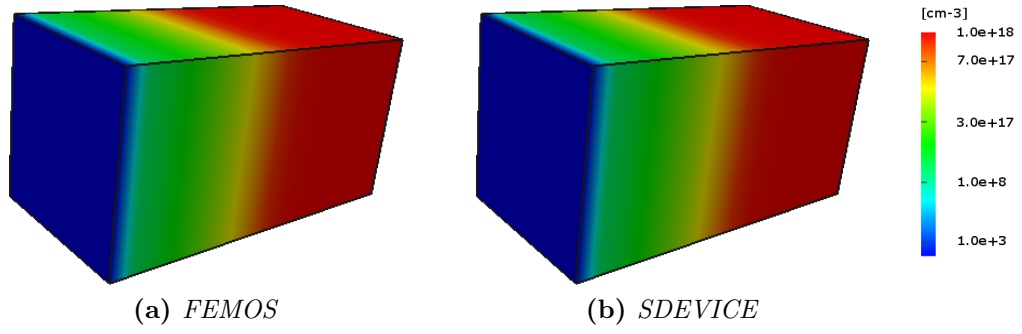




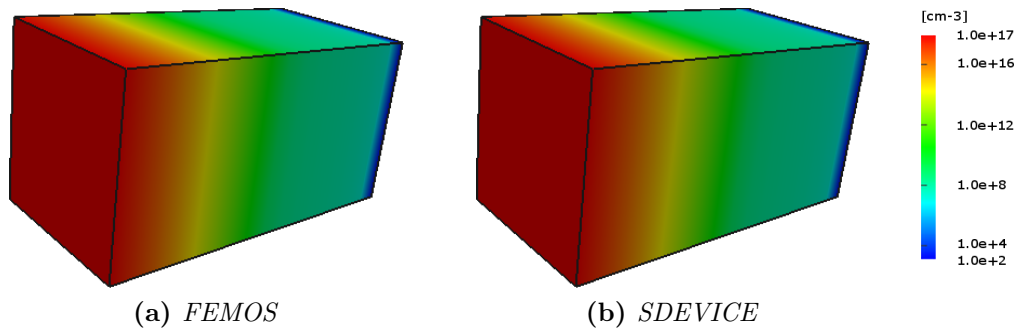
**Figure 4.2:** 1D plots of the solutions and the quasi fermi potential levels along the line parallel to the Z-axis and placed at the center of the device. On the left is presented the test case at  $V_A = 0.3[V]$  while on the right at  $V_A = 1.0[V]$ .



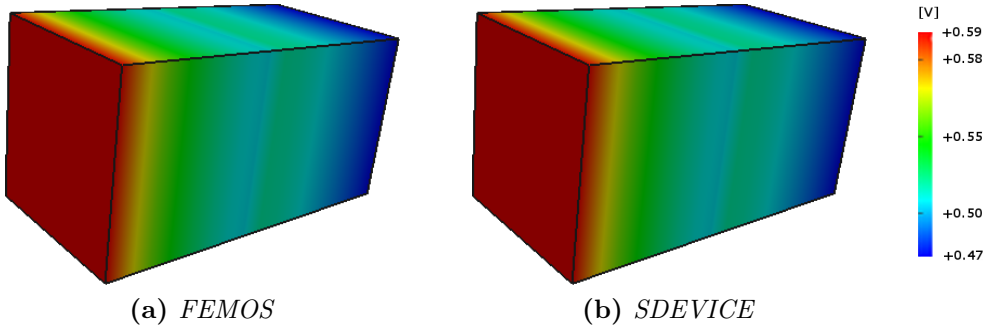
**Figure 4.3:** p-n junction 0.3[V] - Electrostatic Potential.



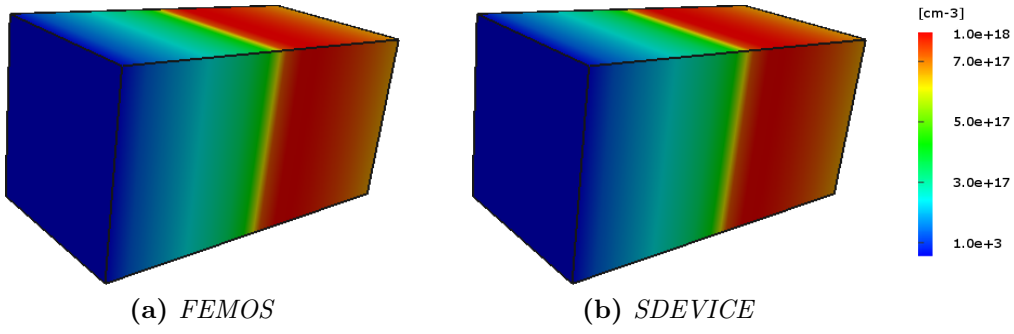
**Figure 4.4:** p-n junction 0.3[V] - Electron density.



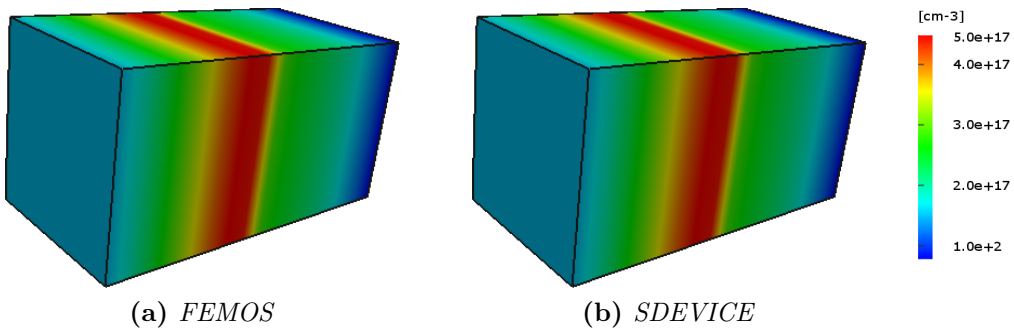
**Figure 4.5:** p-n junction 0.3[V] - Hole density.



**Figure 4.6:** p-n junction 1.0[V] - Electrostatic Potential.



**Figure 4.7:** p-n junction 1.0[V] - Electron density.



**Figure 4.8:** p-n junction 1.0[V] - Hole density.

### Computational cost and initial condition

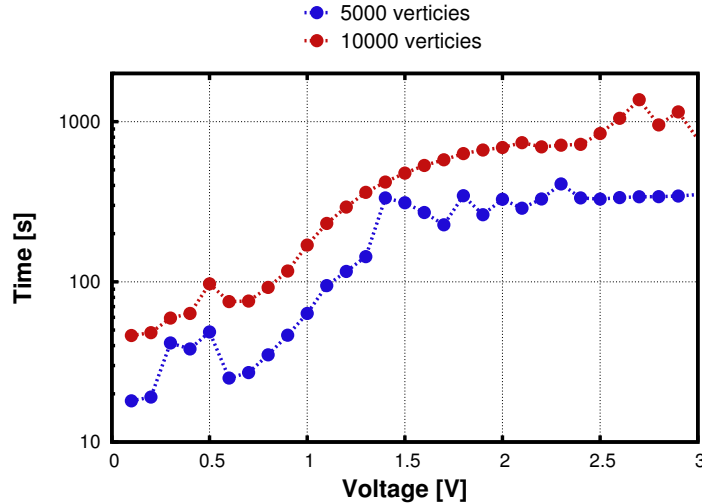
It's well known that the convergence time is strictly related to the kind of chosen initial condition: closest to solution better it is, however predict in every situations the possibly shape of the solutions is hard (if not even impossible). For this reason we have adopted a common and general approach splitting the domain in several regions accordingly to their doping concentration: each of the semiconductor regions are treated as they are in equilibrium with the nearest contact, then the guess value for  $\varphi$  is obtained thanks to the relations (1.20) or (1.21); this choice can be understood for a case close to the equilibrium: this method guarantees good converging performances.

In order to analyze the response of the system at different bias an additional test is realized: in the range between 0.0[V] and 3.0[V] several voltages are applied on the previous device and for each bias point the initial guess is computed as described.

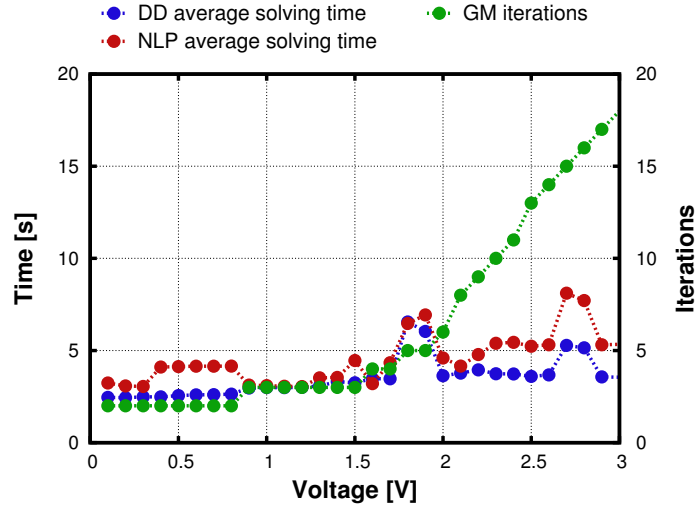
Fig.4.9 shows how the computational cost increases with the increasing of the applied bias. Moreover as expected if the mesh is finer, the time needed to find the solution increases, resulting in a rigid shift of the curve toward high time value.

Let us consider the case with coarse mesh. In Fig.4.10 it's evident how the average time spent to solve the NLP and the DD equations remains almost unchanged. On the contrary the number of GM iterations needed by the system to reach the solution, increases for voltages above  $\approx 1.5[V]$ .

A possible explanation of this could be found comparing solution and

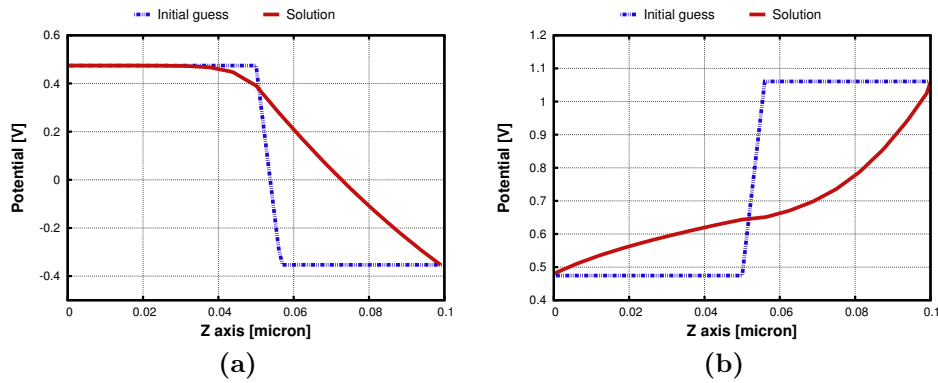


**Figure 4.9:** Total time Gummel Map.



**Figure 4.10:** Time NLP and DD, iteration GM.

initial guess for a bias below and above 1.5[V] (similar considerations may be done for carrier densities). When voltage is low, like in Fig.4.11a ( $V_A = 0.1[V]$ ), the potential shape is well predicted by the initial guess, resulting in a better convergence for the Gummel map algorithm. On the contrary in Fig.4.11b ( $V_A = 1.6[V]$ ) the device operates as a resistance and the potential profile is close to a linear function: this implies that the solution is far from the initial guess as equilibrium condition and the algorithm needs more steps.

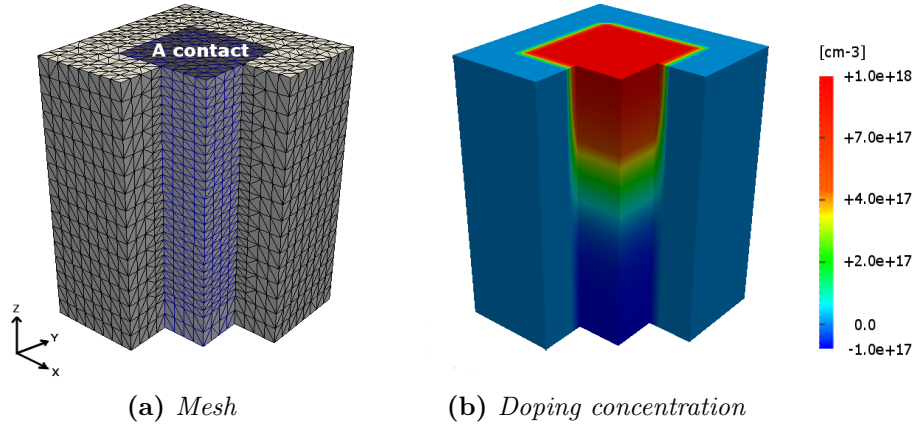


**Figure 4.11:** Initial step for different bias coupled with the solution.

### 4.1.2 p-n junction in oxide

In this test case a silicon p-n junction of  $0.3[\mu m]$  long has been surrounded by an oxide layer of  $0.025[\mu m]$  thick. The section of the silicon part is a  $0.1 \times 0.1[\mu m^2]$  square. We spent 6334 vertices and 33121 elements overall the domain. The structure and the doping are shown in Fig.4.12. The setting of the electrodes is similar to the previously test case and contacts are defined only on silicon surface. Tab.4.2 reports settings, models and parameters used in the simulations.

