Elaborato di Tesi ...

Titolo progetto di tesi ...

Progetto svolto da:
Andrea Bortolossi
Matr. 783023

Anno Accademico 2013–2014

Contents

Chapter 1

Introduction

1.1   Brief history of VLSI devices

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 Gordton 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 precise physic branch: obviously this

strategic choice guarantees more efficiency but it implies a lost in generality. The conseguence is that the work of the model analyst became harder when he have to afford problems located in the middle of different phenomenon.

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 SDE-VICE simulator) and the Navier-Lamè equations (which is well performed by COMSOL simulator). Now the question is: how to put in comunication the different outputs? Take into account that it's not possbile known precisely how the above programs resolves the equations, which implies a relevant risk 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 total control on simulation procedure and the possibility of fully customize, the major drawback is that the improvement of a personal code needs time and human resources, which in many cases are not avaible.

The FEMOS project (*Finite Element Method Oriented Solver*) tries to overcome the above limitations. Even if modern devices present innovative and unexpected behaviour, we can't avoid the treatment of the classical semiconductor devices from the simulation possibilties of FEMOS. This thesis found its origin in the development of this achievement, but as subject covered a spread wide area in terms of models and kind of devices and we focus on:

- development of a finite element based simulator for semiconductor devices which deals with multiple generation/recombination and mobility models;

- check solutions obtained against commercial software (SDEVICE);

- definition and implementation of a new way to compute the current density inside the device;

- extension of the residual method presented in referenza for the 3D case;

- evaluation of the possibility to extend the residual method at the computation of the current density inside the device.

Chapter 2

Semiconductor model

In this chapter we shall present the basic physics properties of semiconductor material accordingly with the quantum mechanics theory [YT09]. The Drift-Diffusion model is then presented.

2.1 Basic Device Physics

This section covers the basic concepts of semiconductor device physics. As the most used material in the fabrication of VLSI devices is silicon, in the follow we will focus on it.

2.1.1 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 a covalent bond. The most important result of the application of quantum mechanics to the description of electrons in a solid is that the allowed energy levels of electrons are grouped into bands. The bands are separated by regions of energy that the electrons in the solid cannot possess: forbidden gaps. The highest energy band that is completely filled by electron at 0[K] is called the *valence band* (E_V). The next higher energy band, separated by a frobidden gap from the valence band, is called the *conduction band* (E_C).

Because in silicon the band gap is on the order of 1 [eV], 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

(a) Metal   **(b)** Insulator
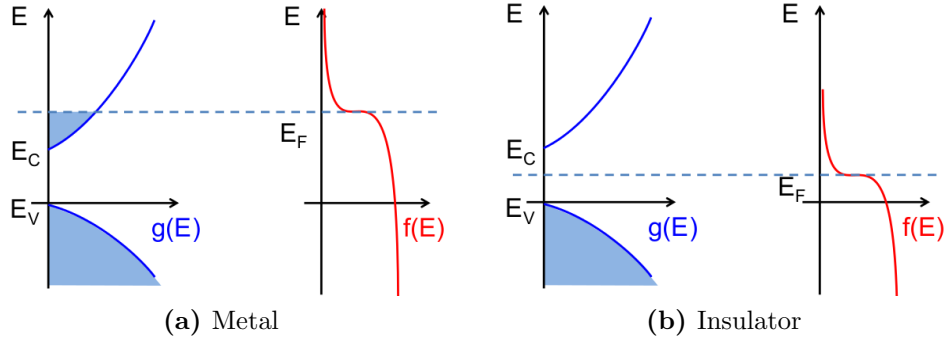
Figure 2.1: Two typical examples of state density occupation (g(E)) and probability distribution (f(E)).

even at absolute zero temperature, this make them good conductors at any temperature.

A suitable formulation of che carrier concentration which is given, for electrons, by the follow integral:

$$n = \int_{E_c}^{\infty} g(E)f(E)\,dE \qquad (2.1)$$

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 a suitable probability ditribution.

The energy distribution of electrons in a solid is governed by the laws of Fermi-Dirac statistics. For a system in thermal equilibrium, the principal result of these statistics is the *Fermi-Dirac distribution function*, which gives the probability that an electronic state at energy E is occupied by an electron,

$$f_D(E) = \frac{1}{1 + exp\left(\dfrac{E - E_f}{k_B T}\right)} \qquad (2.2)$$

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*.

Definition 2.1. 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 (2.2) can be approximated with the Maxwell-Boltzmann statistics for classical particles, which read as follows:

$$f_D(E) \simeq f_{MB}(E) = \begin{cases} exp\left(-\dfrac{E - E_f}{KT}\right) & E \gg E_f \\ 1 - exp\left(-\dfrac{E_f - E}{KT}\right) & E \ll E_f \end{cases} \qquad (2.3)$$

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

Observation 2.1. 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 and/or semiconductors in contact, the Fermi level at thermal equilibrium is flat (spatially constant throughout the region).

In general (2.1) is a Fermi integral of the order 1/2 and must be evaluated numerically. In the case of non-degenerate semiconductor, Fermi levels stay at least $3KT/q$ below the edge of the conduction band (for holes we consider the same approximation above the valence band). The Fermi-Dirac distribution can be approximated by the Maxwell-Boltzmann distribution and (2.1) can be solved in the analytically way, obtaining,

$$n = N_c exp\left(-\frac{E_c - E_f}{KT}\right) \qquad (2.4)$$

$$p = N_v exp\left(-\frac{E_f - E_v}{KT}\right) \qquad (2.5)$$

where N_c and N_v are the *effective density of states*. In intrisic semiconductor $n = p$ and the *intrinsic Fermi level* E_i can be calculated using equations (2.4) and (2.5) as:

$$E_i = E_f = \frac{E_c + E_v}{2} - \frac{KT}{2}ln\left(\frac{N_c}{N_v}\right) \qquad (2.6)$$

By replacing (2.6) in (2.4) we have the expression of the intrinsic carrier concentration $n_i = n = p$:

$$n_i = \sqrt{N_c N_v} exp\left(-\frac{E_g}{2KT}\right) \qquad (2.7)$$

Observation 2.2. Since the thermal energy, $k_B T$ is muc 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 (2.4) and (2.5) can be rewritten in terms of the intrinsic carrier density (n_i) and energy (E_i) :

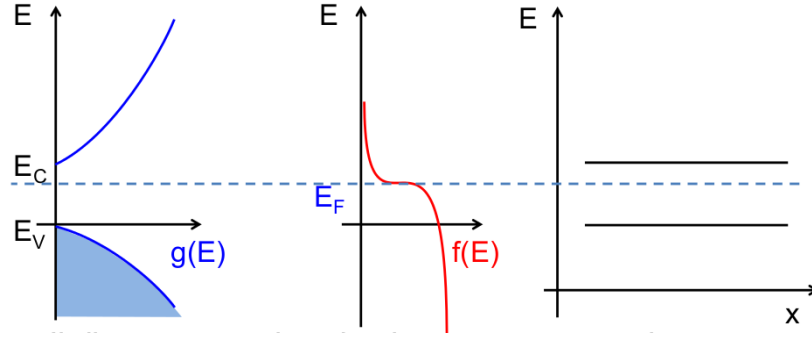$$n = n_i exp\left(\frac{E_f - E_i}{KT}\right) \tag{2.8}$$

$$p = n_i exp\left(\frac{E_i - E_f}{KT}\right) \tag{2.9}$$

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

$$np = n_i^2 \tag{2.10}$$

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

One of the most used graphical tool for the anlysis of the functionality of devices is the band diagram Fig.2.2, which summarizes the informations presented above.

Figure 2.2: Construction of tha band diagram.

2.1.2 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 it's usual add impurities atoms 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.

A column-V atom in a silicon lattice tends to have one extra electron loosely bonded after forming covalent bonds with other silicon atoms. In most cases, the thermal energy at room temperature is sufficient to ionize the impurity atom and free the extra electron to the conduction band. Such type of impurities are called *donors*; they become positively charged when ionized. Silicon material doped with column-V impurities or donors is called *n-type* silicon.

A column-III impurity atom in a silicon lattice tends to be deficient by one electron when forming covalent bonds with other silicon atoms. Such an impurity atom can also be ionized by accepting an electron from the valence band, which leaves a free-moving hole that contributes to electrical conduction. These 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.

In contrast to intrinsic silicon, the Fermi level in an extrinsic silicon is not located at the midgap. 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 example, for an n-type material with a donor impurity concentration, N_d, the charge neutrality condition requires that

$$n = N_d^+ + p \tag{2.11}$$

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

$$p = N_a^- + n \tag{2.12}$$

For the sake of simplicity we consider in this work that at room temperature all impurties are ionized ($N_d = N_d^+$ and $N_a = N_a^-$). Typcally the magnitude of impurities is between $10^{16} * 10^{20} [cm^{-3}]$, while the usual intrinsic carrier concentration is almost $10^{10} [cm_3]$, for this reason we can approximate the equilibrium densities concentration:

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

$$\tag{2.13}$$

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

Take into account this approximation and substituting (2.4) and (2.5) in (2.11) and (2.12), solving the algebraic equationwe have

$$E_c - E_f = KTln\left(\frac{N_c}{N_d}\right) \tag{2.14}$$

$$E_f - E_v = KTln\left(\frac{N_v}{N_a}\right) \tag{2.15}$$

Equation (2.10) is indipendent of the dopant type and Fermi level position. Instead of using N_c, N_v and referring to E_c and E_v equation (2.14) and (2.15) can be written in a more useful form in terms of n_i and E_i defined by equations (2.7) and (2.6):

$$E_f - E_i = KTln\left(\frac{N_d}{n_i}\right) \tag{2.16}$$

$$E_i - E_f = KTln\left(\frac{N_a}{n_i}\right) \tag{2.17}$$

Observation 2.3. The distance between the Fermi level and the intrinsic Fermi level near the midgap is a logarithmic function of doping concentration.

As a consequences:

- non linearity relations between potential and desities,

- exponential dependence of densities from potential.

(a) n-type   **(b)** p-type
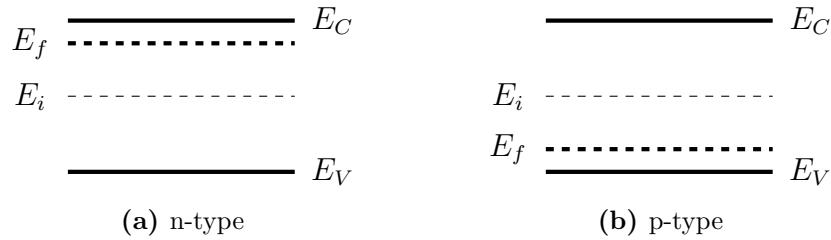
Figure 2.3: Band diagrams of estrinsic silicon.

2.1.3   Densities at nonequilibrium condition

In VLSI device operation a nonequilibrium sistuations is often possible: the densities of one or both types of carriers depart from their equilibrium values

given by (2.8) and (2.9). 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 separate Fermi levels for electrons and holes. They are called *quasi Fermi levels* defined as E_{fn} and E_{fp} replacing equation (2.8) and (2.9):

$$n = n_i exp \left(\frac{E_{fn} - E_i}{k_B T} \right) \tag{2.18}$$

$$p = n_i exp \left(\frac{E_i - E_{fp}}{k_B T} \right) \tag{2.19}$$

In non equilibrium condition quasi Fermi levels have a similar physical interpretation in terms of the state occupancy as the Fermi level.

Observation 2.4. The electron density in the conduction band can be calculated using E_{fn}, and the hole density in the valence band using E_{fp}.

2.1.4 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 and electron or hole concentration gradient.

Drif current - Ohm's law

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

Observation 2.5. 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 defnded as

$$\mathbf{v}_d^e = -\frac{q\mathbf{E}\tau_e}{m_e} \qquad \mathbf{v}_d^h = +\frac{q\mathbf{E}\tau_h}{m_h} \tag{2.20}$$

where $q = 1.602e^{-19}[C]$ is the elementary charge, $\mathbf{E}$ the electric field, τ_η the average time of flight of the carrier between two consecutive interactions with the atoms of the lattice and m_η is the effective mass. The coefficient $q\tau_\eta/m_\eta$ characterizes how quickly a carrier can move through the lattice and it's well known as carrier mobility $[m^2 V^{-1} s^{-1}]$. In general, to include different contributions to the mobility *Matthiessen's rule* is used:

$$\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_I} + \cdots \tag{2.21}$$

where μ_L and μ_I correspond to the lattice and impurity scattering limited components of mobility (for a more detailed description of mobility models see [YT09]).

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

$$\mathbf{J}_n = -qn\mathbf{v}_d^n = qn\mu_n\mathbf{E} = \sigma_n\mathbf{E} \tag{2.22}$$

$$\mathbf{J}_p = +qp\mathbf{v}_d^p = qp\mu_p\mathbf{E} = \sigma_p\mathbf{E} \tag{2.23}$$

The scalar coefficient $qn\mu_n(qp\mu_p)$ is often summerized by the electron (hole) conductivity $\sigma_n(\sigma_p)$.

Relations (2.22) and (2.23) expresses the well known *Ohm' 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*:

$$\mathbf{J}_n = -D_n(-q\nabla n) \tag{2.24}$$

$$\mathbf{J}_p = -D_p(+q\nabla p) \tag{2.25}$$

The proportionally constants D_n and D_p are called the electron and hole diffusion coefficients and have units of $[cm^2 s^{-1}]$. Physically, both drift and diffusion are closely associated with the random thermal motion of carriers and their collisions with the silicon lattice in thermal equilibrium. A simple relationship between the diffusion coefficient and the mobility is the well known *Einstein relation*:

$$D_\eta = \frac{k_B T}{q}\mu_\eta \tag{2.26}$$

Drift-Diffusion transport equations

By considering (2.22), (2.23), (2.24) and (2.25), the total electron and hole current densities are:

$$\mathbf{J}_n = qn\mu_n\mathbf{E} + qD_n\nabla n \tag{2.27}$$

$$\mathbf{J}_p = qp\mu_p\mathbf{E} - qD_p\nabla p \tag{2.28}$$

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

We remark that these constitutive laws can be rewritten in two other ways highlighting different physical explanations of the same phenomenon. Moreover these reinterpreations give different start points for the discrete solver algorithm.

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

$$\mathbf{E} = -\nabla\varphi \tag{2.29}$$

the current densities can be :

$$\mathbf{J}_n = -qn\mu_n\left(\nabla\varphi - \frac{k_BT}{qn}\nabla n\right)$$

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

Considering equations (2.18) and (2.19) the above can be written as:

$$\mathbf{J}_n = -qn\mu_n\nabla\varphi_n \tag{2.30}$$

$$\mathbf{J}_p = -qn\mu_p\nabla\varphi_p \tag{2.31}$$

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

Observation 2.6. 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*. In 1973 Jan Slotboom proposed this change in variables for the two-dimensional numercal simulation of a bipolar transistor:

$$u_n = n_i exp\left(-\frac{\varphi_n}{V_{th}}\right) \tag{2.32}$$

$$u_p = n_i exp\left(\frac{\varphi_p}{V_{th}}\right) \tag{2.33}$$

where $V_{th} = k_B T/q$. Using the above equations into (2.27) and (2.28) we obtain:

$$\mathbf{J}_n = qD_n exp\left(\frac{\varphi}{V_{th}}\right)\nabla u_n \tag{2.34}$$

$$\mathbf{J}_p = -qD_p exp\left(-\frac{\varphi}{V_{th}}\right)\nabla u_p \tag{2.35}$$

This interpretation results is:

Observation 2.7. The drift-diffusion current density in a semiconductor, is a totally diffusive flux of a new kind of carrier and diffusivity coefficient.

2.2   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 simulaiton of semiconductor devices. In this section we'll show how is possible to deduce the DD model.

2.2.1   Drift Diffusion formulation

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

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{2.36}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.37}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{2.38}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.39}$$

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

$$\begin{aligned}\mathbf{D} &= \epsilon\mathbf{E}\\ \mathbf{B} &= \mu_m\mathbf{H}\end{aligned} \tag{2.40}$$

where ϵ is the material dielectric permettivity $[Fcm^{-1}]$ and μ_m is the magnetic permeability $[Hcm^{-1}]$. Since $\nabla \cdot (\nabla \times \mathbf{A}) = 0$ for any vector $\mathbf{A}$,

(2.39) is satisfied by introducing a vector potential $\mathbf{A}$ such that $\mathbf{B} = \nabla \cdot \mathbf{A}$. We replace it in (2.37) and we obtain

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0. \tag{2.41}$$

From this we can state that exist a scalar potential φ such that

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \varphi \tag{2.42}$$

We multiply (2.42) by ϵ, we apply the divergence operator and we obtain using (2.29), (2.40) and (2.38)

$$\rho + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \cdot (\epsilon \nabla \varphi) \tag{2.43}$$

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

$$\nabla \cdot (\epsilon \nabla \varphi) = \rho. \tag{2.44}$$

We apply the divergence operator on the equation [2.36] and we get the *Continuity Equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \tag{2.45}$$

To close the above system we need to specify the mathematical form of the electric charge density (ρ) and the electric conduction current density ($\mathbf{J}$). As we introduced in the preview section, devices are usually formed by extrinsic semiconductor and this causes the presence in the lattice of two kind of charge:

- free charge (ρ_{free}) (free electron and holes carriers),

- fixed charge (ρ_{fixed}) (ionoized dopant impurities).

$$\rho = \underbrace{q(p - n)}_{\rho_{free}} + \underbrace{q(N_D^+ - N_A^-)}_{\rho_{fixed}} \tag{2.46}$$

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

Accordingly with the preview hypotesis and replacing (2.46), (2.27) and (2.28), we can split the continuity equation into the contribute of electrons and holes, the DD model formulation looks as follows:

$$
\begin{cases}
\nabla \cdot (-\epsilon \nabla \varphi) \;=\; q(p - n + N_D^+ - N_A^-) \\[2mm]
-q\dfrac{\partial n}{\partial t} + \nabla \cdot (-q\mu_n n \nabla \varphi + qD_n \nabla n) \;=\; qR \\[2mm]
q\dfrac{\partial p}{\partial t} + \nabla \cdot (-q\mu_p p \nabla \varphi - qD_p \nabla p) \;=\; -qR
\end{cases}
\tag{2.47}
$$

The system is an incompletely parbolic initial value/boundary problem in three scalar unkown dependent variables $\varphi(\mathbf{x}, t)$, $n(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$. Notice that the problem is a nonlinearly coupled system of PDE's, because of the presence of the drift terms ($n\nabla\varphi$ and $p\nabla\varphi$).

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 $\mathbf{J_p}$. We can interpret $R(\mathbf{x}, t)$ as the net rate of generation and recombination.

The stationary form can be easily derivate from (2.47) by not considering the temporal derivative.

2.2.2   Generation and Recombination phenomenon

The modelling of $R(\mathbf{x}, t)$ is one of the most important feature due to the important role in determining the current-voltage characteristic of devices.

It's important to keep in mind that electrons and holes are in continuos fluctuation due to their thermal energy, but the macroscopic result of such a process at equilibrium is that the net recombination rate is identically zero at each point and at each time level. Therefore our interest is to analyze the deviations from this condition.

In every moment the system try to mantain the equilibrium, so it's important underlying that the response with a recombination event happens in order to neutralize an excess of charge, while generation event are usually due to thermal agitation or an external input source.

The phenomenological model for the net recombination rate R is often given by the sequent formulation:

$$
R(n, p) = (pn - n_i^2)F(n, p)
\tag{2.48}
$$

where F is a function modelling the specific recombination/generation (R/G) event.

In the following we present the classical theory about three kind of contribute.

Shockley-Read-Hall recombination (SRH)

Electron and hole generation and recombination can take place directly between the valence band and the conduction band, or inderactly via trap centers in the energy gap (we indicate with E_T the energy level at where these traps live). The latter category includes Shockley-Read-Hall phenomena (SRH), more precisely SRH rate is a two-particle process which matematically expresses the probality that:

R_{SRH} 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,

G_{SRH} an electron is emitted from the valence band to the conduction band, through he mediation of an unoccupied trapping level located in the energy gap.

The following expression is usually employed for the modulating function F:

$$F_{SRH}(n,p) = \frac{1}{\tau_n \left(p + n_i cosh \left(\frac{E_T}{k_B T} \right) \right) + \tau_p \left(n + n_i cosh \left(\frac{E_T}{k_B T} \right) \right)} \quad (2.49)$$

the quantiaties τ_n and τ_p are called *carrier lifetimes* and are phisically defined as the reciprocals of the capture rates per single carrier associated with the energy trap distribution within the semiconductor energy gap. Their typical order of magnitude lies in the range $10^{-3} \mu s \div 1 \mu s$.

Parameter	Unit	Electrons	Holes
τ	s	1.0×10^{-5}	3.0×10^{-6}
E_T	eV	0.0	0.0

Table 2.1: List of parameters in the electron and hole mobility models including scattering from lattice thermal vibrations

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 following expression is usually employed for the modulating function F:

$$F_{AU}(n,p) = C_n n + C_p p \qquad (2.50)$$

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

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

Table 2.2: List of parameters in the electron and hole Auger generation/recombination model.

Impact ionization (II)

The impact ionization mechanism is a three-particle generation process and it is dissimilar from the previously phenomenon because we can't express its

contribute with a relation like (2.48). The high energy carrier generation is triggered by the presence of very high electric fields. Due to these fields 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 of generated carriers is produced within the region. There are several different models for the II generation, inside our code we implemented the van Overstraeten - de Man referenza manuale sdevice model based on the Chynoweth law referenza dentro sdevice:

$$G_{II}(n,p) = \alpha_n n |\mathbf{v}_n| + \alpha_p p |\mathbf{v}_p| \tag{2.51}$$

where:

$$\alpha(E_{ava}) = \gamma exp\left(-\frac{\gamma b}{E_{ava}}\right) \tag{2.52}$$

$$\gamma = \frac{tanh\left(\dfrac{\hbar\omega_{op}}{2k_B T_0}\right)}{tanh\left(\dfrac{\hbar\omega_{op}}{2k_B T}\right)} \tag{2.53}$$

The factor γ with the optical phono energy $\hbar\omega_{op}$ expresses the temperature dependence of the phonon gas against which carriers are accelerated. E_{ava} is the driving force which takes into account how the electric field influence the generation event. There are two possibilties to compute this quantity:

- compute the component of the elctrostatic field in the direction of the current

$$E_{ava}^{n,p} = \frac{\mathbf{E} \cdot \mathbf{J}_{n,p}}{||\mathbf{J}_{n,p}||} \tag{2.54}$$

- consider the module of the quasi fermi gradient

$$E_{ava}^{n,p} = |\nabla\varphi_{n,p}| \tag{2.55}$$

2.2.3   Mobility models

In the following section we illustrate the most commonly used phenomenological models to describe carrier mobilities. More precisely we want describe the several mechanisms that characterize the average time of flight (2.20). Scattering phenomenon slow down the motion of carriers throughout the lattice and the three main physical principles governing these events are:

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

Table 2.3: List of parameters in the electron and hole of van Overstraeten-de Man model

Parameter	Unit	Electrons	Holes
μ^0	$cm^2 V^{-1} s^{-1}$	1417.0	470.5
β	1	2.5	2.2

Table 2.4: List of parameters in the electron and hole mobility models including scattering from lattice thermal vibrations

- interaction with the thermally generated vibrations of the silicon atoms;

- presence of ionized dopant impurities in the crystal;

- reduction to the velocity saturation at high electric fields.

Scattering from thermal vibrations

Intuitively, carrier mobility is to be a decreasing function of temperature, as we expect collisions to become more and more frequent as T gets higher. This idea is commonly represented using a simple power law of the form

$$\mu_\nu^L = \mu_\nu^0 \left(\frac{T}{T_0} \right)^{-\beta_\nu} \qquad \nu = n, p \qquad (2.56)$$

where μ_ν^0 is the low-field mobility, β_ν are positive numbers and T_0 is a reference temperature typically $T_0 = 300[K]$.

Scattering from Ionized Impurities

Dopant ionized impurities represent local perturbations of the periodic distribution of silicon atoms. They strongly influence the carrier motion through electrostatic interaction, reducing the mobility of electrons and holes. The

model used to simulate doping-dependent mobility in silicon was proposed by *Masetti*:

$$\mu = \mu_{min1} exp\left(-\frac{P_c}{N_{tot}}\right) + \frac{\mu^L - \mu_{min2}}{1 + \left(\frac{N_{tot}}{C_r}\right)^{\alpha}} - \frac{\mu_1}{1 + \left(\frac{C_s}{N_{tot}}\right)^{\beta}} \tag{2.57}$$

where $N_{tot} = N_D^+ + N_A^-$, μ_ν^L is given by (2.56), μ_{min1} and μ_{min2} are a positive quantities representing the minimum value of μ, P_c, C_r and C_s are reference doping values.

Parameter	Unit	Electrons	Holes
μ_{min1}	$cm^2V^{-1}s^{-1}$	52.2	44.9
μ_{min2}	$cm^2V^{-1}s^{-1}$	52.2	0
μ_1	$cm^2V^{-1}s^{-1}$	43.4	29.0
P_c	cm^{-3}	0	9.23×10^{16}
C_r	cm^{-3}	9.68×10^{16}	2.23×10^{17}
C_s	cm^{-3}	3.43×10^{20}	6.10×10^{20}
α	1	0.680	0.719
β	1	2.0	2.0

Table 2.5: 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 is completely wrong as it would predict an unbounded carrier velocity in the material as $|\mathbf{E}| \to \infty$. On the contrary, carrier scattering with lattice phonos produces a limitation to carrier velocity according to the following mathematical expression

$$\lim_{|\vec{E}| \to \infty} \mu|\mathbf{E}| = v_{sat} \tag{2.58}$$

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

$$\mu = \frac{\mu_L}{\left[1 + \left(\dfrac{\mu_L |\mathbf{E}|}{v_{sat}} \right)^\beta \right]^{1/\beta}} \tag{2.59}$$

$$v_{sat} = v_0 exp \left(\frac{300}{T} \right)^{v_{exp}} \qquad \beta = \beta_0 \left(\frac{T}{300} \right)^{\beta_{exp}}. \tag{2.60}$$

Parameter	Unit	Electrons	Holes
v_0	$cm\ s^{-1}$	1.07×10^7	8.37×10^6
v_{exp}	1	0.87	0.52
β_0	1	1.109	1.213
β_{exp}	1	0.66	0.17

Table 2.6: List of parameters in the electron and hole mobility models including scattering from velocity saturation.

Chapter 3

Resolution of the system

The aim of this chapter is to introduce notation and to present the physical model exposed in the previous chapter, in a more precisely and mathematical form. A briefly overview on the avaible and used iteration algorithms is then treated.

3.1   Geometry and boundary conditions

Consider the stationary form of problem (2.47), in order to proceed we have to close the *Poisson equation* and the *Drift Diffusion equation* for electrons and holes with suitable boundary conditions over the domain.

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

$$n = p = 0$$
$$\mathbf{J}_n = \mathbf{J}_p = \mathbf{0} \tag{3.1}$$

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. Moreover 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 material. This event is well performed by suitable Dirichlet boundary conditions, therefore in the follow we indicate $\partial\Omega_c = \Gamma_D$.

$$\begin{aligned} \varphi &= \varphi_D \\ n &= n_D \qquad on\,\Gamma_D. \\ p &= p_D \end{aligned} \qquad (3.2)$$

In the case where the neighbouring material is a perfect insulator, hypotesis (3.1) takes in place and therefore no Drift-Diffusion equations are solved, this implies that (3.2) reduces to the sole condition on the potential.

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

$$\begin{aligned} \mathbf{D} \cdot \mathbf{n} &= 0 \\ \mathbf{J}_n \cdot \mathbf{n} &= 0 \qquad on\,\Gamma_N \\ \mathbf{J}_p \cdot \mathbf{n} &= 0 \end{aligned} \qquad (3.3)$$

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 (3.3) reduces to the sole one related with the Poisson equation.

Finally when we solve a Drift-Diffusion equation, if $\Omega_{ox} \neq \varnothing$ boundary shape could be change. Although the treatment of boundaries doesn't change

(a)   (b)
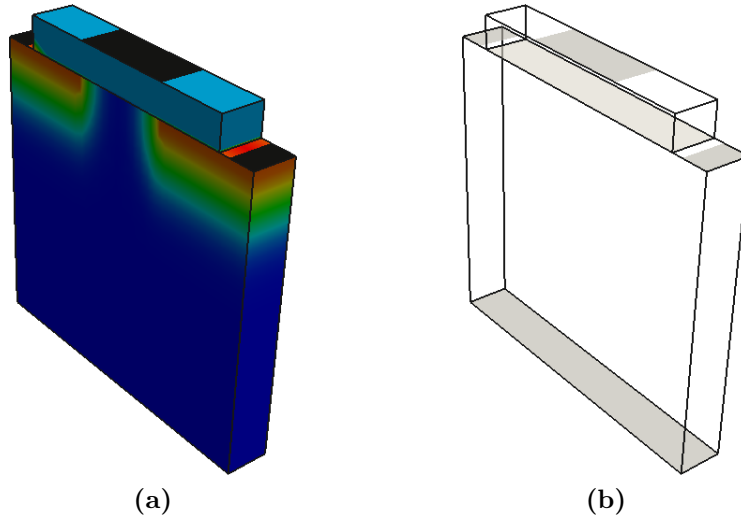
Figure 3.1: (a) Mos device with net dopant concentration distributed accordingly to a gaussian profile and Γ_D colored in black. The oxide layer is colored in light blue. (b) Outline of the Mos device with Γ_{int} in light gray.

if you consider the suitable restrictions and extensions

$$\begin{aligned}
\Gamma_{D,Si} &= \Gamma_D \cap \partial\Omega_{Si} \\
\Gamma_{N,Si} &= \Gamma_N \cap \partial\Omega_{Si} \cup \Gamma_{int}.
\end{aligned} \tag{3.4}$$

In silicon material the ideal condition on contacts is accomplished if both thermodynamical equilibrium and charge neutrality are reached. These two conditions correspond to the following algebraic system for n_D and p_D

$$\begin{cases}
p_D n_D & = & n_i^2 \\
p_D - n_D + N_D - N_A & = & 0
\end{cases}. \tag{3.5}$$

Solving (3.5) on $\Gamma_{D,Si}$ we obtain:

$$n_D = \frac{D + \sqrt{D^2 + 4n_i^2}}{2} \tag{3.6}$$

$$p_D = \frac{-D + \sqrt{D^2 + 4n_i^2}}{2} \tag{3.7}$$

where $D := N_D - N_A$ is the net doping concentration. Furthermore we can approximate that the quasi fermi potential level of silicon at contact is aligend with the external applyed voltage. As a conseguence we can easily calculate potential condition on $\Gamma_{D,Si}$ using (2.8) and (2.9)

$$\varphi_D = \varphi_f + V_{th} ln\left(\frac{n_D}{n_i}\right) = \varphi_f - V_{th} ln\left(\frac{p_D}{n_i}\right) \tag{3.8}$$

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

As we conveyed the boundary conditions, we are able to write in the closed way the stationary form of problem (2.47) as follows:

$$\begin{aligned}
-\Delta\epsilon\varphi - q(p - n) &= qD & in\,\Omega = \Omega_{ox} \cup \Omega_{Si} \\
\varphi &= \varphi_D & on\,\Gamma_D \\
\nabla\varphi \cdot \mathbf{n} &= 0 & on\,\Gamma_N
\end{aligned}$$

$$\begin{aligned}
\nabla \cdot (q\mu_n n\nabla\varphi - qD_n\nabla n) &= -qR & in\,\Omega_{Si} \\
n &= n_D & on\,\Gamma_{D,Si} \\
\nabla n \cdot \mathbf{n} &= 0 & on\,\Gamma_{N,Si}
\end{aligned} \tag{3.9}$$

$$\begin{aligned}
\nabla \cdot (-q\mu_p p\nabla\varphi - qD_p\nabla p) &= -qR & in\,\Omega_{Si} \\
p &= p_D & on\,\Gamma_{D,Si} \\
\nabla p \cdot \mathbf{n} &= 0 & on\,\Gamma_{N,Si}
\end{aligned}$$

The high coupled nonlinear nature of system (3.9) makes an analytical treatment very difficult, if not even impossible. For this reason, numerical schemes must be used to compute an approximate solution.