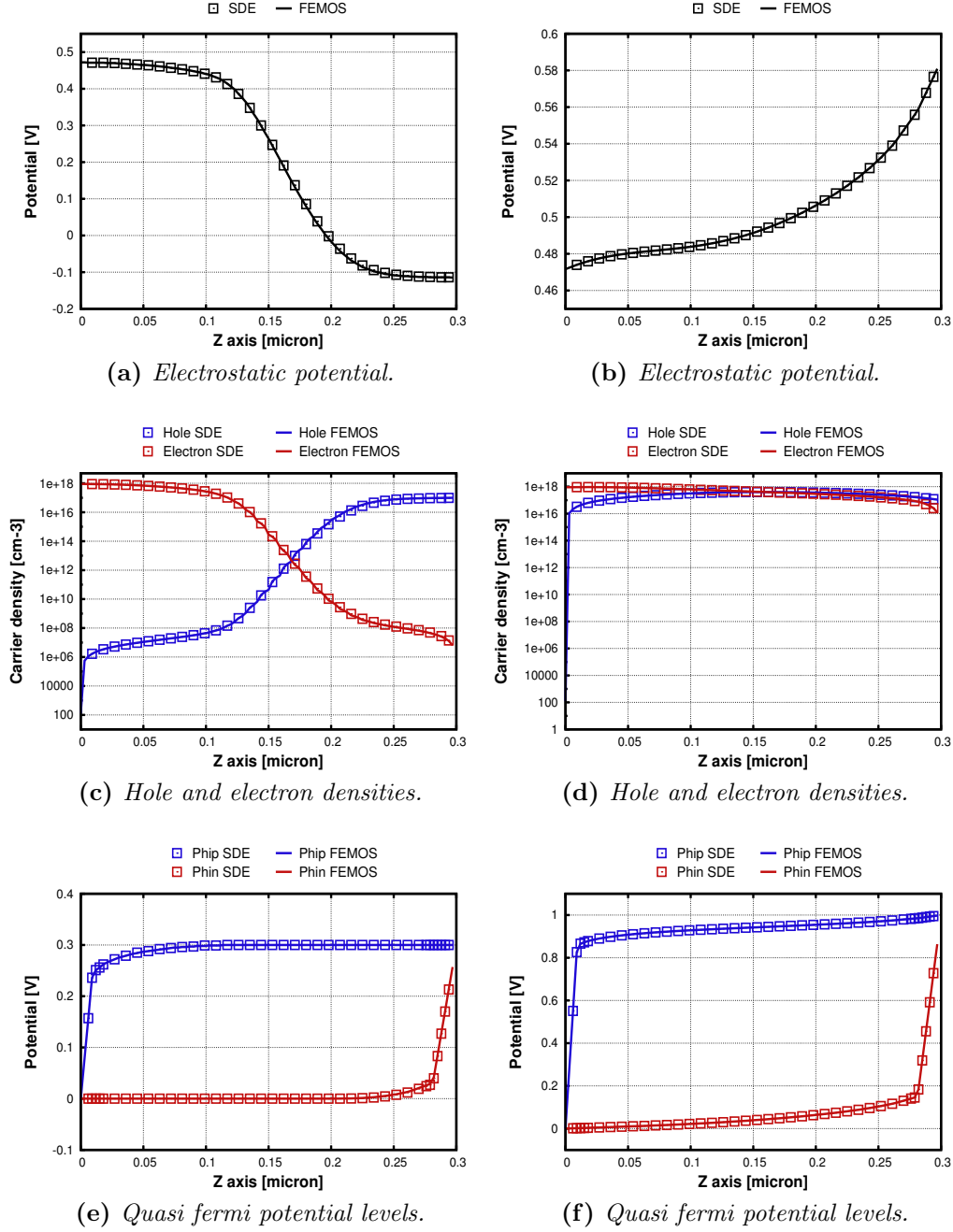
Fig.4.13 shows the solutions and the quasi fermi potential levels along a line parallel to the Z-axis and placed at the center of the device. The main features are similar to the previous test case, also for the boundary layers at contact in carriers and quasi fermi potential. Fig.4.14 ÷ 4.16 show the 3D solutions for the test at  $0.3[V]$ , while Fig.4.17÷4.19 refer to the case at  $1.0[V]$ . Both the 1D cuts and 3D plots agreement with the commercial software is very good.



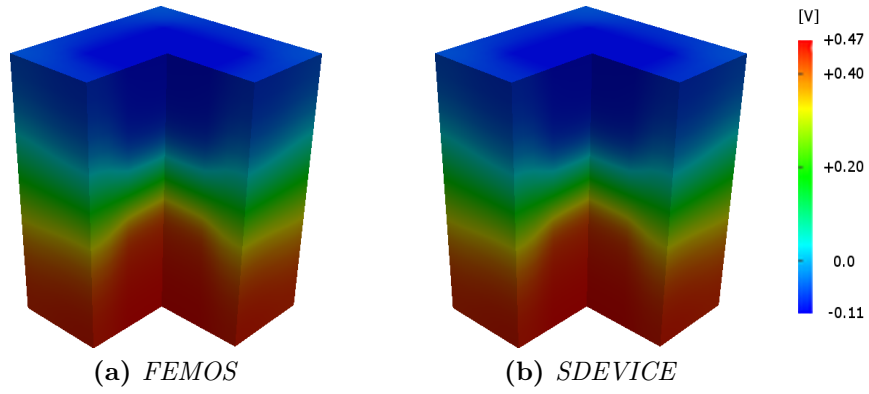
**Figure 4.12:** Test case p-n junction in oxide.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{Ox}$
$V_A = 0.3$	$\mu_n = 1417, \mu_p = 470.5$	SRH, Auger	11.6	3.9
$V_A = 1.0$	$\mu_n = 1417, \mu_p = 470.5$	SRH, Auger	11.6	3.9

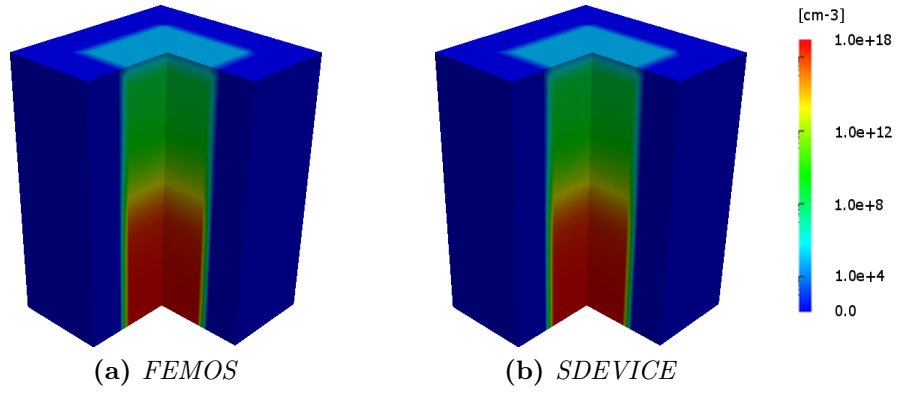
**Table 4.2:** p-n junction in oxide - list of settings, parameters and models.



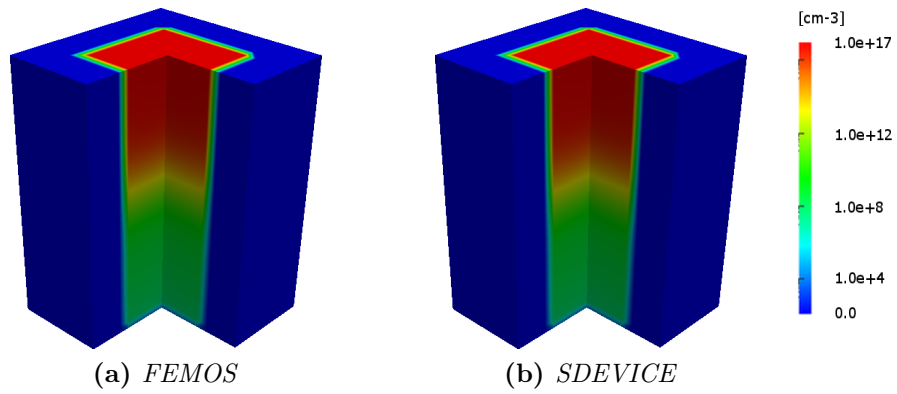
**Figure 4.13:** 1D plots of the solutions and the quasi fermi potential levels along the line parallel to the Z-axis and placed at the center of the device. On the left test case at  $V_A = 0.3[V]$  is reported while on the right at  $V_A = 1.0[V]$ .