3.2   Iteration algorithms

The most used algorithms are *the fully coupled Newton's method* and *the decoupled Gummel map*. First of all we introduce a more compact form of (3.9):

$$\mathbf{F}(\mathbf{U}) = \mathbf{0} \tag{3.10}$$

where:

$$\mathbf{U} := [\varphi, n, p]^T \qquad \mathbf{F}(\mathbf{U}) := \left[\begin{array}{c} F_1(\mathbf{U}) \\ F_2(\mathbf{U}) \\ F_3(\mathbf{U}) \end{array} \right] \tag{3.11}$$

and having set:

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

The non linear problem (3.10) is the abstract generalization of the search of a zero for a real function $f : \mathbb{R} \to \mathbb{R}$. The vector function $\mathbf{F}$ is a nonlinear differential operator and the associated problem which we intend to resolve is: given a functional space V and the operator $\mathbf{F} : V \to V$, find $\mathbf{U} \in V$ such that (3.10) is satisfied.

In our application, the function space V is tipycally a subset of the Sobolev space $[H^1(\Omega)]^d$ (in this case with d we intend the number of component of $\mathbf{F}$). Sobolev spaces are based on the space of square integrable functions on Ω which reads as follows

$$L^2(\Omega) := \left\{ v : \int_{\Omega} |v|^2 d\Omega = ||v||^2_{L^2(\Omega)} < +\infty \right\}. \tag{3.12}$$

The general form of a Sobolev space for an integer $m \geq 0$ is

$$H^m(\Omega) := \left\{ v : D^\alpha v \in L^2(\Omega), \forall |\alpha| \leq m \right\}. \tag{3.13}$$

On this space, we shall use the semi-norm

$$|v|^2_{m,\Omega} = \sum_{|\alpha|=m} ||D^\alpha v||^2_{L^2(\Omega)} \tag{3.14}$$

and the norm

$$||v||^2_{m,\Omega} = \sum_{k \le m} |D^\alpha v|^2_{k,\Omega} \tag{3.15}$$

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

$$H^1_0 := \left\{ v : v \in H^1(\Omega), v|_{\partial\Omega} = 0 \right\} \tag{3.16}$$

$$H^1_{0,\Gamma_D} := \left\{ v : v \in H^1(\Omega), v|_{\Gamma_D} = 0 \right\} \tag{3.17}$$

For $v \in H^1_0(\Omega)$ we have the *Poincaré inequality*

$$|v|_{0,\Omega} \le C(\Omega)|v|_{1,\Omega} \tag{3.18}$$

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

The above setting of function spaces are used widely in the proceeding of this work.

3.2.1 Abstract Newton's method

Before introducing the abstract Newton's method we have to define some useful concept.

Definition 3.1 (Frechèt differentiable)**.** Let be X and Y two vector spaces. Given $f, g \in X$ and a functional $F : X \to Y$, the functional F is Frechét differentiable if it exists a linear bounded operator $A_f : X \to Y$ such that:

$$\lim_{||g|| \to 0} \frac{||F(f+g) - F(f) - A_f(g)||_Y}{||g||_X} = 0 \tag{3.19}$$

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

Considering the functional operator (3.11) 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.

$$\mathbf{F}'_{ij}(\mathbf{U})[\mathbf{V}]_j := \lim_{\eta \to 0} \frac{F_i(\mathbf{U} + \eta[\mathbf{V}]_j) - F_i(\mathbf{U})}{\eta} \qquad \mathbf{V} \in V \tag{3.20}$$

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 abstract Newton's method reads as follows:

Newton's method

Let be $\mathbf{F}$ a function operator, given $\mathbf{U}^0 \in V$, for all $k \geq 0$ until convergence, solve the following linear problem:

$$\begin{aligned} \mathbf{F}'(\mathbf{U}^k)\delta\mathbf{U}^k &= -\mathbf{F}(\mathbf{U}^k) \\ \mathbf{U}^{k+1} &= \mathbf{U}^k + \delta\mathbf{U}^k \end{aligned} \tag{3.21}$$

The application of Newton's method has transformed the original problem (3.10) into the *fixed-point problem* of finding $\mathbf{U} \in V$ such that

$$\mathbf{U} = T_{\mathbf{F}}(\mathbf{U}) \tag{3.22}$$

where

$$T_{\mathbf{F}}(\mathbf{U}) = \mathbf{F}'(\mathbf{U})^{-1}(\mathbf{F}'(\mathbf{U})\mathbf{U} - F(\mathbf{U})) \tag{3.23}$$

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

Theorem 3.1. *Let $\mathbf{U} \in V$ be a solution of problem (3.10). Assume that $\mathbf{F}'$ is Lipschitz continuos in the ball $\mathcal{B}(\mathbf{U}, \delta)$, i.e., that there exists $K > 0$ such that:*

$$||\mathbf{F}'(\mathbf{v}) - \mathbf{F}'(\mathbf{z})||_{L(V,V)} \leq K||\mathbf{v} - \mathbf{z}||_V \qquad \forall \mathbf{v}, \mathbf{z} \in \mathcal{B}(\mathbf{U}, \delta), \mathbf{v} \neq \mathbf{z} \tag{3.24}$$

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 (3.21) converges quadratically to $\mathbf{U}$, i.e., there exists $C > 0$ such that, for a suitable $k_0 \geq 0$ we have:

$$||\mathbf{U} - \mathbf{U}^{k+1}||_V \leq C||\mathbf{U} - \mathbf{U}^k||_V^2 \qquad \forall k \geq k_0 \tag{3.25}$$

3.2.2 Fully coupled Newthon's method

If we consider the linearization of the entire system (3.9) the relative Jacobian matrix is a 3x3 matrix and the problem reads as

$$
\begin{bmatrix} F_{1,\varphi} & F_{1,n} & F_{1,p} \\ F_{2,\varphi} & F_{2,n} & F_{2,p} \\ F_{3,\varphi} & F_{3,n} & F_{3,p} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta n \\ \delta p \end{bmatrix} = \begin{bmatrix} -F_1(\varphi,n,p) \\ -F_2(\varphi,n,p) \\ -F_3(\varphi,n,p) \end{bmatrix}. \tag{3.26}
$$

Every 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 $\delta\varphi$, δn and δp we note that the structure of the relative discretizated matrix is a 3x3 block matrix system where every block is a $N_{dof} \times N_{dof}$ matrix:

$$
\begin{bmatrix} \mathbf{K}_{1,\varphi} & \mathbf{K}_{1,n} & \mathbf{K}_{1,p} \\ \mathbf{K}_{2,\varphi} & \mathbf{K}_{2,n} & \mathbf{K}_{2,p} \\ \mathbf{K}_{3,\varphi} & \mathbf{K}_{3,n} & \mathbf{K}_{3,p} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta n \\ \delta p \end{bmatrix} = \begin{bmatrix} -\mathbf{F}_1(\varphi,n,p) \\ -\mathbf{F}_2(\varphi,n,p) \\ -\mathbf{F}_3(\varphi,n,p) \end{bmatrix}. \tag{3.27}
$$

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 particularly important to provide a very good initial guess vector $[\varphi_0, n_0, p_0]$, therefore we have to know for every equation a possibly shape of the solution.

Finally variable in play have different order of magnitude and the jacobian matrix is often quite ill-conditioned and needs appropriate scaling and balancing 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, although we chose an alternative algorithm knwon as *Decoupled Gummel Map.*

3.2.3 Gummel map algorithm

In 1964 H. K. Gummel proposed an orginal and alternative iterative algorithm in order to solve the system (3.9) in a semiconductor device in one spatial dimension.

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 (2.8) and holes (2.9)

$$
F_1(\varphi) = \nabla \cdot (-\epsilon\nabla\varphi) - q(n_i(e^{((\varphi_p-\varphi)/V_{th})} - e^{((\varphi-\varphi_n)/V_{th})}) + D). \tag{3.28}
$$

We have decoupled the system and the new algorithm reads as

(Step 0) Give a suitable initial condition for φ^0 and set a positive parameter $toll_{GM} > 0$ (Gummel Map tollerance)

(Step 1) Fix a positive parameter $toll_{NLP} > 0$ (Non Linear Poisson tollerance), solve the linearized Non Linear Poisson equation (NLP) in Ω using the Newton's method untill $||F_1(\varphi^{k+1})|| > toll_{NLP}$:

$$
\begin{cases}
\nabla \cdot (-\epsilon_{Si}\nabla\delta\varphi^k) + \dfrac{1}{V_{th}}\sigma_{Si}^k\delta\varphi^k = f_{Si}^k & in \ \Omega_{Si} \\
\nabla \cdot (-\epsilon_{ox}\nabla\delta\varphi^k) = f_{ox}^k & in \ \Omega_{ox} \\
\delta\varphi^k = 0 & on \ \Gamma_D \\
\nabla\delta\varphi^k \cdot \mathbf{n} = 0 & on \ \Gamma_N \\
\varphi^{k+1} = \varphi^k + \delta\varphi^k
\end{cases}
\tag{3.29}
$$

having set,

$$
\sigma_{Si}^k(\varphi^k) = qn_i\left[e^{((\varphi_p-\varphi^k)/V_{th})} - e^{((\varphi^k-\varphi_n)/V_{th})}\right]
$$

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

$$
f_{ox}^k(\varphi^k) = \nabla \cdot (-\epsilon\nabla\varphi^k)
$$

computed accordingly to the definition of the Frechét derivate.

(Step 2) Solve the Linear Electron Contintuity Equation (LEC):

$$
\begin{cases}
\nabla \cdot (q\mu_n n\nabla\varphi^i - qD_n\nabla n) = -qR(n^{i-1}, p^{i-1}) & in \ \Omega_{Si} \\
n = n_D & on \ \Gamma_{D,Si} \\
\nabla n \cdot \mathbf{n} = 0 & on \ \Gamma_{N,Si}
\end{cases}
\tag{3.30}
$$

(Step 3) Solve the Linear Hole Contintuity Equation (LHC):

$$
\begin{cases}
\nabla \cdot (-q\mu_p p\nabla\varphi^i - qD_p\nabla p) = -qR(n^{i-1}, p^{i-1}) & in \ \Omega_{Si} \\
p = p_D & on \ \Gamma_{D,Si} \\
\nabla p \cdot \mathbf{n} = 0 & on \ \Gamma_{N,Si}
\end{cases}
\tag{3.31}
$$

(Step 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. A concisely scheme is presented in Fig.3.2. Unfortunately there isn't any convergence result for this method like (3.1), although there are several advantages which make Gummel map algorithm preferable to the Fully Coupled Newton's Method. In fact simulations experience shows that the Gummel process is much more insesitive 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 stage cost: at each iteration step the Gummel algorithm requires the successive solution of three problems, each one of size equal to $N_{dof} \times N_{dof}$.
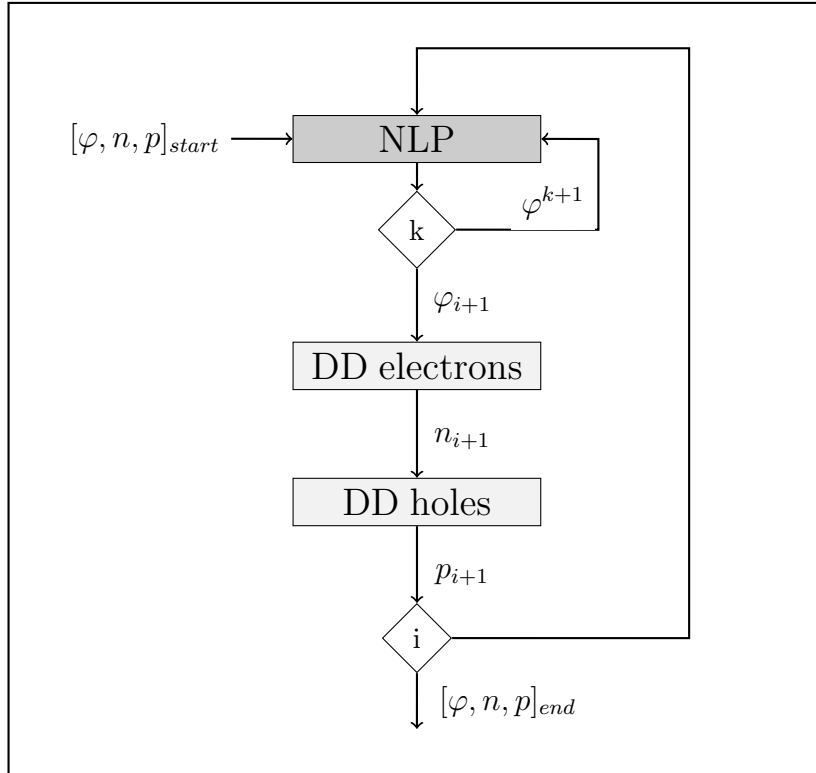
Figure 3.2: Gummel map algorithm

We spent some considerations over step 2 and step 3. Accordingly with (2.48) 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). In fact we can denote the general R/G term as follows:

$$
\begin{array}{ll}
R_n^{i-1}(n) &= \sigma_n^{i-1} n - f^{i-1} \\
R_p^{i-1}(p) &= \sigma_p^{i-1} p - f^{i-1}
\end{array}
\tag{3.32}
$$

where

$$
\sigma_n = \frac{p^{i-1}}{F(p^{i-1}, n^{i-1})} \qquad \sigma_p = \frac{n^{i-1}}{F(p^{i-1}, n^{i-1})}
$$
$$
f = \frac{n_i^2}{F(p^{i-1}, n^{i-1})}.
\tag{3.33}
$$

We can rewrite systems (3.30) and (3.31)

$$
\begin{cases}
\nabla \cdot (q\mu_n n \nabla \varphi^i - q D_n \nabla n) + q\sigma_n^{i-1} n = q f^{i-1} & in \ \Omega_{Si} \\
n = n_D & on \ \Gamma_{D,Si} \\
\nabla n \cdot \mathbf{n} = 0 & on \ \Gamma_{N,Si}
\end{cases}
\tag{3.34}
$$

$$
\begin{cases}
\nabla \cdot (-q\mu_p p \nabla \varphi^i - q D_p \nabla p) + q\sigma_p^{i-1} p = q f^{i-1} & in \ \Omega_{Si} \\
p = p_D & on \ \Gamma_{D,Si} \\
\nabla p \cdot \mathbf{n} = 0 & on \ \Gamma_{N,Si}
\end{cases}
\tag{3.35}
$$

Systems (3.34) and (3.35) are just a special treatment of the continuity equations in the Gummel Decoupled Algorithm. 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 algbraic 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 tha classical *Gauss-Seidel* method for the iterative solution in linear algebraic systems.