**Figure 4.14:** p-n junction in oxide 0.3[V] - ElectrostaticPotential.

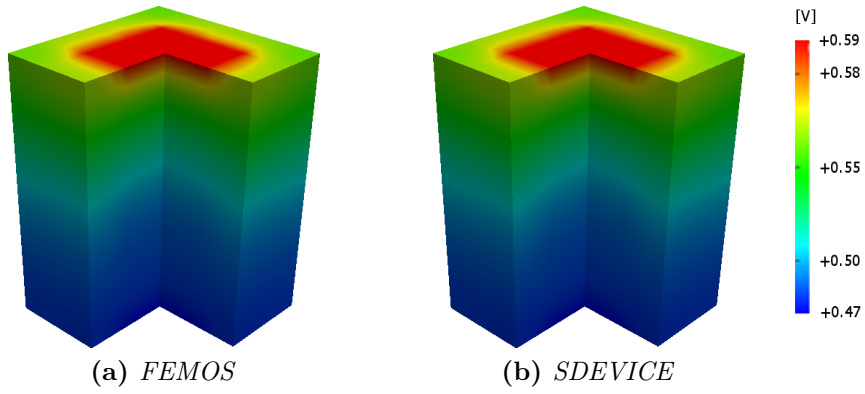


**Figure 4.15:** p-n junction in oxide 0.3[V] - Electron density.

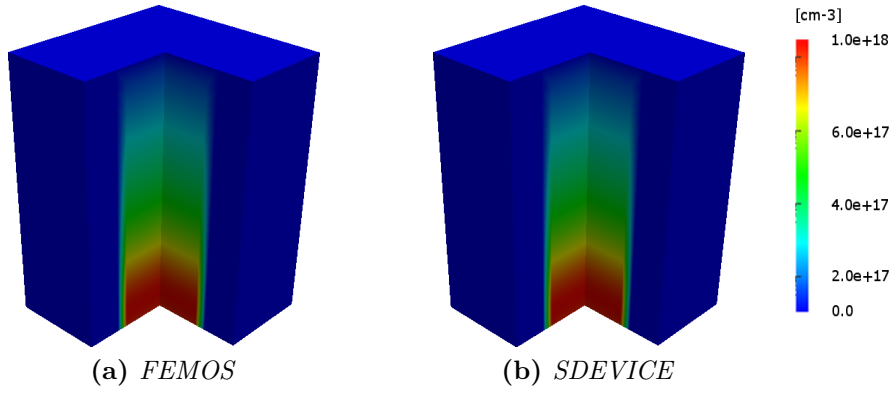


**Figure 4.16:** p-n junction in oxide 0.3[V] - Hole density.

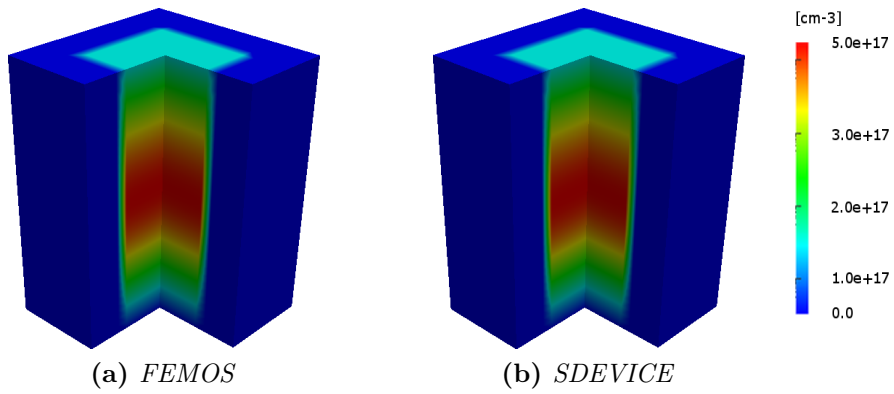




**Figure 4.17:** p-n junction in oxide 1.0[V] - Electrostatic Potential.



**Figure 4.18:** p-n junction in oxide 1.0[V] - Electron density.

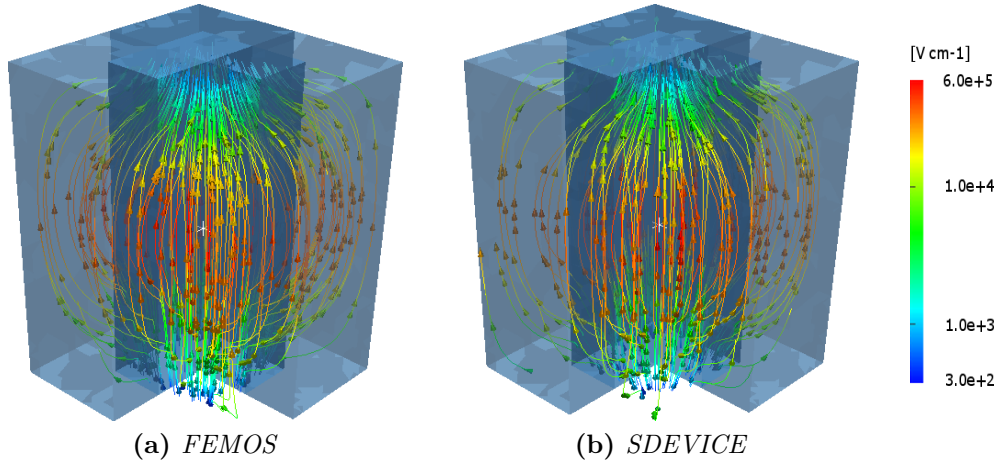


**Figure 4.19:** p-n junction in oxide 1.0[V] - Hole density.

### 3D effect on the electric field

Fig.4.14 shows how at low bias the electrostatic potential behaves different in the oxide with the respect to silicon. Because we imposed  $\nabla\varphi \cdot \mathbf{n} = 0$  on the oxide boundary no field lines of the electric field could cross that boundary and as consequence field lines could start and end only at contact A or B. The electric field inside the device is due only to the displacement effect in the junction of the silicon which imposes the electric response also in the oxide material where the solution is not anymore linear. Fig.4.20 reports the field lines of the electric field in the case  $V_A = 0.3[V]$  compared with the commercial software. The lower magnitude of the electric field in the oxide, results a more diffused potential.

At high bias (Fig.4.17) the influence of the contacts (A,B) becomes higher and the electrostatic potential is much more similar in the two subdomain.



**Figure 4.20:** Test case dide p-n in oxide 0.3[V] - Electric field.

It is important to note that displacement formulation approach doesn't satisfied in a strong manner the priciple of action and reaction. This is equivalent to say that given two element  $K_i, K_j \in \mathcal{T}_h$  such that  $K_i \cap K_j = f_i$  where  $f_i \in \mathcal{F}_h$  and let be  $\mathbf{n}_i$  the outward normal vector of  $\partial K_i$  and  $\mathbf{n}_j$  the outward normal vector of  $\partial K_j$ , the following equations are satisfied only in a weak way:

$$\mathbf{D}|_{K_i} \cdot \mathbf{n}_i = \mathbf{D}|_{K_j} \cdot \mathbf{n}_j \quad (4.1)$$

$$\mathbf{J}_n|_{K_i} \cdot \mathbf{n}_i = \mathbf{J}_n|_{K_j} \cdot \mathbf{n}_j \quad (4.2)$$

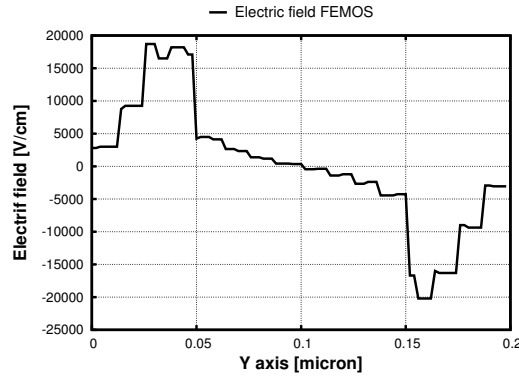
$$\mathbf{J}_p|_{K_i} \cdot \mathbf{n}_i = \mathbf{J}_p|_{K_j} \cdot \mathbf{n}_j \quad (4.3)$$

In fact taking in consideration equation (4.1), we can state that along the interface silicon-oxide the following equality holds

$$\epsilon_{ox} \mathbf{E}|_{K_i} \cdot \mathbf{n}_i = \epsilon_{Si} \mathbf{E}|_{K_j} \cdot \mathbf{n}_j \quad \Rightarrow \quad [\mathbf{E}|_{K_i}]_y = \frac{\epsilon_{Si}}{\epsilon_{ox}} [\mathbf{E}|_{K_j}]_y \quad (4.4)$$

Accordingly with the parameters used in the simulations the jump of the electric field component from oxide to semiconductor is around 3.00. Fig.4.21 shows the plots of  $E_y$  along a line parallel to the Y-axis crossing both oxide and silicon: the ratio expressed by (4.4) is almost 3.6 at  $y = 0.05[\mu m]$  and  $y = 0.15[\mu m]$  where the interfaces are located.

Furthermore each tetrahedral interfaces of the partition are affected by this problem, which means that the normal component of the electric field from one element of the grid to the neighbouring is not conserved even if the material is homogeneous and there is no charge at the interface. Even the presence of this drawback, the solutions are acceptable. But if we would satisfy equation (4.1) in a strong manner, the use of mixed-hybrid formulation is needed which ensures the conservation of the flux also under possible strong discontinuities of material properties.



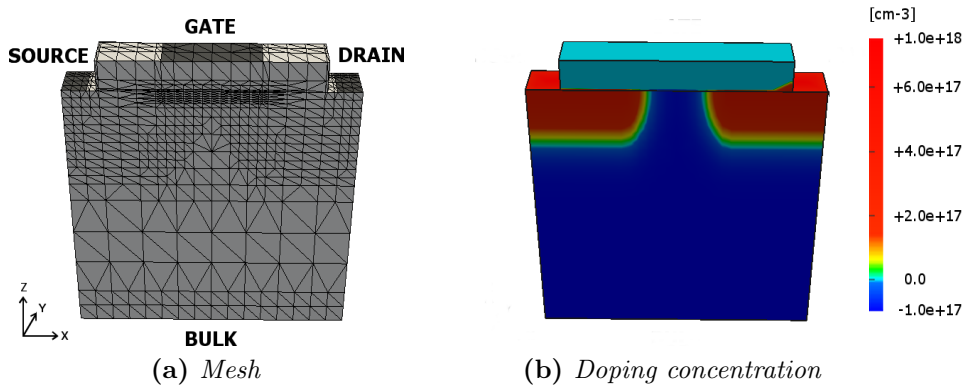
**Figure 4.21:**  $E_y$  along a line parallel to Y-axis,  $z = 0.22[\mu m]$  and  $x = 0.1[\mu m]$ .

### 4.1.3 MOSFET n-channel

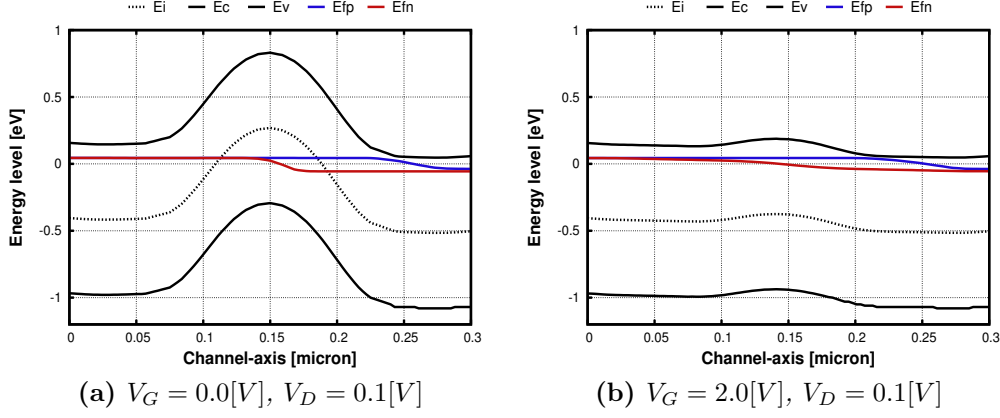
The description of MOSFET device can be found in [YT09]: it is a four-terminal device with the electrodes designated as gate (G), source (S), drain (D) and substrate or bulk (B). The gate electrode is usually made of metal or heavily doped polysilicon and is separated from the substrate by a thin silicon dioxide. The surface region under the gate oxide between source and drain is called *channel* region. Because the current in a MOSFET is due by carriers of one polarity, the MOSFET is usually referred as a unipolar or majority-carrier device: n-channel (n-MOSFET) and p-channel (p-MOSFET) are considered in the test. A *n-MOSFET* (p-MOSFET) consists of a p-type (n-type) silicon substrate into which two n-regions (p-regions) are designed as source and drain. The n-regions (p-regions) are doped accordingly with a gauss profile as occurring by a real implantation process. Fig.4.22a and Fig.4.22b shows the geometry and the doping concentration for the n-MOSFET with a coarse mesh (2739 vertices and 12338 elements). We note that the mesh has been refined where the more interesting phenomena occur: along channel and at drain/source contacts.

If sufficiently large positive voltage is applied to the gate, the silicon surface is inverted to n-type (p-type), which forms a conducting channel between the source and drain: applying a small positive voltage to the drain (source) the electrons (holes) start to flow from source (drain) to drain (source) and therefore a current is generated.

The visualization of the MOSFET working principle has been better clarified in Fig.4.23 where the band profile along the channel axis has been reported for two different gate bias. The voltage applied to the gate tends to decrease the energy barrier in the channel region: a little drain voltage causes the flow of the electrons. Fig.4.24a shows the profile of carrier concentrations



**Figure 4.22:** Geometry of MOS n-channel.



**Figure 4.23:** Energy band levels for nMOSFET along channel.

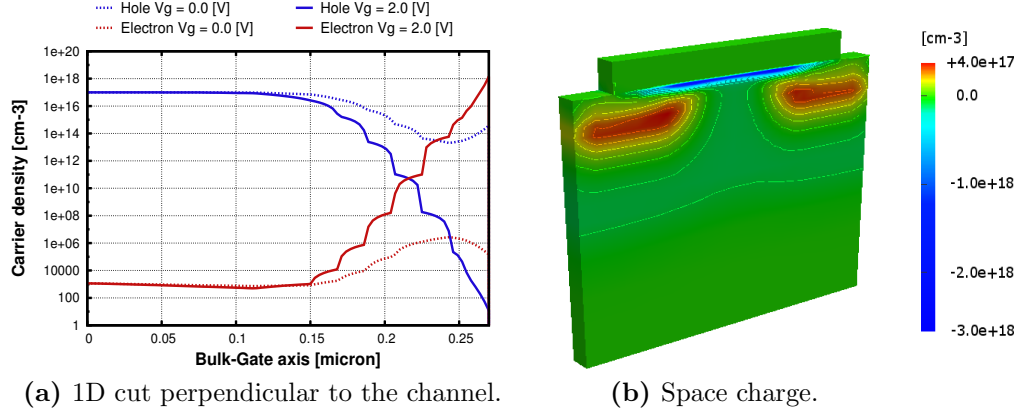
in the middle of the channel with a cut perpendicular to the gate, for off-state and on-state: the inversion occurs when the gate voltage is higher than the well known threshold voltage. Fig.4.24b shows the 3D view of the n-channel space charge after inversion.

Finally Fig.4.25 reports the streamline plot of the electric field inside the device for FEMOS and the SDEVICE in the case of on-state MOSFET.

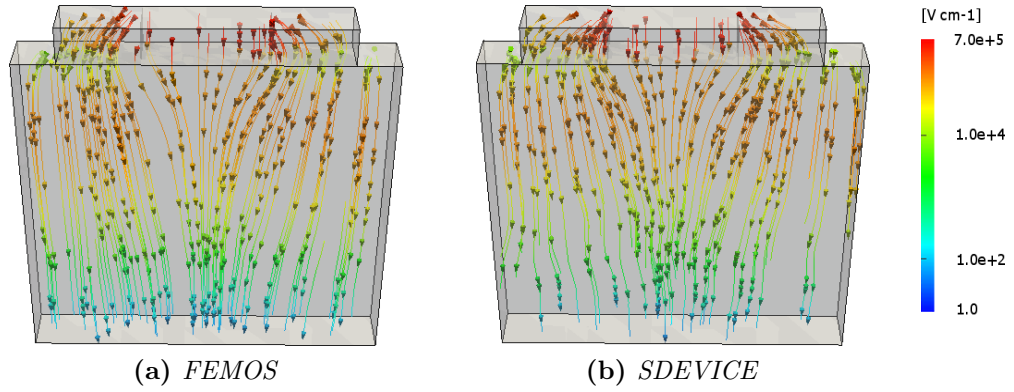
Settings, parameters and models used for these simulations are summarized in Tab.4.3. Fig.4.26÷4.28 shows the 3D view of the electrostatic potential, electron and hole densities obtained by FEMOS and by the commercial code for the off-state ( $V_G = 0.0[V]$ ), while Fig.4.29÷4.31 for the on-state ( $V_G = 2.0[V]$ ): the agreement is very good.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_G = 0.0[V]$ , $V_D = 0.0[V]$ , $V_S = V_B = 0.0[V]$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH, Auger, II	11.6	3.9
$V_G = 2.0[V]$ , $V_D = 0.1[V]$ , $V_S = V_B = 0.0[V]$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH, Auger	11.6	3.9

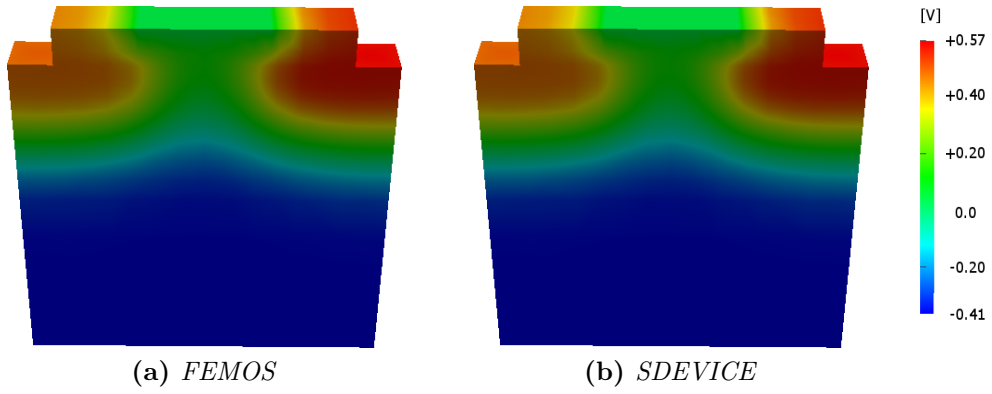
**Table 4.3:** nMOSFET - list of settings, parameters and models.



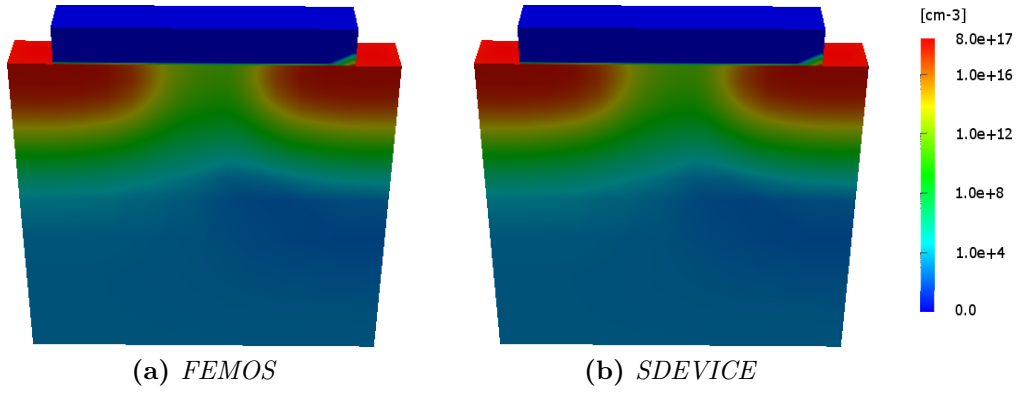
**Figure 4.24:** Channel of the nMOSFET.



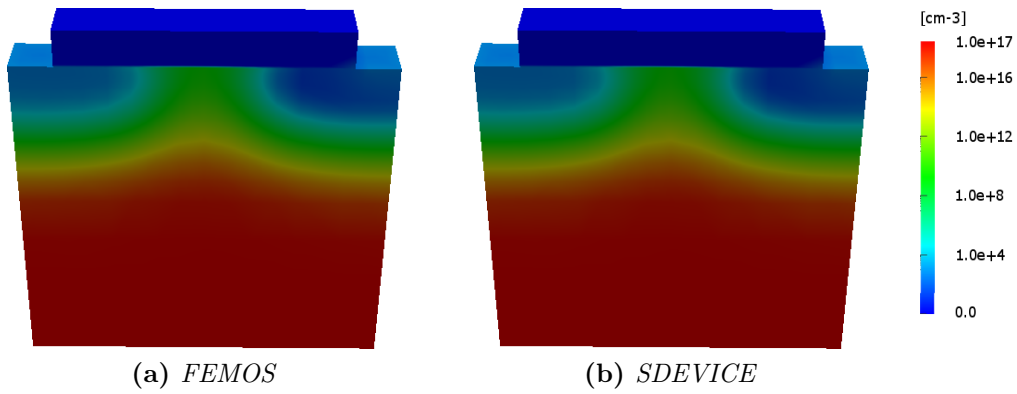
**Figure 4.25:** Electric field density -  $V_G = 2.0[V]$ .



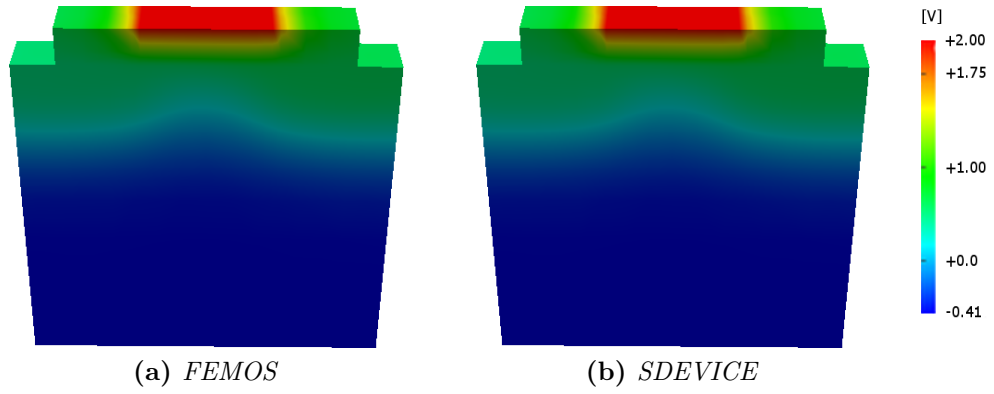
**Figure 4.26:** Electrostatic potential -  $V_G = 0.0[V]$ .



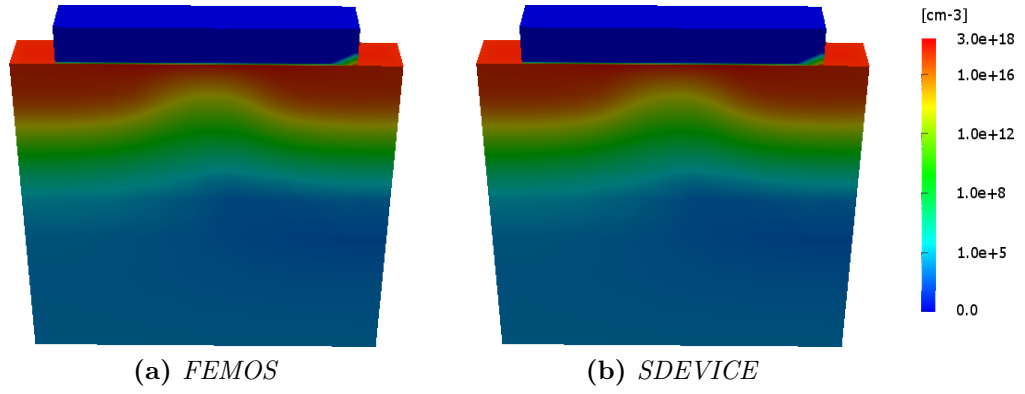
**Figure 4.27:** Electron density -  $V_G = 0.0[V]$ .



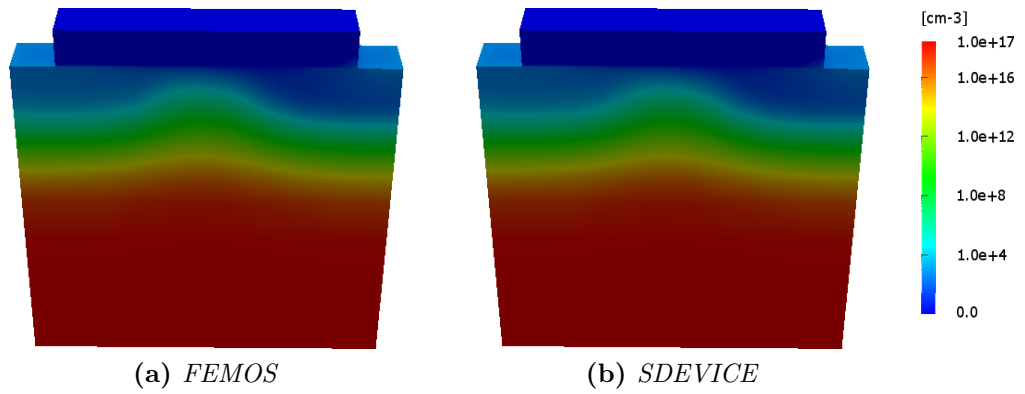
**Figure 4.28:** Hole density -  $V_G = 0.0[V]$ .



**Figure 4.29:** Electrostatic potential -  $V_G = 2.0[V]$ .



**Figure 4.30:** Electron density -  $V_G = 2.0[V]$ .



**Figure 4.31:** Hole density -  $V_G = 2.0[V]$ .



### Inverse polarization

In section 3.3.3 we pointed that the discretization scheme (EAFE) can't satisfy the discrete maximum principle in 3D simulations unless we satisfy condition (3.48). Therefore it is possible encounter negative solution and we usually have to deal with this problem when the concentration of electrons and holes become low.

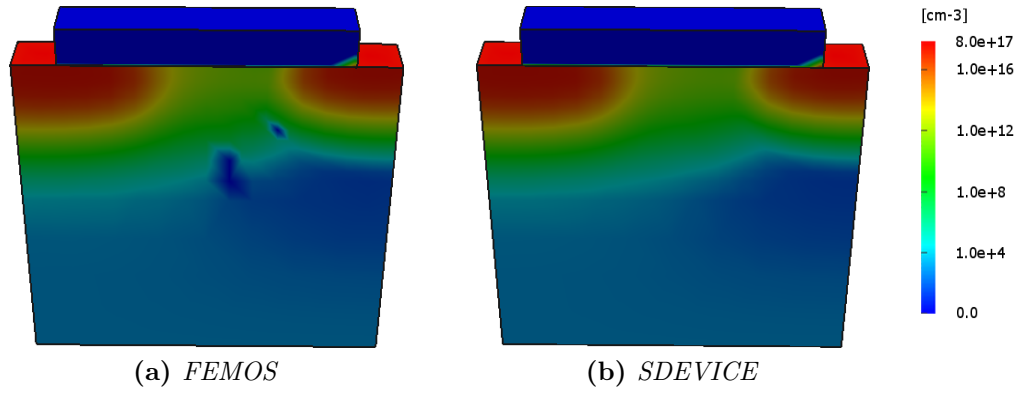
In order to highlight this possible critical situation n-channel MOSFET has put in inverse polarization by grounded all contacts except the drain which is ramped to 0.5[V]: Tab.4.4 reports the needed indications for the simulation.