Chapter 4

Finite element discretization

In this section we shall present the classical variational formulation of problems (3.29), (3.34) and (3.35). For each kind of PDE problem we give a briefly presentation of the well-posedness analysis. Finally we describe the finite element discretization.

4.1   Non Linear Poisson Equation weak form

Let us write problem (3.29) in a more compact form

$$
\begin{cases}
\nabla \cdot (-\epsilon \nabla \delta\varphi^k) + \sigma^k \delta\varphi^k & = & f^k & \quad in\ \Omega \\
\delta\varphi^k & = & 0 & \quad on\ \Gamma_D \\
\nabla \delta\varphi^k \cdot \mathbf{n} & = & 0 & \quad on\ \Gamma_N \\
\varphi^{k+1} & = & \varphi^k + \delta\varphi^k
\end{cases}
\tag{4.1}
$$

having set

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

where $\mathcal{I}_A(\mathbf{x})$ is equal to 1 if $\mathbf{x} \in A$ and 0 otherwise. System (4.1) constitutes a classical Diffusion-Reaction (DR) problem in Ω, respect the variable $\delta\varphi^k$. Now we multiply the equation with a test function $v \in H^1_{\Gamma_D}$ and we integrate over all the domain

$$
-\int_\Omega \epsilon \Delta \delta\varphi^k v\, d\Omega + \int_\Omega \sigma^k \delta\varphi^k v\, d\Omega = \int_\Omega f^k v\, d\Omega \qquad \forall v \in H^1_{\Gamma_D}(\Omega). \tag{4.2}
$$

Applying the Green-formula on (4.2) and then considering the boundary conditions, we get the weak formulation which reads as: find $\delta\varphi^k \in H^1_{\Gamma_D}(\Omega)$ such that

$$\int_\Omega \epsilon \nabla \delta\varphi^k \nabla v \, d\Omega + \int_\Omega \sigma^k \delta\varphi^k v \, d\Omega = \int_\Omega f^k v \, d\Omega \qquad \forall v \in H^1_{\Gamma_D}(\Omega) \qquad (4.3)$$

We are able to define the following bilinear form

$$a : H^1_{\Gamma_D}(\Omega) \times H^1_{\Gamma_D}(\Omega) \to \mathbb{R}, \qquad a(u,v) = \int_\Omega \epsilon \nabla u \nabla v \, d\Omega + \int_\Omega \sigma^k uv \, d\Omega \quad (4.4)$$

and the linear and bounded functional

$$F : H^1_{\Gamma_D}(\Omega) \to \mathbb{R}, \qquad F(v) = \int_\Omega f^k v \, d\Omega \qquad\qquad (4.5)$$

In order to prove the existence and uniqueness of the solution, we apply the *Lax-Millgram theorem* [Sal10] to the weak formulation (4.3). The well-posedness is ensured by several and physical hypotesis:

- $\epsilon \in L^\infty(\Omega)$ and $\exists m$ s.t. $0 < m \le \epsilon$ (a.e.) in Ω;

- $\forall k \ge 0 \; \sigma^k \in L^\infty(\Omega)$ and $\exists m$ s.t. $0 < m \le \sigma^k$ (a.e.) in Ω_{Si}.

We define some useful quantities:

$$\epsilon_M = max_\Omega \epsilon \quad \epsilon_m = min_\Omega \epsilon$$
$$\sigma_M = max_\Omega \sigma \quad \sigma_m = max_\Omega \sigma = 0$$

Take into account the above hypotesis one can demonstrate:

- **Continuity of the bilinear form**,

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

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

- **Coercivity of the bilinear form**,

  $\forall u \in H^1_{\Gamma_D}$

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

- **Continuity of the functional,**

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

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

4.2 Continuity Equations weak form

Without loss of generality we can consider only the electron continuity equation. System (3.34) 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 (2.32), we can rewrite system (3.34)

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

(4.6)

We can easily obtain the weak formulation repeating the same steps made for the linearized Non Linear Poisson problem. Therefore the weak formulation of the Electron Continuity equation is: find $u_n \in H^1_{\Gamma_{D,Si}}(\Omega)$ such that:

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

(4.7)

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. We have to make some 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 Ω_{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 Ω_{Si}.

As we made before, we define the relative bilinear form

$$a(u,v) = \int_{\Omega_{Si}} qD_n e^{(\varphi^i/V_{th})}\nabla u_n \nabla v\, d\Omega + \int_{\Omega_{Si}} \sigma_n^{i-1}e^{(\varphi^i/V_{th})}u_n v\, d\Omega \qquad (4.8)$$

and the linear and bounded functional

$$F(v) = \int_{\Omega_{Si}} f^{i-1} v \, d\Omega \tag{4.9}$$

Now the well-posedness of this problem is ensured repeating the same passages showed in the previous section.

4.3   Numerical approximation

In this section we shall introduce briefly the classical Galerkin's method for the approximation of a weak formulation on Ω. Every weak formulation could be represented in a more compact and generic form as, find $u \in V$ such that

$$a(u, v) = F(v) \qquad \forall v \in V \tag{4.10}$$

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

$$V_h \subset V, \quad dim V_h < \infty \, \forall h > 0 \tag{4.11}$$

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

$$a(u_h, v_h) = F(v_h) \ \forall v_h \in V_h \tag{4.12}$$

Let be $\mathcal{T}_h$, a finite partition of Ω, 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 defined functions:

$$X_h^r(\Omega) := \{v_h \in C^0(\bar{\Omega}) : v_h|_K \in \mathbb{P}_r, \forall K \in \mathcal{T}_h\} \tag{4.13}$$

and the relative space where functions vanish on boundaries

$$X_{h,\Gamma_D}^r(\Omega) := \{v_h \in X_h^r : v_h|_{\Gamma_D} = 0\}. \tag{4.14}$$

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

$$dim \mathbb{P}_r := \frac{(r+1)^3}{2} + \frac{(r+1)^2}{2} + \frac{r(r+1)(2r+1)}{12} - \frac{r(r+1)^2}{2} - \frac{r(r+1)}{4} \tag{4.15}$$

More precisely we approximate $H^1_{\Gamma_D}(\Omega)$ with $X^1_{h,\Gamma_D}(\Omega)$ and $H^1_{\Gamma_{D,Si}}(\Omega_{Si})$ with $X^1_{h,\Gamma_{D,Si}}(\Omega_{Si})$. Therefore accordingly with (4.15) we have

$$dim\mathbb{P}_1 = 4$$
$$dimX_h^1 = N_h$$
$$dimX_{h,\Gamma_D}^1 = N_h - N_g$$

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 $j = 1, \ldots, N_h$ such that:

$$u_h = \sum_{j=1}^{N_h} u_j \psi_j \tag{4.16}$$

Because every functions of V_h is a linear combination of ψ_i, we can test equation (4.12) only for each basis function rather than $\forall v_h \in V_h$. The result of the complete discretization is find u_j, with $j = 1, \ldots, N_h$ such that:

$$\sum_{j=1}^{N_h} u_j a(\psi_j, \psi_i) = F(\psi_i) \qquad \forall i = 1, \ldots, N_h \tag{4.17}$$

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

$$\sum_{j=1}^{N_h} u_j \sum_{K \in \mathcal{T}_h} a_K(\psi_j, \psi_i) = \sum_{K \in \mathcal{T}_h} F_K(\psi_i) \qquad \forall i = 1, \ldots, N_h \tag{4.18}$$

4.3.1 Geometrical discretization

We set every elements $K \in \mathcal{T}_h$ as a tetrahedral of volume $|K|$; we suppose that there exists a constant $\delta > 0$ such that:

$$\frac{h_K}{\rho_K} \leq \delta \qquad \forall K \in \mathcal{T}_h \tag{4.19}$$

where $h_k = diam(K) = max_{x,y \in K}|x - y|$ and ρ_K is the diameter of the sphere inscribed in the tetrahedral K. Condition (4.19) is the so called *mesh regularity condition* [Qua08] and it ensures an istropic partition. We denote with $\mathcal{E}_h$, $\mathcal{V}_h$ and $\mathcal{F}_h$ the set of all the edges, verteces and faces of $\mathcal{T}_h$ respectively, and for each $K \in \mathcal{T}_h$ we denote by ∂K and $\mathbf{n}_{\partial K}$ the boundary of the element and its outward unit normal.

4.3.2   Linearized Non Linear Poisson equation

As regards the linearized NLP equation we have:

$$a(\psi_j, \psi_i) = \int_\Omega \epsilon \nabla \psi_j \nabla \psi_i \, d\Omega + \int_\Omega \sigma^k \psi_j \psi_i \, d\Omega \qquad (4.20)$$

and the relative restriction on the element is

$$a_K(\psi_j, \psi_i) = \int_K \epsilon \nabla \psi_j \nabla \psi_i \, dK + \int_K \sigma^k \psi_j \psi_i \, dK \qquad (4.21)$$

Equation (4.21) 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 ϵ is a piece-wise constant function, which changes on different material regions. $\mathcal{T}_h$ is built such that every K belongs to a single region, while it is possible that vertices belong to different regions. Therefore ϵ is constant over each elements and integral in (4.21) 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 we obtain the contributions of the local system matrix A_K^k:

$$[A_K^k]_{ij} = \epsilon_K L_{ij} + \frac{|K|}{4} \sigma_i^k \qquad (4.22)$$

having set

$$
\begin{aligned}
L_{ij} &\simeq \int_K \nabla \psi_i \nabla \psi_j \, d\Omega \\
\sigma_i^k &= \sigma^k(\mathbf{x}_i)
\end{aligned}
\qquad (4.23)
$$

Here follows the construction of the right hand side, based on the local contribution approximated with trapezoidal rule:

$$[F_K]_i^k = f_i^k |K|/4 \simeq \int_\Omega f^k \psi_i \, d\Omega \qquad (4.24)$$

The local contributions of each element K can 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 $\mathcal{J}_K : \mathcal{V}_{\mathcal{T}_h} \to \mathcal{V}_K$ the map which connects I to its corresponding local index $i = 1, \ldots, 4$ in the element K. Then we have

$$A_{IJ}^k = \sum_{\substack{\forall K \in \mathcal{T}_h \, s.t. \\ \mathcal{J}_K(I), \mathcal{J}_K(J) \subset \mathcal{V}_K}} [A_K]_{ij}^k \qquad (4.25)$$

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

$$b_I^k = \sum_{\substack{\forall K \in \mathcal{T}_h \, s.t. \\ \mathcal{J}_K(I) \subset \mathcal{V}_K}} [F_K]_i^k \qquad (4.26)$$

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:

$$\begin{cases} A^k \delta\varphi^k &= \mathbf{b}^k \\ \varphi^{k+1} &= \varphi^k + \delta\varphi^k \end{cases} \qquad (4.27)$$

As every iteration procedure, problem (4.27) needs a suitable convergence break criterion. A good method is based on checking the satisfaction of the fixed point equation (3.10) by the k-th solution, then the inner loop of the Gummel Map reads as: given a tolerance *toll* > 0 solve problem (4.27) untill:

$$||\mathbf{b}(\varphi^{k+1})||_2 > toll \qquad (4.28)$$

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

Damping

Nevertheless the theorem (3.1), the system (4.27) 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 termed 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:

$$\tilde{A}(\varphi_k) = \frac{1}{t_k}A(\varphi_k) \tag{4.29}$$

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 or s_k such that the modified Newton method exhibits superior convergence properties compared to the classical Newton method. For the case (4.29) there's a simple criterion suggested by Deuflhard [Deu74]. t_k is taken from the interval $(0,1]$ in such a manner that for any norm,

$$||A(\varphi_k)^{-1}\mathbf{b}(\varphi_k - t_kA(\varphi_k)^{-1}\mathbf{b}(\varphi_k))|| < ||A(\varphi_k)^{-1}\mathbf{b}(\varphi_k)|| \tag{4.30}$$

Condition (4.30) 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 descents. This condition can be easily evaluated only if the Jacobian matrix is factored into triangular matrices because the evaluation of the argument of the norm on the left hand side of (4.30) is then 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 $D(\varphi_k)$:

$$||D(\varphi_k)^{-1}\mathbf{b}(\varphi_k - t_kD(\varphi_k)^{-1}\mathbf{b}(\varphi_k))|| < ||D(\varphi_k)^{-1}\mathbf{b}(\varphi_k)|| \tag{4.31}$$

This is the criterion developed in our code. However the value to use for t_k is a question of trial and error. Frequently one chooses the following sequences:
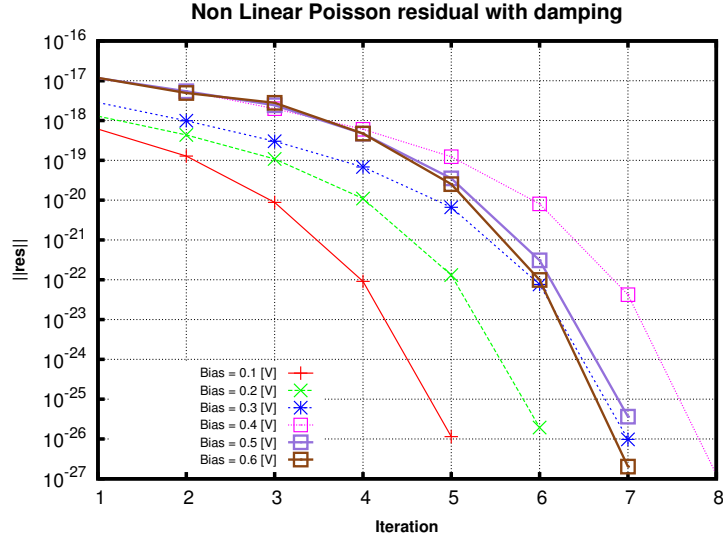
$$t_k = \frac{1}{2^i} \tag{4.32}$$

$$t_k = \frac{1}{2^{\frac{i(i+1)}{2}}} \tag{4.33}$$

obiuvsly i is the subiterations of damping reached when satisfied (4.31). Sufficiently close to the solution (4.30) (and so (4.31)) will be satisfied with $t_k = 1$ so that the convergence properties of the classical Newton method are recovered.
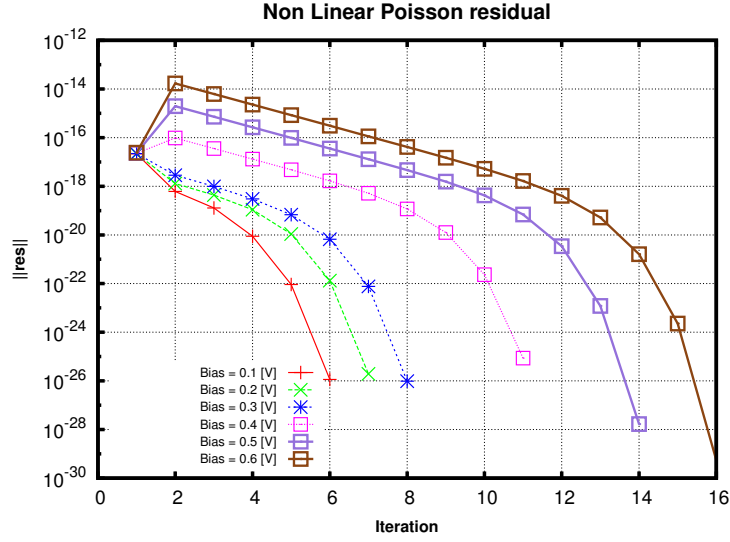
Figure 4.1: Number of iteration against reasidual for different bias in a diode case test. (a) NLP solving without damping, (b) NLP solving with Deuflhard damping criterion.
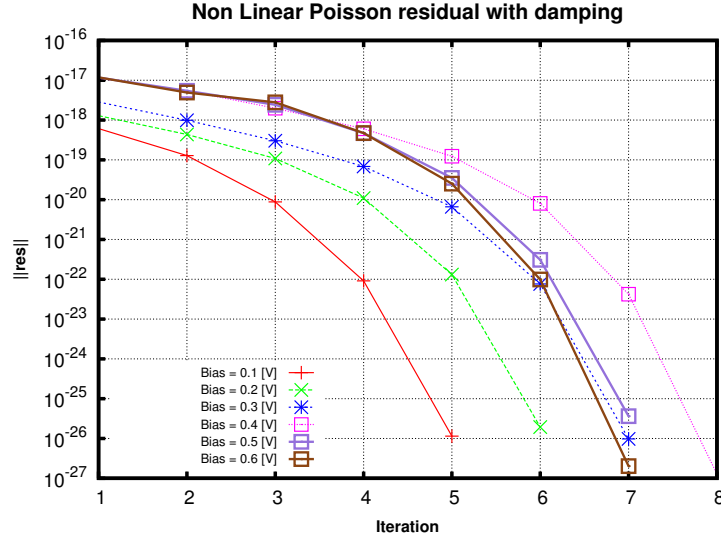
Figure 4.2: Behaviour of the t_k coefficients.

4.3.3   Continuity equations

As regards the LEC equation we have:

$$a(u,v) = \int_{\Omega_{Si}} qD_n e^{(\varphi^i/V_{th})} \nabla\psi_j \nabla\psi_i \, d\Omega + \int_{\Omega_{Si}} \sigma_n^{i-1} e^{(\varphi^i/V_{th})} \psi_j \psi_i \, d\Omega \quad (4.34)$$

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 φ.

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:

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

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

The weak form (4.34) 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 [DF05]. First of all it's convenient to reformulate problem (3.34) by using the relation (2.32) and (2.34) in a more generic form. This yields the following equivalent form of the strong ADR model problem

$$
\begin{cases}
\nabla \cdot (\mathbf{J}_n(n)) + \sigma n = f & in \ \Omega_{Si} \\
\mathbf{J}_n = q D_n e^{(\varphi/V_{th})} \nabla (e^{(-\varphi/V_{th})} n) & in \ \Omega_{Si} \\
n = n_D & on \ \Gamma_{D,Si} \\
\mathbf{J}_n \cdot \mathbf{n} = 0 & on \ \Gamma_{N,Si}
\end{cases}
\tag{4.35}
$$

We report here the weak formulation of (4.35) which is well investigated in [DF05], find $\mathbf{J}_n \in [L^2(\Omega)]^d$ and $n \in H^1_{\Gamma_{D,Si}}(\Omega)$ such that

$$
- \int_{\Omega_{Si}} \mathbf{J}_n \cdot \nabla v \, d\Omega + \int_{\Omega_{Si}} \sigma n v \, d\Omega = \int_{\Omega_{Si}} f v \, d\Omega \qquad \forall v \in H^1_{\Gamma_{D,Si}}(\Omega)
\tag{4.36}
$$

$$
\int_{\Omega_{Si}} (q D_n e^{(\varphi/V_{th})})^{-1} \mathbf{J}_n \cdot \mathbf{q} \, d\Omega + \int_{\Omega_{Si}} \nabla(e^{(\varphi/V_{th})n}) \cdot \mathbf{q} \, d\Omega = 0 \qquad \forall \mathbf{q} \in [L^2(\Omega)]^d
\tag{4.37}
$$

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

$$
\Sigma_h := \left\{ \mathbf{q}_h \in [L^2(\Omega)]^d : \mathbf{q}|_K \in [\mathbb{P}_0]^d \forall K \in \mathcal{T}_h \right\}
\tag{4.38}
$$

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

$$
\mathbf{q}^h_{1,2,3} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}
\tag{4.39}
$$

Therefore we can easily deduced the restriction on a generic element of (4.36) and (4.37) and write the relative discretized bilinear form

$$
\begin{cases}
a_h^K(n_h, v_h) &= \int_K \mathbf{J}_{n,h}^K(n_h) \nabla v_h \, dK + \int_K \sigma n_h v_h \, dK \\
\\
F(v_h)^K &= \int_K f v_h \, dK \\
\\
\mathbf{J}_{n,h}^K &= D_K (q D_n e^{(\varphi/V_{th})}) \nabla (e^{(-\varphi/V_{th})} n_h)
\end{cases}
\tag{4.40}
$$

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

$$D_K(qD_n e^{(\varphi/V_{th})}) = \begin{cases} \mathcal{M}_K(qD_n e^{(\varphi/V_{th})})\,\mathcal{I} \\[2mm] \mathcal{H}_K(qD_n e^{(\varphi/V_{th})})\,\mathcal{I} \\[2mm] \dfrac{1}{|K|}\sum_{i=1}^{6}\mathcal{H}_{e_i}(qD_n e^{(\varphi/V_{th})})|e_i|s_i\mathbf{t}_i\mathbf{t}_i^T \end{cases} \tag{4.41}$$

These different approaches in the computation of the average of the diffusion coefficient are responsible for the quite different numerical perfomance of the relative methods. We already presented the standard average and the armonic average and we discussed briefly the advantages of them. Now we are interested in the latter method which is an exponentially fitted treatment over the entire subdomain K. From (4.40) we immeditaly obtain:

$$\mathbf{J}_h^K = \frac{1}{|K|}\sum_{i=1}^{6}|e_i|s_i j_{e_i}\mathbf{t}_i \tag{4.42}$$

Having defined the flux vector over K in the form (4.42), it is clearly possible to construct a fimily of Galerkin finite element approximations for the continuity equations by a proper choice of the quantities j_{e_i}.

Edge flux density computation

We introduce an approximate flux density computation along the edges of ∂K based on the classical one dimensional *Sharfetter-Gummel* method. For each edge e_i, the tangential component of $J_h^K(n_h)$ is defined as:

$$j_{e_i} \simeq \mathcal{H}_{e_i}(qD_n e^{(\varphi/V_{th})})\frac{\mathcal{B}(\delta_i(\varphi/V_{th}))n_{h,k} - \mathcal{B}(-\delta_i(\varphi/V_{th}))n_{h,j}}{|e_i|} \tag{4.43}$$

where

$$\delta_i(\varphi/V_{th}) = \frac{\varphi_k - \varphi_j}{V_{th}} = 2\frac{(\mathbf{b}_K \cdot \mathbf{t}_{e_i})|e_i|}{2\mathcal{H}_{e_i}(qD_n e^{(\varphi/V_{th})})} = 2\gamma_i \tag{4.44}$$

$$\mathcal{B}(z) = \begin{cases} \dfrac{z}{e^z - 1} & z \neq 0 \\[2mm] 1 & z = 0 \end{cases} \tag{4.45}$$

being $\mathbf{b}_K$ the relative drift component of the problem on K and $|\gamma_i|$ the Pèclet number associated with the edge e_i.

The discretization scheme

Given the choice for j_{e_i} and substituting the equation for $\mathbf{J}_h^K$ in the bilinear form of (4.40), we can compute the local system matrix over each element which reads as follows:

$$\Phi_K = \begin{bmatrix} -\begin{pmatrix} a_{e12}\mathcal{B}_{12}L_{21}^K+ \\ a_{e13}\mathcal{B}_{13}L_{31}^K+ \\ a_{e14}\mathcal{B}_{14}L_{41}^K \end{pmatrix} & a_{e12}\mathcal{B}_{12}L_{21}^K & a_{e13}\mathcal{B}_{13}L_{31}^K & a_{e14}\mathcal{B}_{14}L_{41}^K \\ a_{e21}\mathcal{B}_{21}L_{12}^K & -\begin{pmatrix} a_{e21}\mathcal{B}_{21}L_{12}^K+ \\ a_{e23}\mathcal{B}_{23}L_{32}^K+ \\ a_{e24}\mathcal{B}_{24}L_{42}^K \end{pmatrix} & a_{e23}\mathcal{B}_{12}L_{32}^K & a_{e24}\mathcal{B}_{12}L_{42}^K \\ a_{e31}\mathcal{B}_{31}L_{31}^K & a_{e31}\mathcal{B}_{32}L_{32}^K & -\begin{pmatrix} a_{e31}\mathcal{B}_{31}L_{31}^K+ \\ a_{e32}\mathcal{B}_{32}L_{32}^K+ \\ a_{e34}\mathcal{B}_{34}L_{34}^K \end{pmatrix} & a_{e34}\mathcal{B}_{34}L_{34}^K \\ a_{e41}\mathcal{B}_{41}L_{41}^K & a_{e42}\mathcal{B}_{42}L_{42}^K & a_{e43}\mathcal{B}_{43}L_{43}^K & -\begin{pmatrix} a_{e41}\mathcal{B}_{41}L_{41}^K+ \\ a_{e42}\mathcal{B}_{42}L_{42}^K+ \\ a_{e43}\mathcal{B}_{43}L_{43}^K \end{pmatrix} \end{bmatrix} \tag{4.46}$$

$$A_K = \Phi_K + \frac{|K|}{4} diag(\sigma) \tag{4.47}$$

$$\mathbf{F}_K = \frac{|K|}{4}(f_1, f_2, f_3, f_4)^T \tag{4.48}$$

For the assembling of the system matrix and the imposition of the essential boundary conditions we refer to the previous section.

Vale la seguente uguaglianza?

$$L_{ij} = s_{ij}\mathbf{t}_{ij} \cdot \nabla \psi_i \tag{4.49}$$

Chapter 5

Simulation results

In this chapter we shall present a collection of the most interesting results obtained to validate the numerical performance of discretization method described in the previously chapters.

5.1 Test cases

We focus our tests over three kind of semiconductor devices:

- **p-n junction**

- **p-n junction in oxide**

- **MOS n-channel**

For everyone of these case test we shall investigate the correction of the FEMOS solution against the results obtained with SDEVICE.

5.1.1 p-n junction

In this example we consider a p-n junction also known as diode. In Fig.5.1 we present the partition used and the doping profile of this test case. 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 verticises used is 4933. The doping concentration is obtained setting a constant profile of acceptor over all the domain ($N_A = 1.0 \times 10^{17}$) overhelmed by a doping profile of acceptor ($N_D = 1.0 \times 10^{18}$) bounded on one side of the device. As you can see we defined the two different areas of doping with an almost abrupt junction. Two contacts are defined, then on the first (A) is applied a voltage of $0.3[V]$ while the

(a) *Mesh*   (b) *Doping concentration* $[cm^{-3}]$ (c)
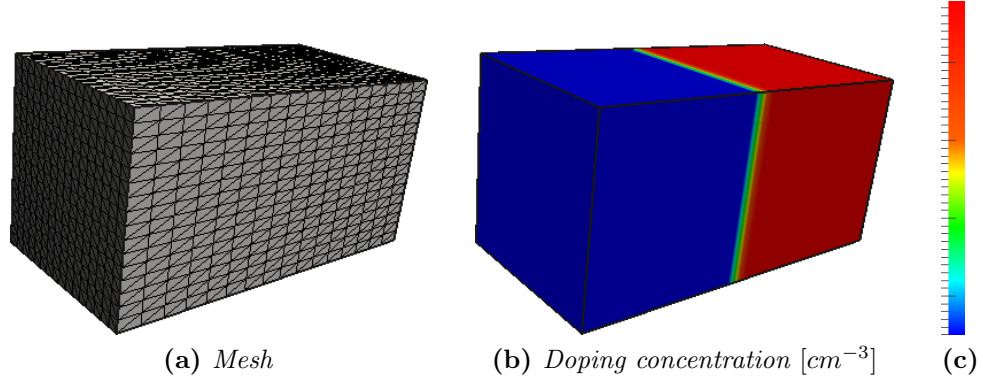
Figure 5.1: Test case diode p-n.

opposite one (B) is grounded. In this situation the diode is direct polarized. The relative electric permettivity of the silicon is $\epsilon_{SI} = 11.6[Fcm^{-1}]$ while the mobility model used is the thermal model (2.56) with a constant temperature of $300[K]$.

In 5.4, 5.5 and 5.6 we validated FEMOS results against SDEVICE. As we aspected from theory the drop in potential is almost bounded around the junction, and as we use an asymmetric doping, it is major extended on the p-side. Notice that on the minority contact, which means A contact for electron solution and B contact for hole solution, the Dirichlet condition causes a big drop in concetration (from 1.0×10^{15} to 1.0×10^3 $[cm^{-3}]$) qua introduco questo commento per legarmi poi ai problemi che abbiamo riscontrato nel conto della corrente.

Now we are interested to analyze the computational performance of this numerical scheme. We know that the main difficulty is the initialization of the variables and obiouvsly the convergence time is strictly related to the kind of initial step we use. However predict in every situations the possibly shape of the solutions is hard, if not even impossible. For this reason a common and general approach is used: we split the domain in several regions accordingly to their doping concentration. Once obtained this partition we treat every of those semiconductor regions as they are in equilibrium with the nearest contact. This hypotesis allow us to compute over the entire domain the initial step for φ, n and p thanks to the relations (2.18) and (2.19). Even this choice of initial step some difficulties may occur, especially if we consider the device far away the equilibrium condition. We can state this effect looking at the figure Fig.5.2, where computational costs are shown against the polarization. We notice that at $1.0[V]$ the computational time increase really quickly, and then it remains stable. This voltage level is almost the $built-in$ of this p-n junction, which means ....
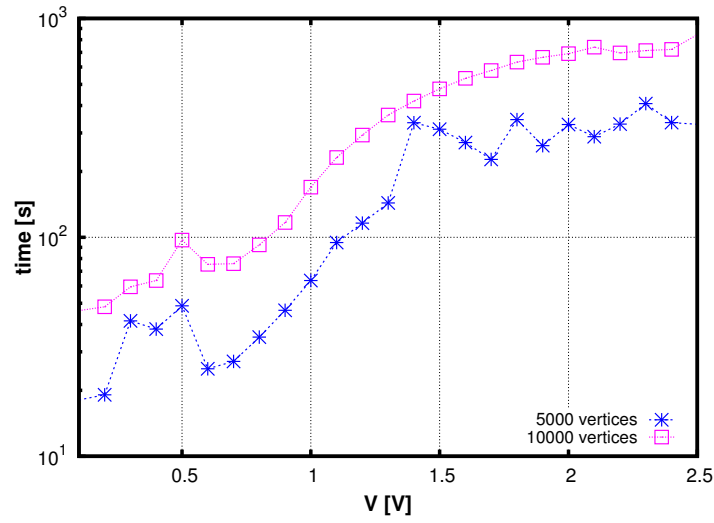
Figure 5.2: Computational times for different bias in two mesh cases.