Fig.4.32 reports the electron density computed with FEMOS and SDEVICE using the mesh presented in Fig.4.22b: the results are comparable, but near the drain-bulk junction FEMOS presents some points with negative concentrations. Increasing drain bias the phenomenon will tend to spread over a larger area, untill it affects irremediably the computation of the results. As we anticipated in chapter 3, the most practice technique to limitate this problem is the addition of degrees of freedom around the problematic regions. Fig.4.33a represents a finer mesh with 13000 points and 67388 elements. Using this mesh the correctness of the solution is recovered Fig.4.33b. Fig.4.34 shows how the satisfaction of the (3.48) changes between the different meshes: increasing the number of vertices over the critical region a better fulfillment of (3.48) is guranteed. Results suggest that in order to treat this situation it may be useful implement a suitable a-posteriori error estimation and adaptive mesh refinement techniques.

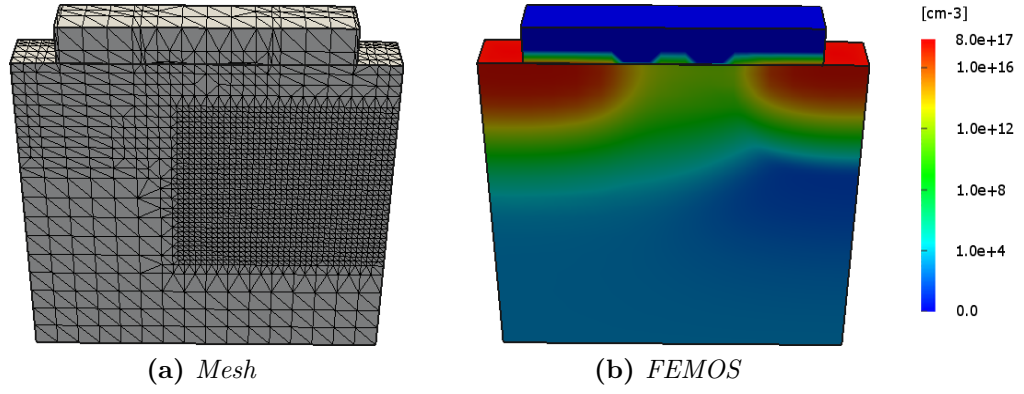
Finally Fig.4.35 and Fig.4.36 reports FEMOS electrostatic potential and hole density for the finer mesh compared with the commercial tool: the agreement is very good.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_G = 0.0, V_D = 0.5,$ $V_S = V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH, Auger, II	11.6	3.9

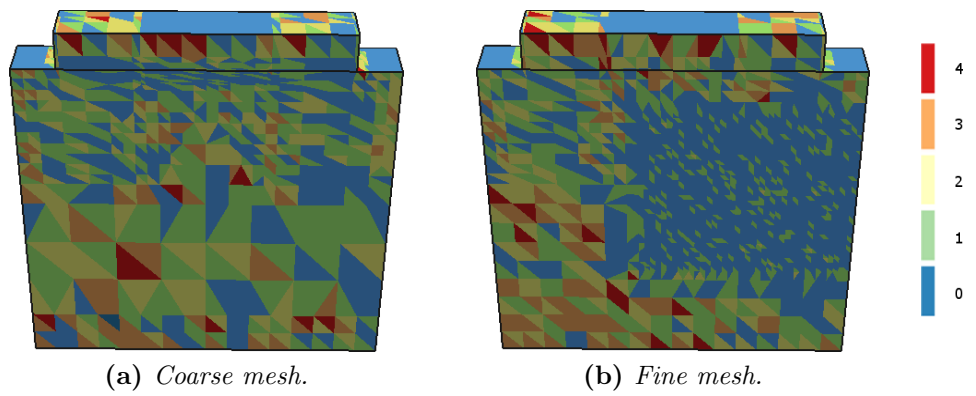
**Table 4.4:** nMOSFET (inverse) - list of settings, parameters and models.



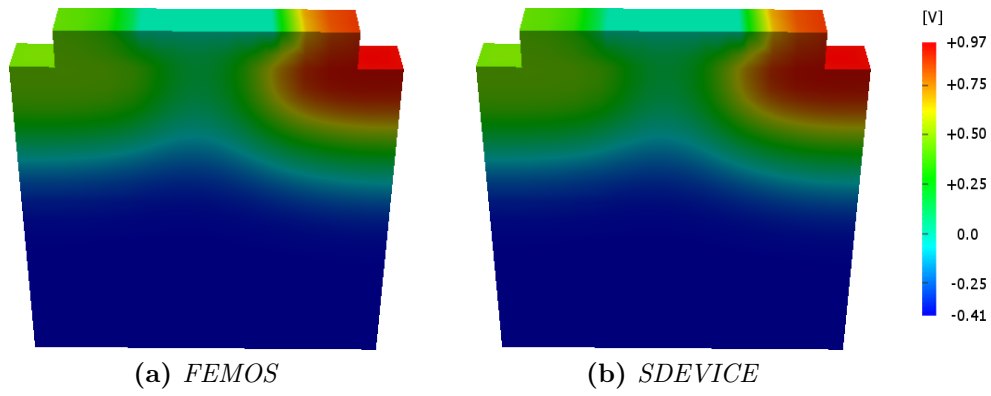
**Figure 4.32:** Negative carriers spots for the electron density solution.



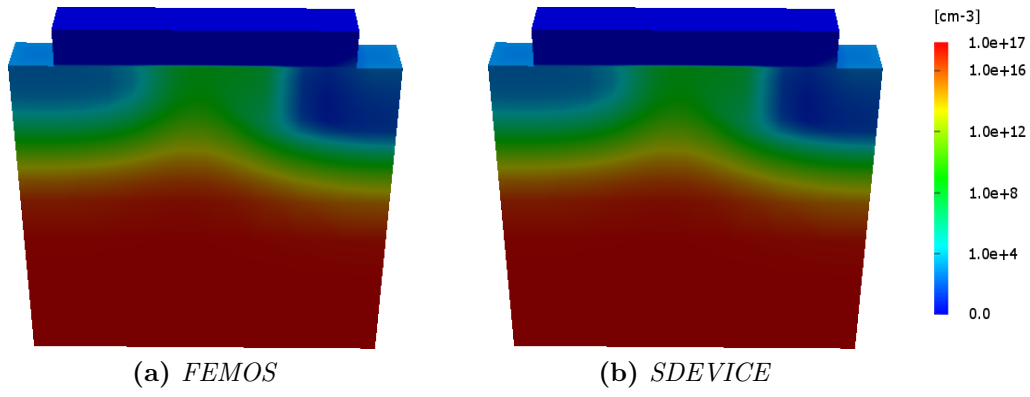
**Figure 4.33:** Electron density with finer mesh.



**Figure 4.34:** Zikatanov condition.



**Figure 4.35:** Electrostatic potential - Inverse polarization.



**Figure 4.36:** Hole density - Inverse polarization.

## 4.2 Calculation of the current at contacts

During the analysis of an electric device, one of the most useful information is the electrical response at terminals. In order to accomplish this target we have to compute the integral of the electron and hole current density over a generic 2D electrode. We refer here to the procedure found in [GS74] (*residue method*) for the 2D case: the analysis is easily extendable to the 3D case if we consider also [HEML00]. Moreover we remark that the method can be successfully applied to a wide spread of applications, including contact charges, carrier quantum probability fluxes and heat fluxes.

A contact is defined by a surface: we can consider  $\Gamma_{D,Si} = \bigcup_{c=1}^d \Gamma_c$  where  $d$  is the number of terminals on the device and  $\forall c = 1, \dots, d$ ,  $\Gamma_c$  is the  $c$ -th contact. For each contact we need to compute the total current  $I_c$  as:

$$\mathcal{I}_c = \mathcal{I}_c^n + \mathcal{I}_c^p \quad (4.5)$$

where  $\mathcal{I}_c^n$  and  $\mathcal{I}_c^p$  are the contribution of the electron and hole current. For a given contact  $\Gamma_c$ , the fluxes of the current density is the flow of:

$$\mathcal{I}_c^\nu = \int_{\Gamma_c} \mathbf{J}_\nu(\nu) \cdot \mathbf{n} d\Gamma \quad \nu = \{n, p\} \quad (4.6)$$

where  $\mathbf{n}$  is the unit outward normal of the domain boundary. It's well known that the evaluation of boundary integrals is a difficult task. Most of the problems (4.6) arise from singularities in spatial derivatives of the approximate solution  $n_h$  or  $p_h$  near the contact edges, due to a change in the boundary condition type from Dirichlet to Neumann.

Let  $\eta$  be the set of all vertices of the partition  $\mathcal{T}_h$  for the discretized electron continuity problem (3.34). We can split the set of total nodes in contact node  $\eta_g \in \Gamma_{D,Si}$  and the complementary part  $\eta_n \in \Gamma_{N,Si}$ . We define an *auxiliary flux*  $H_h$  on  $\Gamma_{D,Si}$  as

$$H_h = \sum_{i \in \eta_g} H_{h,i} \psi_i \quad (4.7)$$

Now given the spaces:

$$\begin{aligned} \mathcal{V}_h &= \text{span}\{\psi_i\}_{i \in \eta_n} \\ \mathbf{V}_h &= \text{span}\{\psi_i\}_{i \in \eta_g} \\ \mathcal{S}_h &= \{u \in \mathcal{V}_h \oplus \mathbf{V}_h : u|_{\Gamma_{D,Si}} = n_D\} \end{aligned}$$

it's possible write a modified form of Glakerin's method which reads as:

find  $n_h \in \mathcal{S}_h$  and  $H_h \in \mathbf{V}_h$  such that

$$(W_h, H_h)_{\Gamma_{D,Si}} = a(W_h, n_h) - F(W_h) \quad \forall W_h \in \mathcal{V}_h \oplus \mathbf{V}_h \quad (4.8)$$

where  $a(\cdot, \cdot)$  is the bilinear form (3.34) and  $F(\cdot)$  the relative functional. Equation (4.8) splits into two subproblems:

$$0 = a(w_h, n_h) - F(w_h) \quad \forall w_h \in \mathcal{V}_h \quad (4.9)$$

$$(W_h, H_h)_{\Gamma_{D,Si}} = a(W_h, n_h) - F(W_h) \quad \forall W_h \in \mathbf{V}_h \quad (4.10)$$

Problem (4.9) is identical to the unmodified case and can be treated as before or using a different discretization scheme. Once obtained the solution  $n_h$  problem (4.10) is fully decoupled from (4.9) and we can determine  $H_h$  as follows

$$(H_h, \psi_i)_{\Gamma_{D,Si}} = a(\psi_i, n_h) - F(\psi_i) \quad \forall i \in \eta_g \quad (4.11)$$

[HEML00] demonstrates that  $H_h$  defines the conserved total flux along  $\Gamma_{D,Si}$  and accordingly with the boundary condition, the following equality is obtained

$$\int_{\Gamma_{D,Si}} H_h d\Gamma = - \int_{\Omega_{Si}} qR d\Omega \quad (4.12)$$

On the other hand if we apply the Divergence theorem on (2.30) we can state

$$\int_{\Gamma_{D,Si}} \mathbf{J}_n \cdot \mathbf{n} d\Gamma = \int_{\Omega} -qR d\Omega \quad (4.13)$$

Equation (4.12) and (4.13) lead us to conclude that for all contacts holds

$$\mathcal{I}_c^n = \int_{\Gamma_c} H_h d\Gamma \quad (4.14)$$

In order to compute (4.14) let be  $\eta_c$  the set of nodes of the contact  $\Gamma_c$ , the following equalities hold

$$\sum_{l \in \eta_c} \int_{\Gamma_c} H_h \psi_l d\Gamma = \int_{\Gamma_c} H_h \sum_{l \in \eta_c} \psi_l d\Gamma = \int_{\Gamma_c} H_h d\Gamma \quad (4.15)$$

Accordingly with (4.15) we can reinterpret  $(H_h, \psi_i)_{\Gamma_{D,Si}}$  as the contribution to the flux at node  $i$  and therefore the current at contact  $c$  is given by summing this quantity over the vertices  $\eta_c$ .

The residue method is: given the system matrix  $A$  of the Drif-Diffusion equation, the solution  $n_h$  and the right hand side  $\mathbf{b}$ , the contribution to the total contact current  $\forall c = 1, \dots, d$  is

$$\mathcal{I}_c^n = (An_h - \mathbf{b}) \cdot \mathbb{I}_c \quad (4.16)$$

where

$$[\mathbb{I}_c]_i := \begin{cases} 0 & i \notin \eta_c \\ 1 & i \in \eta_c \end{cases} \quad (4.17)$$

These results hold also for the hole continuity equation.