The graph Fig.5.3 shows us that this problem is related to the convergence of the Gummel Map and not to the shape of the system matrix of the equations. Indeed while the number of iterations increase like the curve of total time computation, the times refer to the computational time of NLP and DD solver equation remain stable.
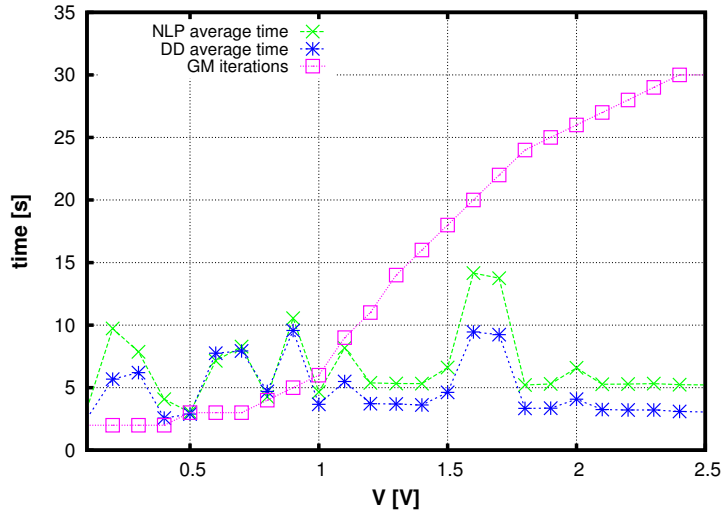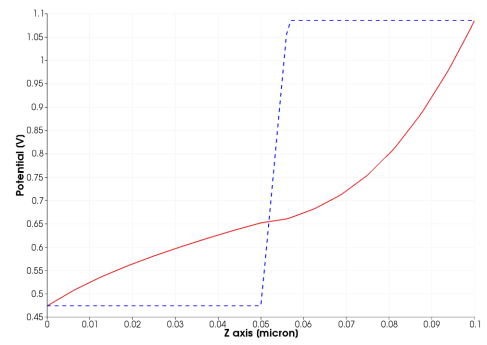

Figure 5.3: Computational times and iterations.

We can see that the shape of the solutions after the buil-in level are really different from the initial step referenza figura plot over line del potenziale.
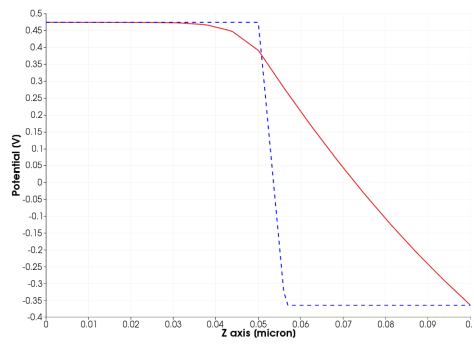
(a)   (b)

So if we are far away from the equilbrium ... Il sistema e fuori dall'equilibrio le soluzioni differiscono molto dalla soluzione iniziale

(a) *FEMOS*   **(b)**   **(c)** *SDEVICE*
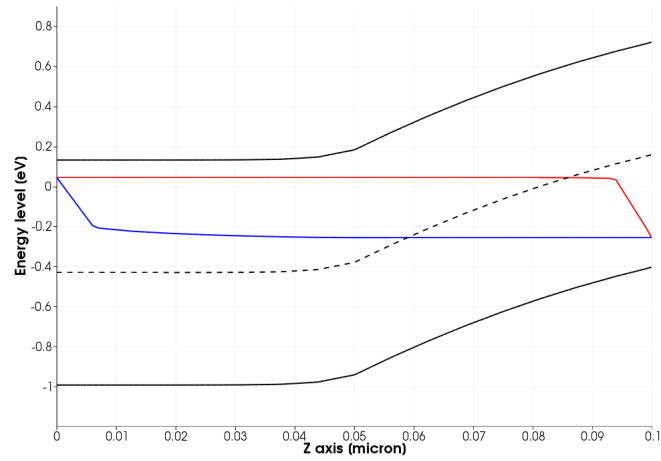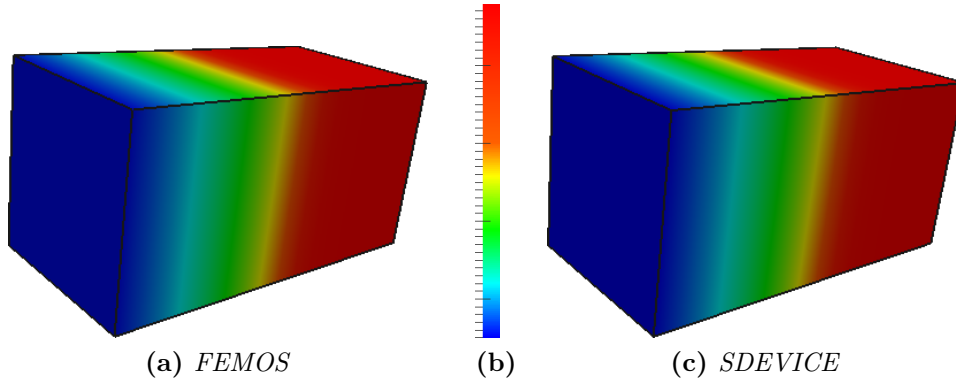
Figure 5.4: Test case dide p-n 0.3[V] potential.

(a) *FEMOS*   **(b)**   **(c)** *SDEVICE*
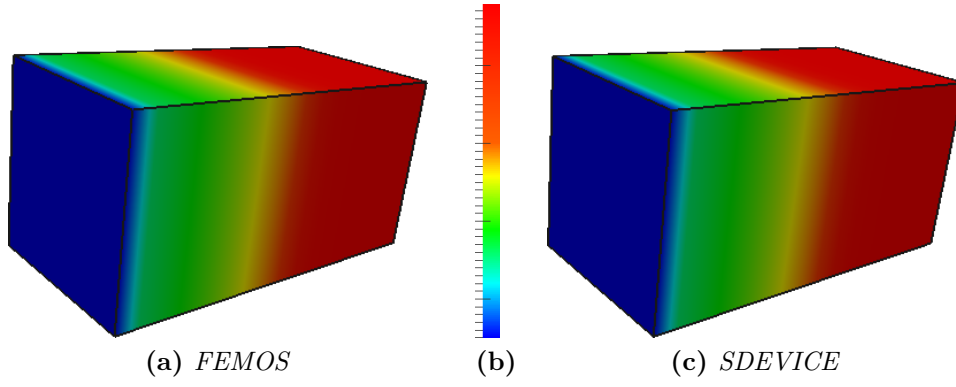
Figure 5.5: Test case dide p-n 0.3[V] electron density.

(a) *FEMOS*   **(b)**   **(c)** *SDEVICE*
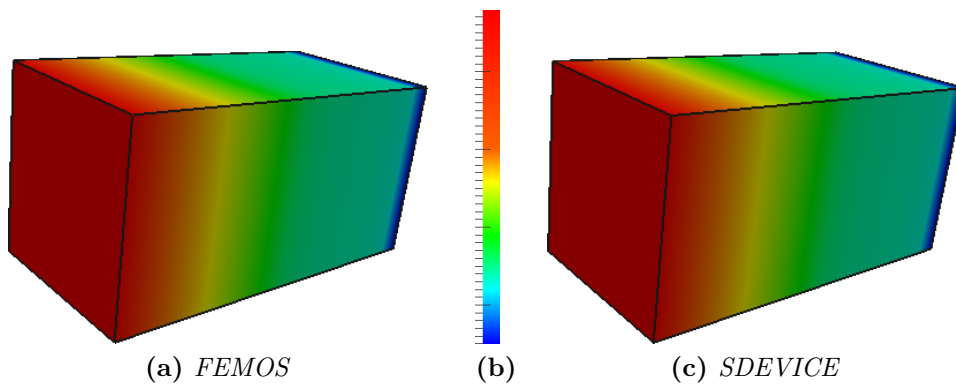
Figure 5.6: Test case dide p-n 0.3[V] hole density.

5.1.2   p-n junction in oxide

In this test case we surrounded the previous structure with a layer of oxide material ($\epsilon_{ox} = 3.9[Fcm^{-1}]$). This layer is about $0.025[\mu m]$ thick and notice that the contacts are defined only on the silicon region. As the physical behaviour in the oxide is much less complicated than in the semiconductor part, we chose to use a smaller step of discretization in this zone. We spent 6334 verticies overall the domain. The structure and the doping are well defined in Fig.5.7. The setting of the electrode is the same of the previously simulation. We still validate our solutions with the commercial software in Fig.5.9, Fig.5.10 and Fig.5.11.

(a) *Mesh*   **(b)** *Doping concentration*
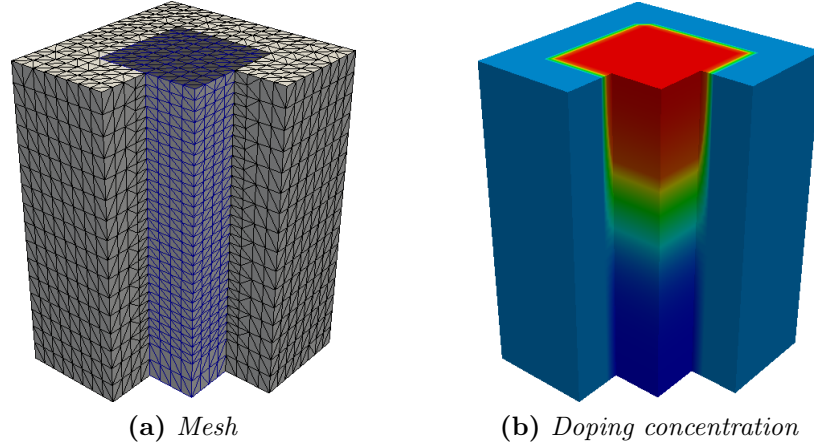
Figure 5.7: Test case p-n junction in oxide.

It's interesting notice the behaviour of the potential inside the oxide layer, indeed it seems to be more diffused than in the silicon. This phenomena is clear if we think about the relation (2.29) between electric field and potential. As we imposed $\nabla\varphi \cdot \mathbf{n} = 0$ on the oxide boundary no field lines of the electric field could cross that boundary. The consequence is that all the field lines could start and end only from the contact A or B. Therefore the behaviour of the the electric field inside the device is due only to the displacement effect on the junction of the silicon, which imposes the electric response in the oxide material, see Fig.5.8. The conseguence is a lower electric field in the oxide which implies a more diffused potential. We remind that the electric field variable is post-processed by the electric potential, this is a conseguence of adopting a displacement formulation approach.
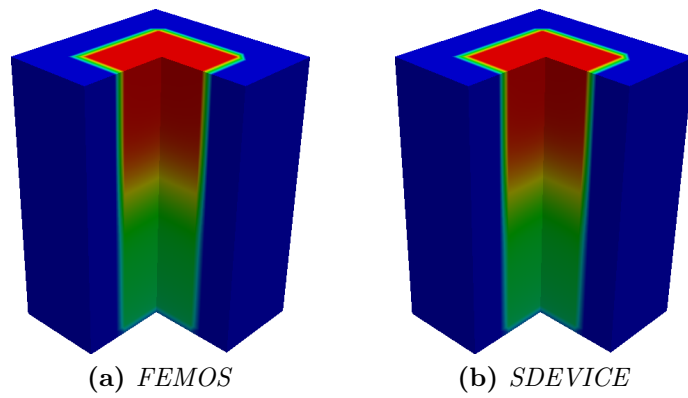
(a) *FEMOS* **(b)** *SDEVICE*

Figure 5.8: Test case dide p-n in ox 0.3[V], electric field.

(a)

(b)

(c)

(a) *FEMOS*   (b) *SDEVICE*
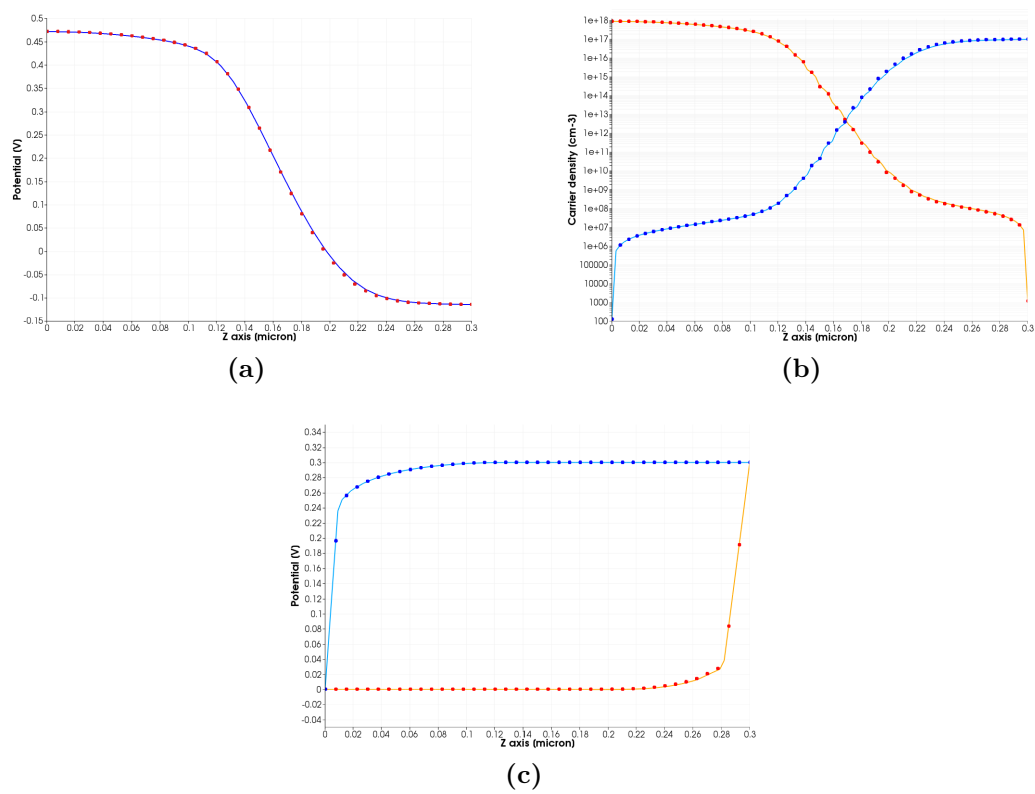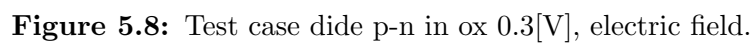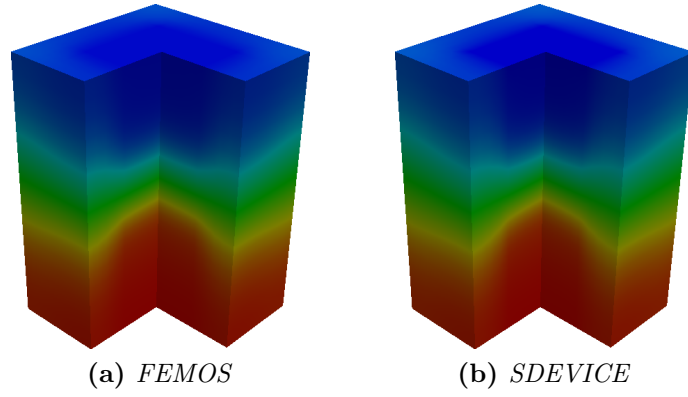
Figure 5.9: Test case dide p-n in ox 0.3[V].

(a) *FEMOS*   (b) *SDEVICE*
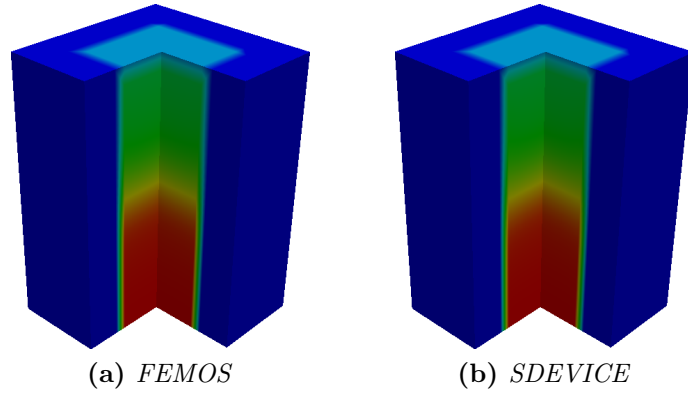
Figure 5.10: Test case dide p-n in ox 0.3[V].