### 4.2.1 Simulation results

The residue method is applied and compared for the devices previously presented with the SDEVICE results. In this section the different mobility and recombination/generation model has been tested and verified along with commercial code solution.

#### p-n junction

Considering p-n junction of section 4.1.1 we grounded the B contact and then A contact is ramped from  $-7.5[V]$  to  $2.0[V]$  in order to obtain the well known diode characteristic. Tab.4.5 reports the parameters used in simulation. In Fig.4.37 we plot the electron and hole current at contact A. Diode breakdown voltage is appearing around  $-7.0[V]$  and it's quite well aligned with the SDEVICE value.

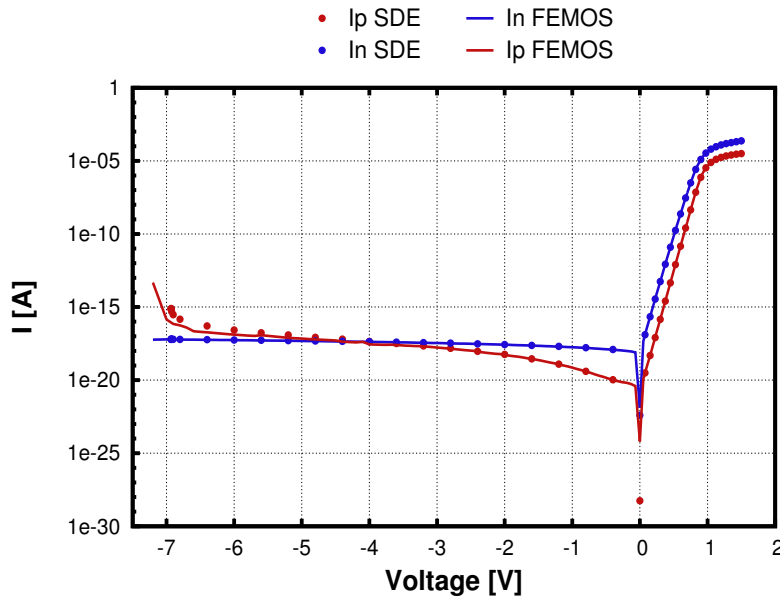


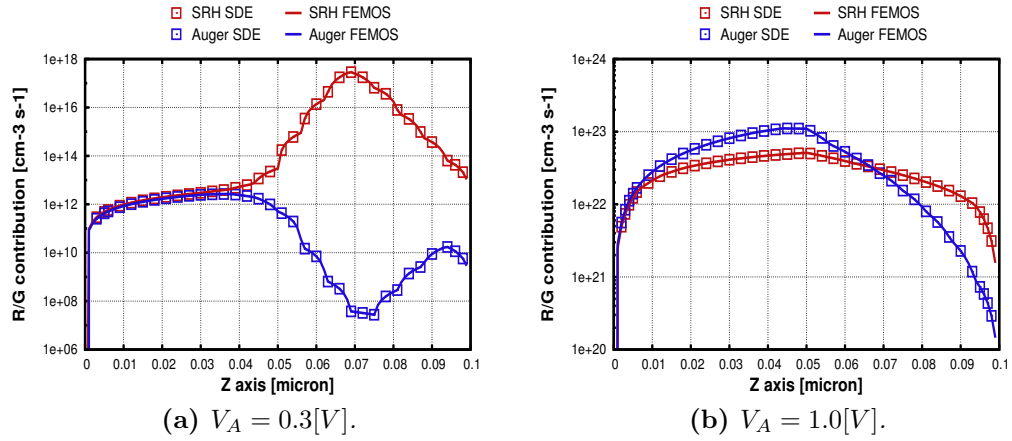
Figure 4.37: Diode characteristic.

Fig.4.38 shows the different behaviour of the SRH and Auger R/G contribution at two voltages (plots are made along a line parallel to the Z-axis and placed at the center of the device). Under the built-in both SRH and Auger are not significative: Auger contribution decreases in the depletion region due to its strictly dependence of the carrier concentrations. At the

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$
$V_A = -7.5 \div 1.5,$ $V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH II, Auger	11.6

**Table 4.5:** p-n junction (characteristic) - list of settings, parameters and models.

built-in: the R/G phenomena are quite increased and distributed over the entire device causing its saturation. The agreement with the commercial tool is very good.

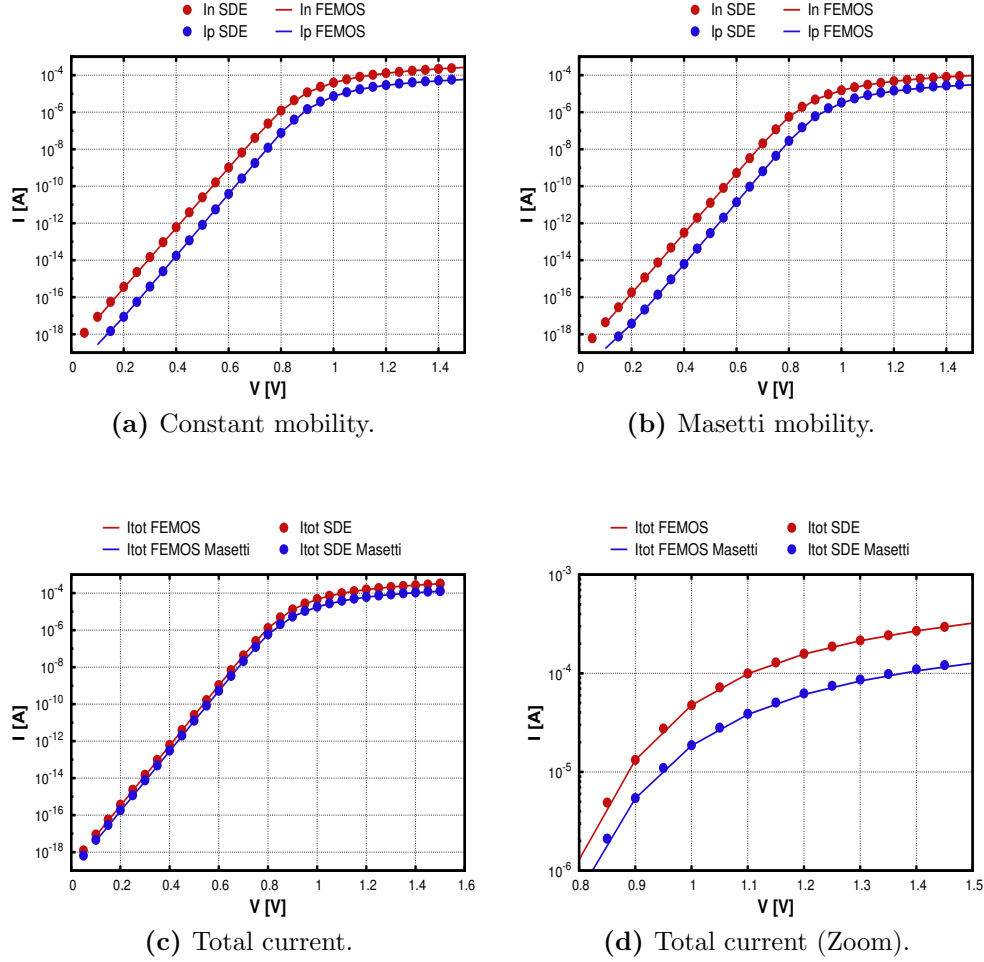


**Figure 4.38:** SRH and Auger RG contribution.

### p-n junction in oxide

In the case of diode in oxide we analyzed the influence of the Masetti model. Fig.4.39a reports electron and hole current at contact A using constant mobility model, while Fig.4.39b using the Masetti one. Applying Masetti a decreasing of the current flow due to the scattering with impurity dopants is found as shown in 1D plot proposed by Fig.4.39d. Agreement with sdevice is very good.



**Figure 4.39:** p-n junction in oxide current at contact - Direct polarization.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_A = 0.0 \div 1.5,$ $V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH II, Auger	11.6	3.9
$V_A = 0.0 \div 1.5,$ $V_B = 0.0$	Masetti	SRH II, Auger	11.6	3.9

**Table 4.6:** p-n junction in oxide - list of settings, parameters and models.

### MOSFET n-channel/p-channel

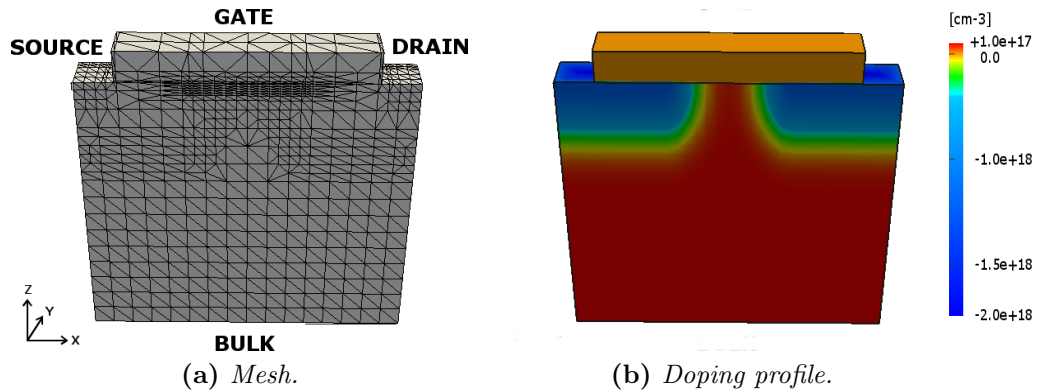
Results for the nMOSFET and the pMOSFET are shown in this section. For n-channel we refer to Fig.4.22a while for p-channel Fig.4.40 shows mesh (2618 vertices and 11514 elements) and doping (a gaussian profile in source and drain regions). In order to validate the code we tested the following device conditions:

1.  $I_D - V_G$  characteristic at low drain bias with several mobility models;
2.  $I_D - V_G$  characteristic for different drain bias;
3.  $I_D - V_D$  characteristic in off-state (inverse polarization).

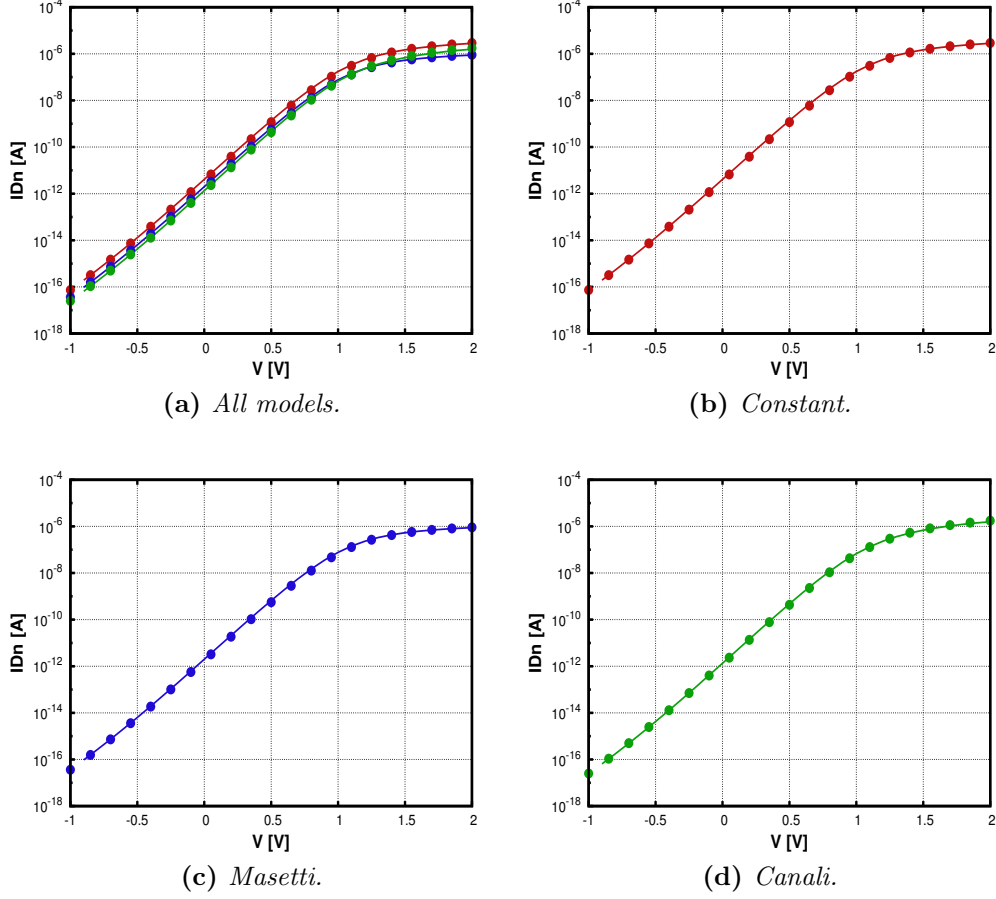
Tab.4.7 and Tab.4.8 show simulation settings of the first test case for the nMOSFET and the pMOSFET respectively. Fig.4.41 shows the results for the nMOSFET: Fig.4.41a shows how the current densities change as the results of different mobility models. Fig.4.41b÷4.41d confirm the good agreement with the commercial tool.