(a) *FEMOS*   (b) *SDEVICE*

**Figure 5.11:** Test case dide p-n in ox 0.3[V].

5.1.3 MOSFET n-channel

The functionality of a *MOSFET* device is quite complex, therefore we give a summary of its basic operation. For a more detailed description see referenza per mos. The metal-oxide-semiconductor field-effect transistor (MOSFET) is the building block of VLSI circuits in microprocessors and dynamic memories. Because the current in a MOSFET is transported predominantly by carriers of one polarity, the MOSFET is usually referred to as a unipolar or majority-carrier device. In this test case we consider an *n-channel MOSFET* or *nMOSFET* which consists of a p-type silicon substrate into which two n-regions are formed. Its structure is well defined in Fig.5.12. 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 the *channel* region. We describe briefly the basic operation of a MOSFET device. When there is no voltage applied to the gate there is no current flow between the source and drain, while if a sufficiently large positive voltage is applied to the gate, the silicon surface is inverted to n-type, which forms a conducting channel between the source and drain. In the latter case if we apply a voltage to the drain the electrons start to flow from source to drain and therefore a current in the opposite direction is generated. As we know already where the phenomena shall be more interesting we performed an oriented meshing which is high fitted along the channel region, at the contact and over the deplation region Fig.5.12. The n-regions are implanted accordingly with a gaussian profile.

(a) *Mesh*   (b) *Doping concentration*   (c)
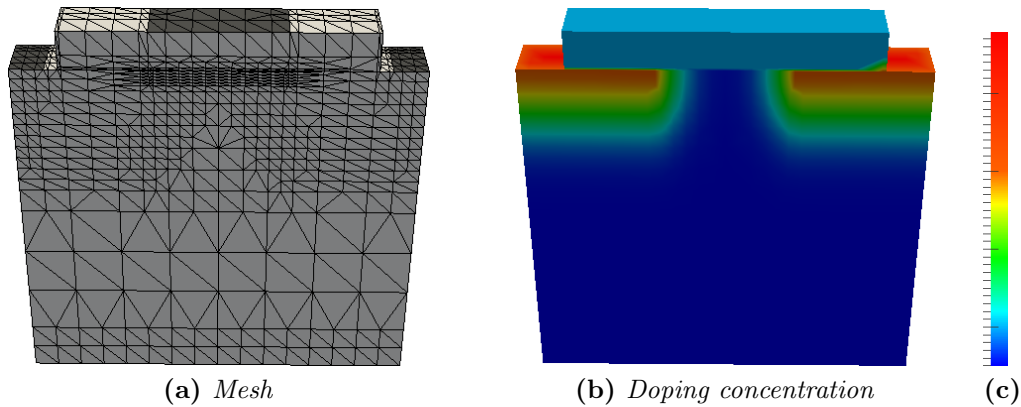
Figure 5.12: Geometry of the test case MOS n-channel.

We match our result in the operating phase of the MOSFET, therefore the setting of the electrode is $V_G = 2.0[V]$, $V_D = 0.1[V]$ and $V_S = V_B = 0.0[V]$. The results are shown in 5.14, Fig.5.15 and Fig.5.16. Whereas the channel region is a p-region, with this polarization its nature is changed to an n-region, for better understand the formation of the channel see ref figura equilibrio which represents the electron and hole density of the same device for a $V_G = 0.0[V]$.
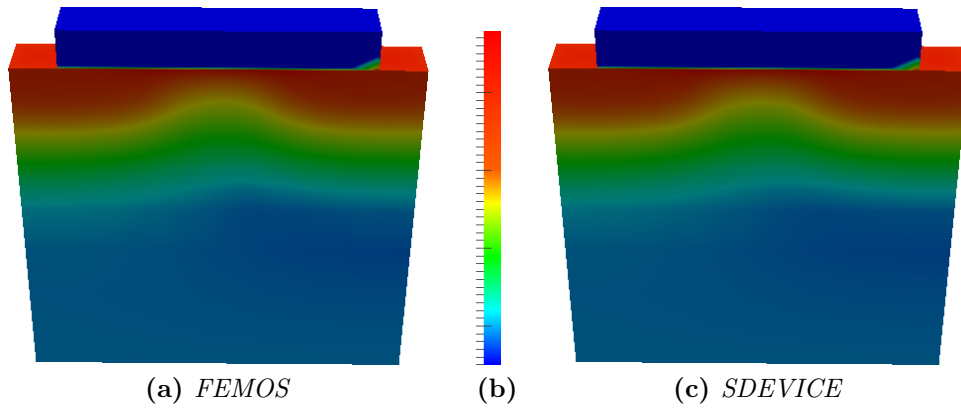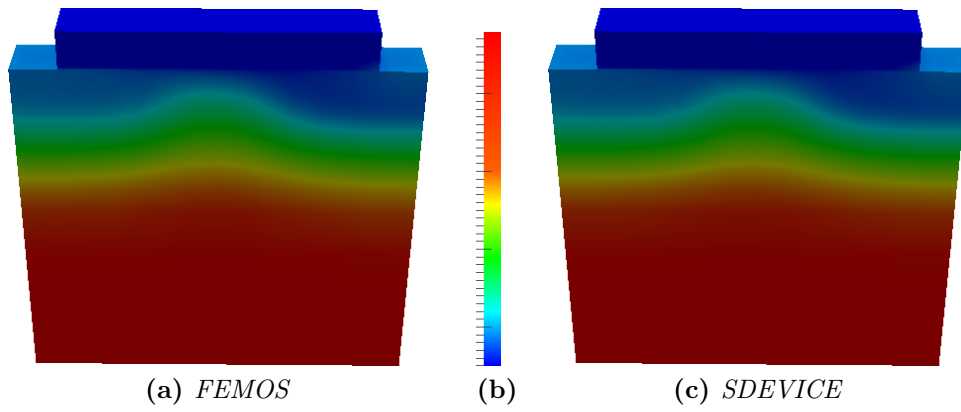

(a)

(b)
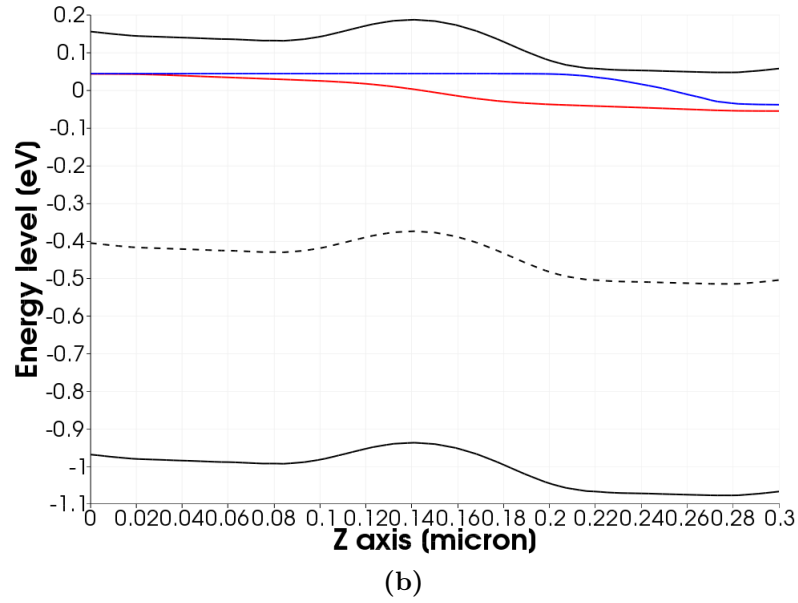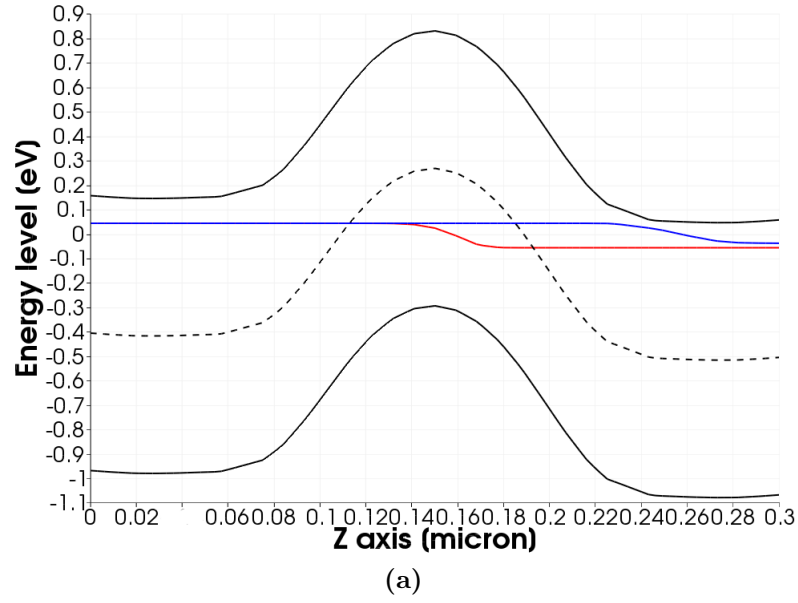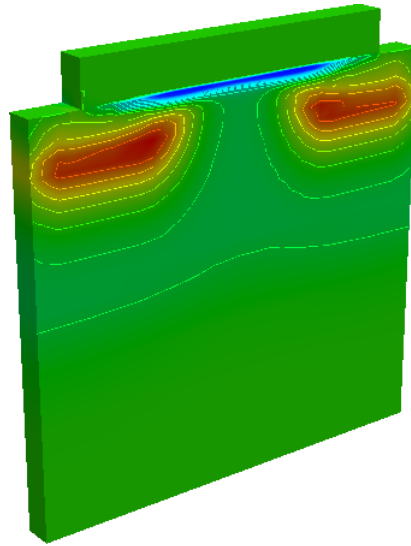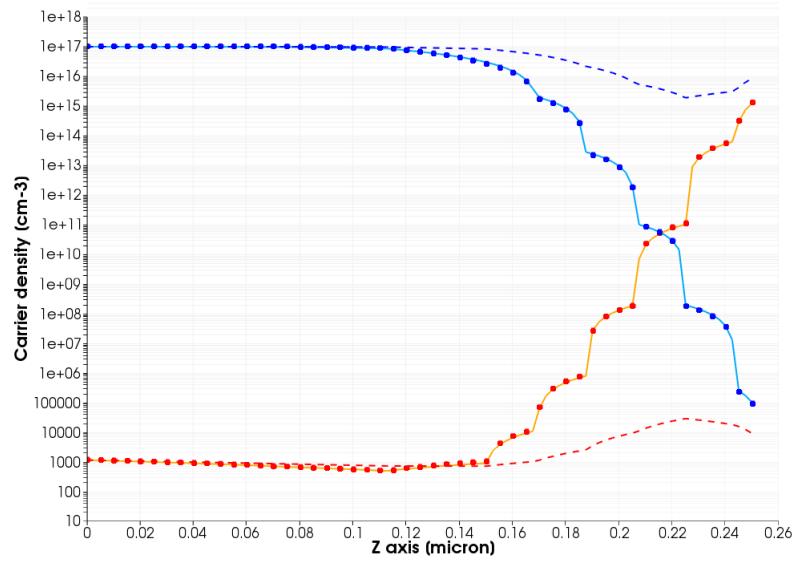
Figure 5.13: Space charge $V_G = 2.0[V]$

(a) *FEMOS*   (b)   (c) *SDEVICE*
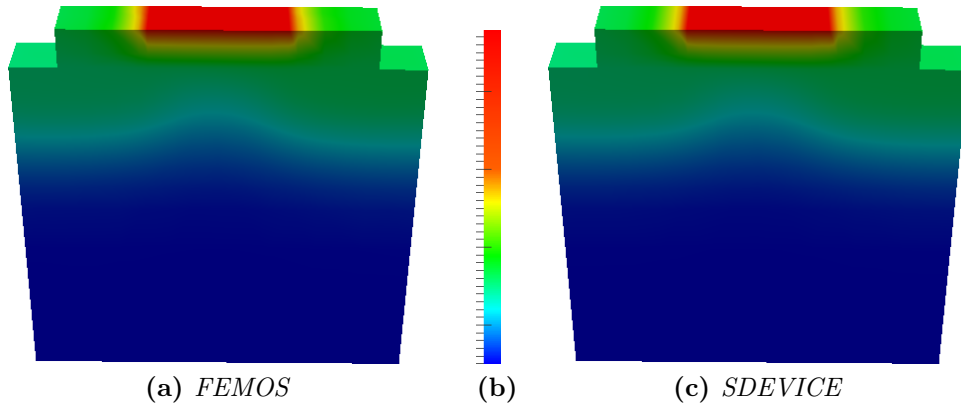
Figure 5.14: Potential Vgate = 2.0 [V].

(a) *FEMOS*   (b)   (c) *SDEVICE*

Figure 5.15: Electron density Vgate = 2.0 [V].

(a) *FEMOS*   (b)   (c) *SDEVICE*

Figure 5.16: Electron density Vgate = 2.0 [V].

Show how decade the GCA the gradual channel approzimation which assumes that the variation of the electric field along the channel is much less than the corresponding variation in the orthogonal direction (perpendicular to the channel).

(a) *FEMOS*   **(b)**   **(c)** *SDEVICE*
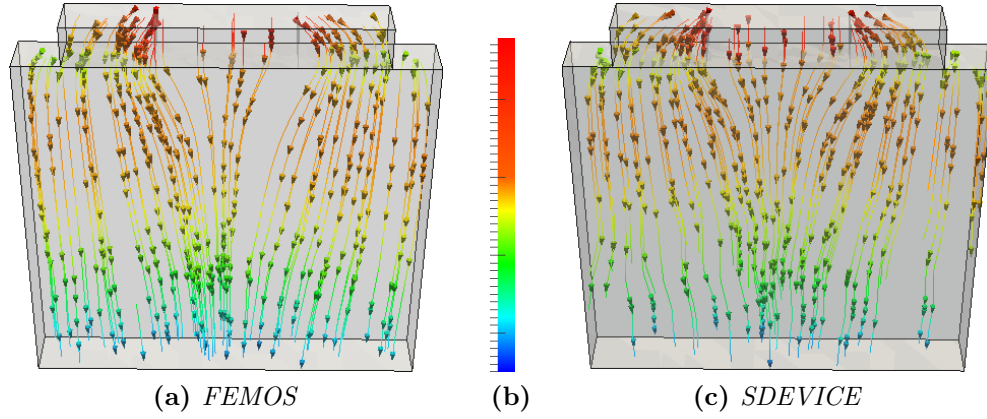
Figure 5.17: Electric field density Vgate = 2.0 [V].

5.2   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 electrode. Consider $\Gamma_{D,Si} = \bigcup_{i=1}^{d} \Gamma_i$ where d is the number of terminals on the device and $\forall i = 1, ..., d$ Γ_i is the i-th contact. The problem is $\forall i$ compute:

$$\mathcal{I}_i = \mathcal{I}_i^n + \mathcal{I}_i^p \qquad (5.1)$$

where $\mathcal{I}_i$ is the total current, $\mathcal{I}_i^n$ is the contribution of the electron current and $\mathcal{I}_i^p$ is the contribution of the hole current. In general, given a contact Γ_i, fluxes of current density to be calculated assume the following form:

$$\mathcal{I}_i^\nu = \int_{\Gamma_i} \mathbf{J}_\nu(\nu) \cdot \mathbf{n} \, d\Gamma_i \qquad \nu = \{n, p\} \qquad (5.2)$$

where as usual $\mathbf{n}$ is the unit outward normal of the domain boundary. It's well known that the evaluation of boundary integrals is a difficult task. Many difficulties in the numerical evalutaion of (5.2) 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 (e.g. from Dirichlet to Neumann) at the contact ends.