Similar test is performed for the pMOSFET, but considering that a p-channel usually operates for negative values of the gate bias, we have to apply a positive polarization on the source terminal. The results are shown in Fig.4.42: the agreement with SDEVICE is still good.

Fig.4.43 presents the characteristic of the nMOSFET at different values of the drain voltage. The analogous test for the pMOSFET is presented in Fig.4.44. Both are very well in agreement with the commercial code.

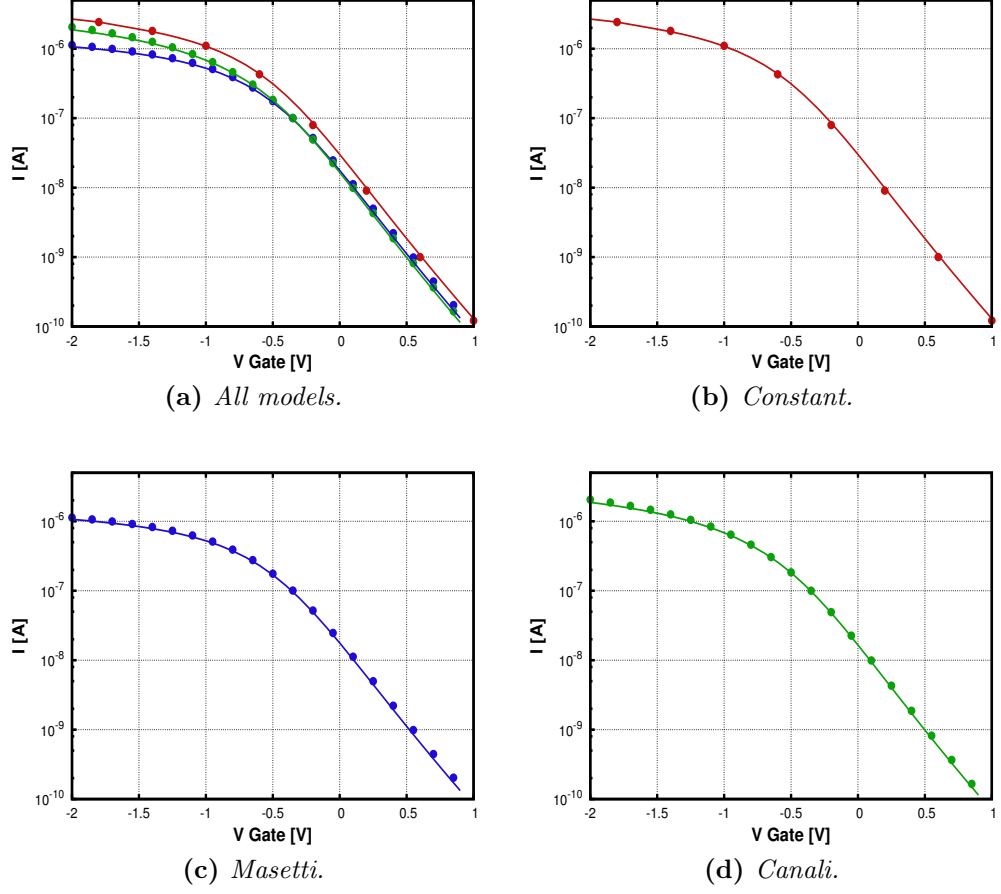


**Figure 4.40:** pMOSFET.

**Figure 4.41:**  $I_D - V_G$  nMOSFET characteristic - mobility models.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_G = -0.5 \div 2.0,$ $V_D = 0.1, V_S = V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH, Auger	11.6	3.9
$V_G = -0.5 \div 2.0,$ $V_D = 0.1, V_S = V_B = 0.0$	Masetti	SRH, Auger	11.6	3.9
$V_G = -0.5 \div 2.0,$ $V_D = 0.1, V_S = V_B = 0.0$	Canali	SRH, Auger	11.6	3.9

**Table 4.7:** nMOSFET (low drain bias characteristic) - list of settings, parameters and models.



**Figure 4.42:**  $I_S - V_G$  pMOSFET characteristic - mobility models.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_G = -1.5 \div 0.5,$ $V_S = 0.1, V_D = V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH, Auger	11.6	3.9
$V_G = -1.5 \div 0.5,$ $V_S = 0.1, V_D = V_B = 0.0$	Masetti	SRH, Auger	11.6	3.9
$V_G = -1.5 \div 0.5,$ $V_S = 0.1, V_D = V_B = 0.0$	Canali	SRH, Auger	11.6	3.9

**Table 4.8:** pMOSFET (low drain bias characteristic) - list of settings, parameters and models.

All values and models used for these simulations are summerized in 4.9 for the nMOSFET and 4.10 for the pMOSFET.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_G = -0.5 \div 2.0$ , $V_D = 0.1, V_S = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -0.5 \div 2.0$ , $V_D = 0.2, V_S = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -0.5 \div 2.0$ , $V_D = 0.5, V_S = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -0.5 \div 2.0$ , $V_D = 1.0, V_S = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -0.5 \div 2.0$ , $V_D = 2.0, V_S = V_B = 0.0$	Canali	SRH, II	11.6	3.9

**Table 4.9:** nMOSFET (different drain bias) - list of settings, parameters and models.

Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_G = -1.5 \div 0.5$ , $V_S = 0.05, V_D = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -1.5 \div 0.5$ , $V_S = 0.1, V_D = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -1.5 \div 0.5$ , $V_D = 0.2, V_D = V_B = 0.0$	Canali	SRH, II	11.6	3.9
$V_G = -1.5 \div 0.5$ , $V_S = 0.5, V_D = V_B = 0.0$	Canali	SRH, II	11.6	3.9

**Table 4.10:** pMOSFET (different drain bias) - list of settings, parameters and models.

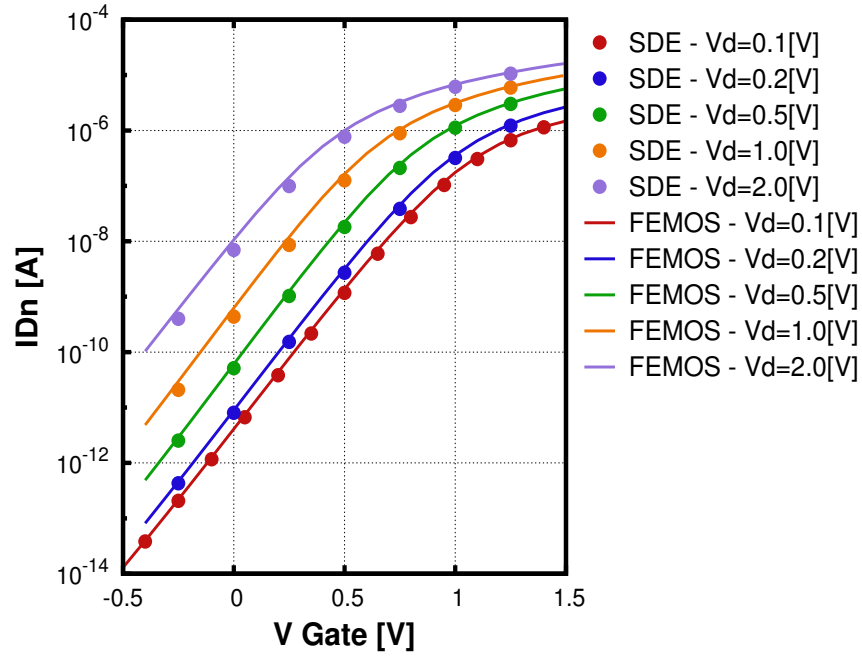


Figure 4.43:  $I_D - V_G$  nMOSFET - several drain voltages.

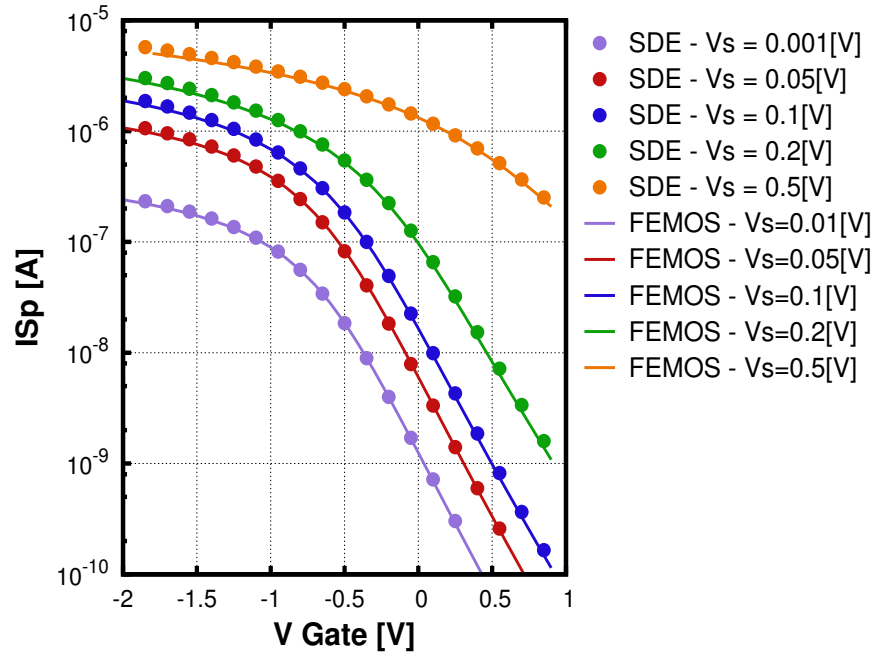


Figure 4.44:  $I_S - V_G$  pMOSFET - several source voltages.

### Impact-ionization model

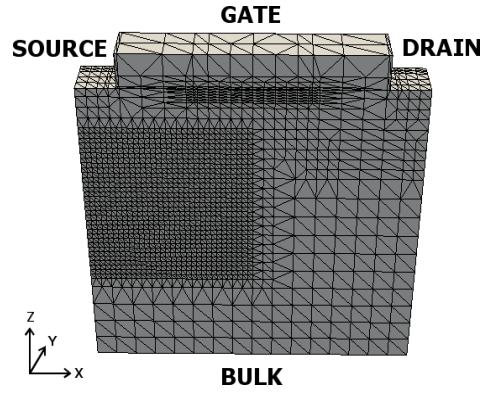
To visualize the effects of impact ionization (Van Overstraeten - de Man model see section 1.2.2) we investigated the devices in off-state with increasing drain (nMOSFET) or source voltage (pMOSFET).

The parameters are summerized in Tab.4.11 for the nMOSFET and in Tab.4.12 for the pMOSFET.

In order to avoid possibly negative concentration also for p-channel, we add degrees of freedom at the source-bulk junction region as presented in Fig.4.45 (15504 vertices and 81587 elements).

For nMOSFET, increasing drain voltage a large amount of generation is produced around the drain-bulk junction as shown in Fig.4.46, where we compare the contribution due to II model computed at  $V_D = 0.5[V]$  between FEMOS and SDEVICE. Some differences arise but the discrepancy doesn't affect the computation of the current at contacts. Fig.4.49 and (4.50) confirm the very good agreement with the commercial tool also for the bulk current. In off-state condition no channel is formed beneath the oxide layer, therefore no preferential path is allowed for electrons (or holes) conduction: carriers can move toward source (drain) and bulk contact.

Due to the difference in the doping level between nMOSFET and pMOSFET the generation of carriers for II is different as shown in Fig.4.46 (nMOSFET) compared with Fig.4.47 (pMOSFET) for  $V_S = V_D = 0.5[V]$  and Fig.4.48 also for  $V_S = 1.2[V]$ .

**Figure 4.45:** pMOSFET mesh.

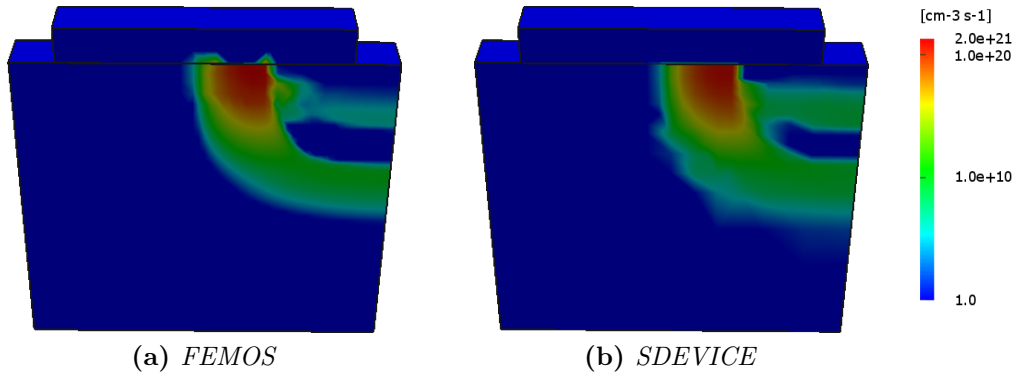
Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_D = 0.0 \div 1.0$ , $V_G = V_S = V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH Auger, II	11.6	3.9

**Table 4.11:** nMOSFET (off-state characteristic) - list of settings, parameters and models.

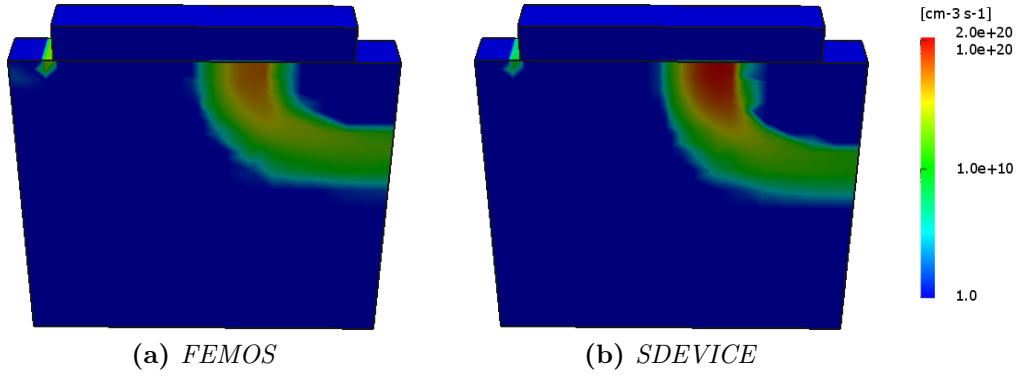
Test case [V]	Mobility model [ $cm^2V^{-1}s^{-1}$ ]	R/G model	$\epsilon_{Si}$	$\epsilon_{0x}$
$V_S = 0.0 \div 1.5$ , $V_G = V_D = V_B = 0.0$	$\mu_n = 1417$ $\mu_p = 470.5$	SRH Auger, II	11.6	3.9

**Table 4.12:** pMOSFET (off-state characteristic) - list of settings, parameters and models.

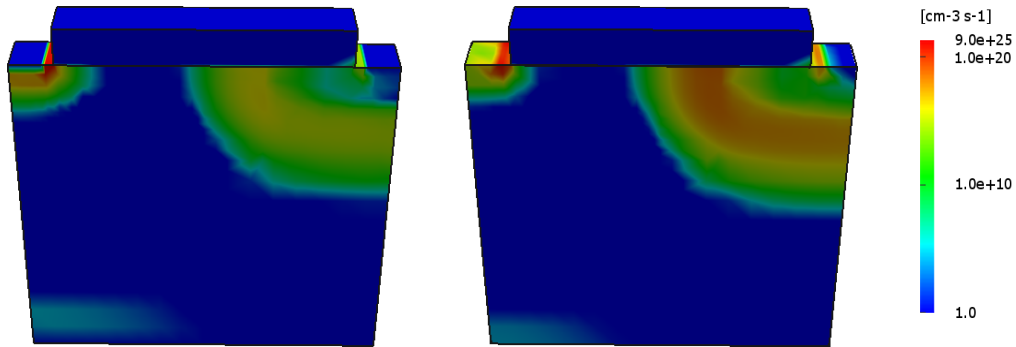




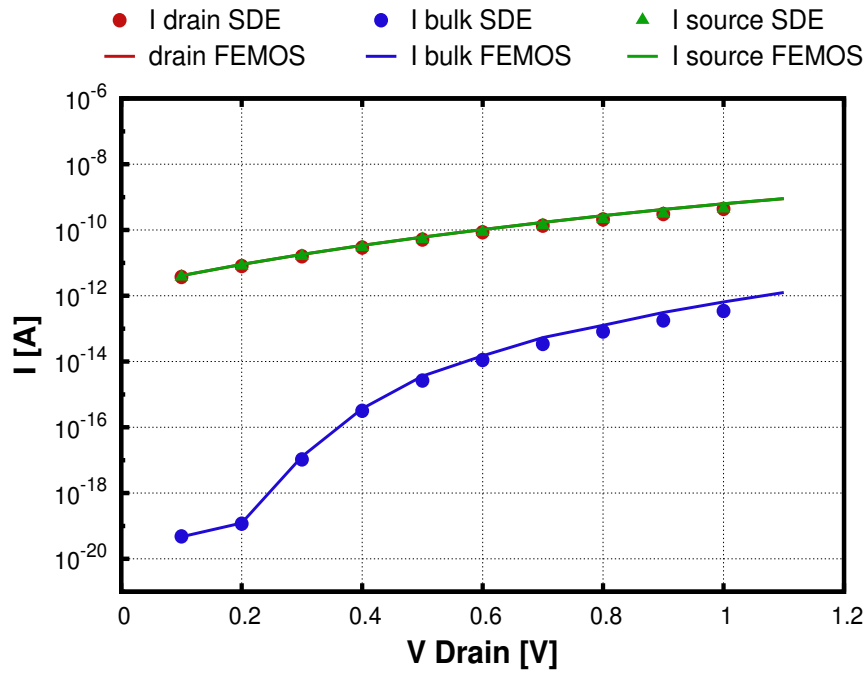
**Figure 4.46:** Contribution of the impact ionization with the Van Overstraeten - de Man model inside the nMOSFET  $V_D = 0.5[V]$ .



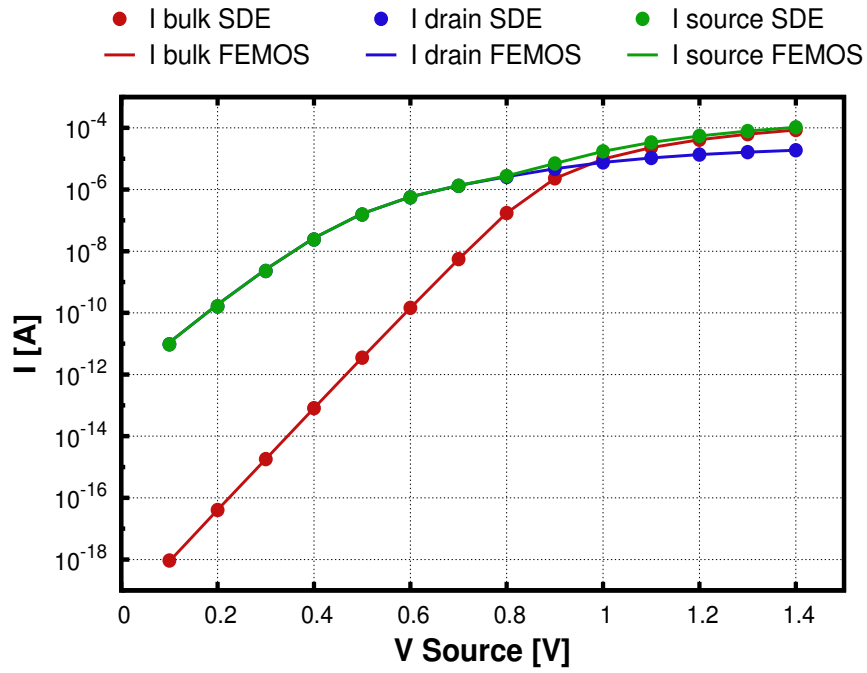
**Figure 4.47:** Contribution of the impact ionization with the Van Overstraeten - de Man model inside the pMOSFET  $V_S = 0.5[V]$ .



**Figure 4.48:** Contribution of the impact ionization with the Van Overstraeten - de Man model inside the pMOSFET  $V_S = 1.2[V]$ .



**Figure 4.49:** Inverse polarization of the nMOSFET.



**Figure 4.50:** Inverse polarization of the pMOSFET.

# Bibliography

- [BCC98] R. E. Bank, W. Coughran, and L. C. Cowsar. The finite volume sharfetter-gummel method for steady convection diffusion equations. *Computing and Visualization in Science*, 1:123–136, 1998.
- [BGB00] Streetman Ben G. and Sanjay Banerjee. *Solid State electronic Devices*. New Jersey: Prentice Hall, 2000.
- [Can75] C. Canali. Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature. *IEEE Transactions on Electron Devices*, 22:1045–1047, 1975.
- [Chy58] A. G. Chynoweth. Ionization rates for electrons and holes in silicon. *Physical Review*, 109:1537–1540, 1958.
- [Deu74] P. Deuffhard. A modified newton method for the solution of ill-conditioned system of nonlinear equations with application to multiple shooting. *Numer. Math*, 22:289–315, 1974.
- [DF05] Carlo De Falco. *Quantum Corrected Drift-Diffusion Models and Numerical Simulation of Nanoscale Semiconductor Devices*. Dottorato di ricerca in matematica e statistica per le scienze computazionali, A.A. 2004/2005.
- [GH92] H Goebel and K. Hoffmann. Full dynamic power diode model including temperature behavior for use in circuit simulations. *Proceedings of the 4th International Symposium on Power Semiconductor Devices & ICs*, pages 130–135, 1992.
- [GS74] R. Gusmeroli and A. S. Spinelli. Accurate boundary integrals calculation in semiconductor device simulation. *Numer. Math*, 22:289–315, 1974.

- [Gum64] H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *Electron Devices*, pages 455–465, 1964.
- [HEML00] Thomas J.R. Hughes, Gerald Engel, Luca Mazzei, and Mats G. Larson. The continuous galerkin method is locally conservative. *Journal of Computational Physics*, 163:467–488, 2000.
- [IE83] Babuška I. and Osborn J. E. Generalized finite element methods: Their performance and their relation to mixed methods. *Numer. Anal.*, 20:510–536, 1983.
- [LH80] W. Lochmann and A. Haug. Phonon-assisted auger recombination in si with direct calculation of the overlap integrals. *Solid State Communications*, 35:553–556, 1980.
- [Lom88] C. Lombardi. A physically based mobility model for numerical simulation of nonplanar devices. *IEEE Transactions on Computer-Aided Design*, 7:1164–1171, 1988.
- [MS83] G. Masetti and M. Severi. Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon. *IEEE Transactions on Electron Devices*, 7:764–769, 1983.
- [PC98] M. Putti and C. Cordes. Finite element approximation of the diffusion operator on tetrahedra. *Society for Industrial and Applied Mathematics*, 19:1154–1168, 1998.
- [PTVF07] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes, The Art of Scientific Computing*. Cambridge University Press, New York, 2007.
- [Qua08] Alfio Quarteroni. *Modellistica Numerica per Problemi Differenziali*. Springer Italia, Milan, 2008.
- [Sal10] Sandro Salsa. *Equazioni a Derivate Parziali, metodi, modelli e applicazioni*. Springer Italia, Milan, 2010.
- [Sde13] *Sentaurus Device User Guide*. Synopsis Inc., 2013.
- [vOdM70] R. van Overstraeten and H. de Man. Measurement of the ionization rates in diffused silicon p-n junctions. *Solid-State Electronics*, 13:583–608, 1970.

- [VOT83] R. Van Overstraeten and M. S. Tyagi. Minority carrier recombination in heavily-doped silicon. *Solid-State Electronics*, 26:577–597, 1983.
- [XZ99] Jinchao Xu and Ludmil Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Mathematics of Computation*, 68:1429–1446, 1999.
- [YT09] Taur Yuan and H. Ning Tak. *Fundamentals of Modern VLSI Devices*. Cambridge University Press, 2009.
- [ZL12] L. T. Zikatanov and R. D. Lazarov. An exponential fitting scheme for general convection-diffusion equations on tetrahedral meshes. *math*, pages 455–465, 2012.



# Ringraziamenti