In the following we present an accurate method for the evaluation of boundary integrals in semiconductor device based on the work [GS74], we shall extend the *residue method* to the 3D case and we confirm the optimal results matching them with SDEVICE. Moreover we remark that the method can be succesfully applied to a wide spread of applications, including contact charges, carrier quantum probability fluxes and heat fluxes.

Before go further with the presentation of the results, it's useful look up the anlysis made in [GS74] and adapt it to our case. Let be A the system matrix of the continuity equation (4.47) and $\mathbf{b}$ the relative right hand side before applying boundary conditions. The linear problem looks as:

$$\sum_{j \in \eta} A_{ij} \nu_j = b_i \qquad \forall i \in \eta \qquad (5.3)$$

where η is the set of all vertices of the partition $\mathcal{T}_h$, ν_j is the generic j-th component of the carrier solution. We can split the set of total nodes in contact node η_g and the complementary part η_n. Notice that the values of ν_j are known on the contacts and (5.3) can be rewritten as follows:

$$\begin{cases} \sum_{j\in\eta_n} A_{ij}\nu_j = b_i - \sum_{j\in\eta_g} A_{ij}\nu_j & \forall i \in \eta_n \\[2em] \sum_{j\in\eta_n} A_{ij}\nu_j = b_i - \sum_{j\in\eta_g} A_{ij}\nu_j & \forall i \in \eta_g \end{cases} \tag{5.4}$$

The first set of equations is the usual linear system which we solve while the second can be used for boundary flux estimation. Now we define a new test function v_i^h as:

$$v_i^h = \sum_{j\in\eta_{gi}} \psi_j \tag{5.5}$$

where η_{gi} is the set of nodes lying on Γ_i. Substituting (5.5) in (5.2) we can estate that:

$$\begin{aligned} \mathcal{I}_i^\nu &= \int_{\Gamma_i} \mathbf{J}_\nu(\nu)\cdot\mathbf{n}\,d\Gamma_i = \sum_{j=1}^{n_d}\int_{\Gamma_j} \mathbf{J}_\nu(\nu)\cdot\mathbf{n}\,v_i^h\,d\Gamma_j \\ &= \sum_{i\in\eta_{gi}}\sum_{j=1}^{n_d}\int_{\Gamma_j} \mathbf{J}_\nu(\nu)\cdot\mathbf{n}\,\psi_i\,d\Gamma_j = \sum_{i\in\eta_{gi}}\int_{\partial\Omega} \mathbf{J}_\nu(\nu)\cdot\mathbf{n}\,\psi_i\,d\partial\Omega \\ &= \sum_{i\in\eta_{gi}}\left[\int_\Omega \nabla\cdot\mathbf{J}_\nu(\nu)\psi_i\,d\Omega + \int_\Omega \mathbf{J}_\nu(\nu)\cdot\nabla\psi_i\,d\Omega\right] \\ &= \sum_{m\in\eta_{gi}}\left[\sum_{j\in\eta} A_{ij}\nu_j\psi_i - b_i\right] \end{aligned} \tag{5.6}$$

The current at contact i is performed by summing the residuals of the matrix A.

Per questo detto metodo dei residui ...

Nel seguito mostriamo alcuni risultati ...

5.2.1   Results

We shall apply the residue method over all the devices presented before in specific condition of function and different modelization of the mobility.

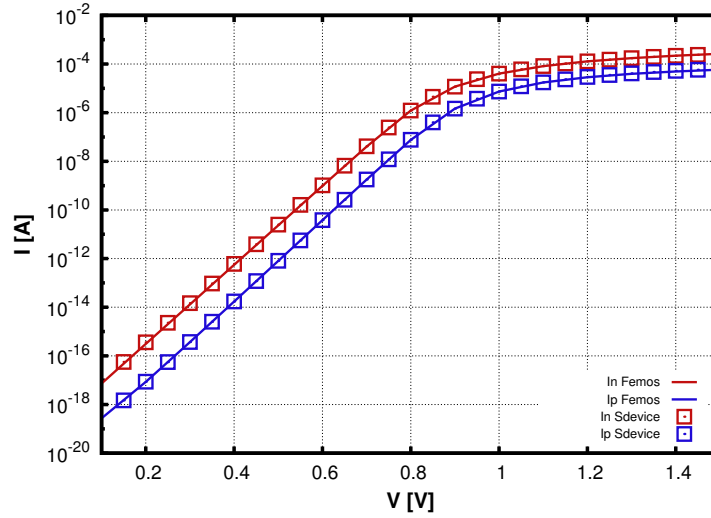**p-n junction**

p-n junction in oxide

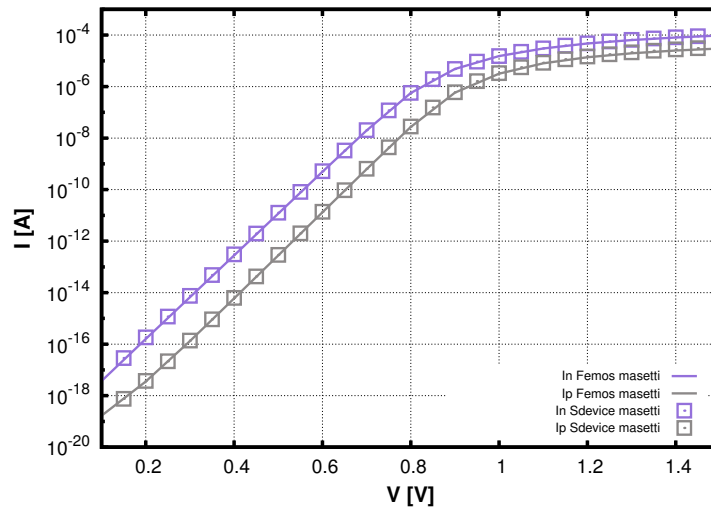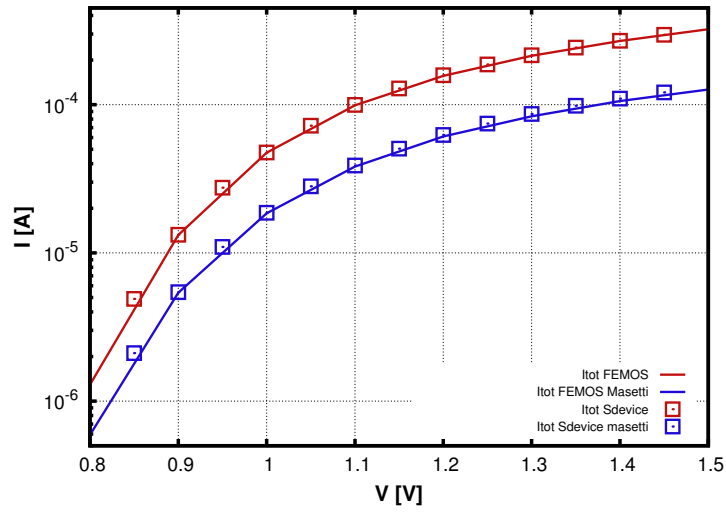**Figure 5.18:** Direct polarization constant mobility.

Figure 5.19: Direct polarization Masetti mobility.

Commenti sull'abbassamento della corrente di contatto dovuto all'abbassamento della mobilit

Polarizzazione diretta mobilit canali (??) viene qualitativamente non è allineato con sedevice

Figure 5.20: Masetti mobility effect Vs. constant mobility effect.

MOS n-channel
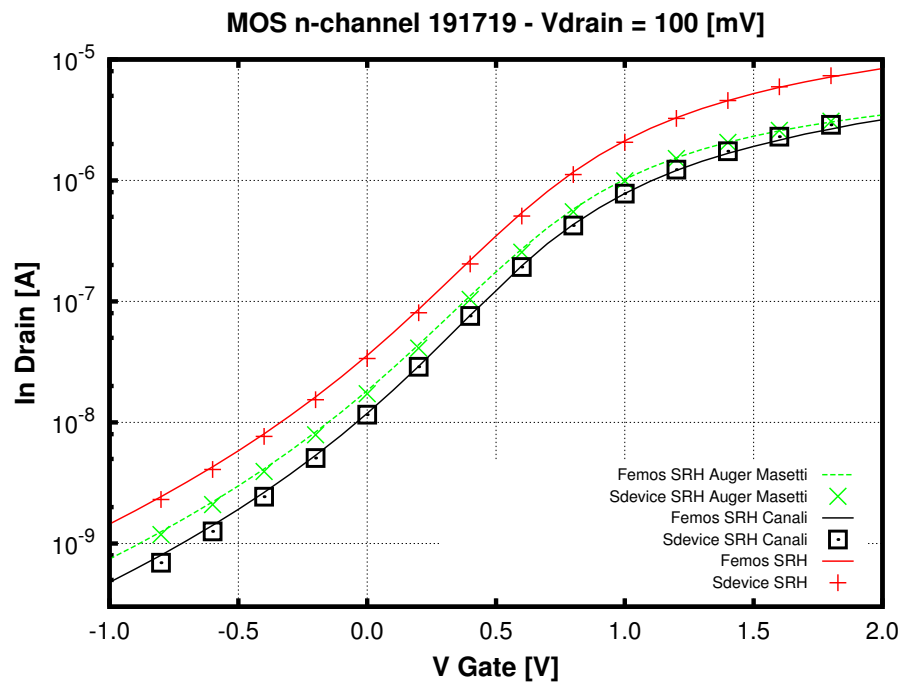
Polarizzazion diretta (OK)

    Polarizzazione diretta masetti

    Polarizzazione inversa

    Stress region contributio impact ionization

    Mos acceso mobilit con Velocit di saturazione (??)

MOS n-channel 191719 - Vdrain = 100 [mV]

Chapter 6

The current calculation problem

In many physical and engineering problems the real interesting variable of the conservation law is the flux in the domain or on specific surfaces and boundaries. The study of micro and nano electronics devices doesn't except this observation, in fact most of all models are oriented to obtain a satisfactory description of the current density. We know that the primal and not mixed formulation for the continuity equation doesn't resolved the flux density. The conseguence of this fact is a binding post-processing of the quantities computed in order to reconstruct the current density of electrons and holes. It's evident which this part covers a lead role in the device simulation: as we are satisfied of the impressive results of the finite element scheme, it will be reather regrettable to lost the accuracy of our simulation during the computation of the current density. About this question many academics propose different solutions and the relative literature is boundless. Nevertheless the problem shows various aspects to take into account, among these there are some which every good method should be respect:

- reduced computational cost;

- easy extension to 3D simulation;

- detains some useful properties like orthogonal conservation across a generic surface of the domain;

- preserve consistency with the numerical scheme adopted.

It's not trivial ensure everyone of these points, thus move on toward a unique choice of a method is a delicate matter. Luckily there's some *main stone* which offers ever a good start point whence achieve new results. Probably

the most known and recognized by the inherent literature is the *Sharfetter-Gummel formula.*
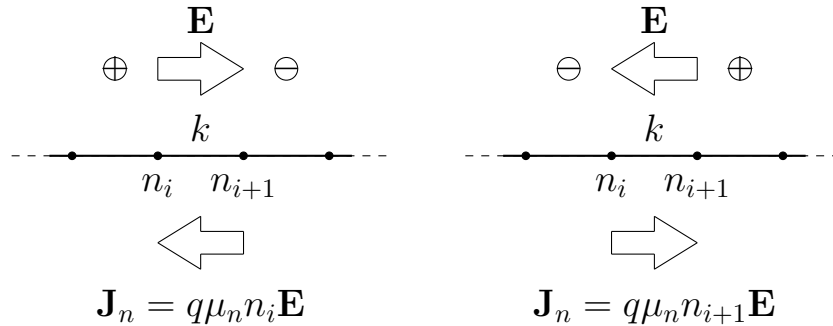
6.1   Scharfetter-Gummel formula

Consider the resolution of the continuity equation along a monodimensional domain. For the sake of simiplicity we contemplate a uniform partition (this hypotesis is not necessary for a more generic analysis). Moreover on every nodes is defined the electrostatic potential φ, and on every elements the relative electrostatic field $\mathbf{E}$. In order to avoid redundant considerations and calculuses, we proceed with our analysis considering only the current density of electrons ($\mathbf{J}_n$).

In 1969 D. Scharfetter and H.K. Gummel (two scientists of Bell Labs), introduced a formula to compute the current density in this case, given φ and the density solution (n) on every nodes. This innovative approach led for the twenty years to follow every simulation which contemplates electric-devices.

We know that the constituve law is composed by a drift component, which depends on the electric field, and a diffusion component, which depends on the variation of the carrier density. Consider a generic element K, we define the drop in voltage $\Delta\varphi^k = \varphi_{i+1} - \varphi_i$. There are three possible situations which are well explain in the picture:

- $\Delta\varphi \gg 0$, mainly drift component from right to left

- $\Delta\varphi \ll 0$, mainly drift component from left to right

- $\Delta\varphi \simeq 0$, mainly diffusion component

With the $Sharfetter-Gummel$ formula it's possibile taking into account every of these situations and solve boundary layer problems which occurs often in presence of strong drift component contribute.

$$J_n^k = q\frac{D_n}{h}\left[n_{i+1}\mathcal{B}\left(\frac{\Delta\varphi^k}{V_{th}}\right) - n_i\mathcal{B}\left(-\frac{\Delta\varphi^k}{V_{th}}\right)\right] \tag{6.1}$$

In the latter case $\Delta\varphi = 0$) the formula became:

$$J_n^k = qD_n\frac{n_{i+1} - n_i}{h} \tag{6.2}$$

which is the correct approximation of the current density using $\mathbb{P}_1$ basis.

Morover many mathematics discover important properties about this method questa parte la vorrei fare meglio. These changes in the direction of the electric field lead to boundary layer problems. It's not possibile afford these situations with a simply upwinding sheme ...

6.2 Extension for the 3D case

The extension of this formula for the 3D case is not trivial. We show the method for the computation of the current density of electrons (the extension for the current density of holes is quite similar). We remark the quasi fermi formula for current density:

$$\mathbf{J}_n = -q\mu_n n\nabla\varphi_n \tag{6.3}$$

where φ_n is the quasi fermi potential level. Let us write (6.3) in function of potential and in a canonic form:

$$\mathbf{J}_n\frac{exp\left(\dfrac{\varphi_n - \varphi}{V_{th}}\right)}{q\mu_n n_i} + \nabla\varphi_n = 0 \tag{6.4}$$

We consider $\mathbf{J}_n \in [L^2(\Omega)]^3$ and $\varphi_n, \varphi \in H^1(\Omega)$. We are able to multiply (6.4) with a generic function $\mathbf{q} \in [L^2(\Omega)]^3$ and then intagrate over the domain Ω:

$$\int_\Omega \frac{exp\left(\dfrac{\varphi_n - \varphi}{V_{th}}\right)}{q\mu_n n_i}\mathbf{J}_n \cdot \mathbf{q}\, d\Omega + \int_\Omega \nabla\varphi_n \cdot \mathbf{q}\, d\Omega = 0 \tag{6.5}$$

We proceed taking the usual discrete space of the constant elemenwise functions:

$$V_h = \left\{w \in L^2(\Omega) : w|_K \in \mathbb{P}_0 \forall K \in \tau_h\right\} \tag{6.6}$$

Now the discrete quantitaties are $\mathbf{J}_n^h \in [V_h]^3$ and $\nabla\varphi_n^h \in V_h$. We desire produce a system of equation on every elements for the three componente of $\mathbf{J}_n$, this is possible with a smart choice of the test function $\mathbf{q}_h \in [V_h]^3$:

$$\mathbf{q}_{1,2,3}^h = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{6.7}$$

From (6.5) we obtain the sequent system of equations defined for every element of the mesh:

$$\int_K \frac{exp\left(\dfrac{\varphi_n - \varphi}{V_{th}}\right)}{q\mu_n n_i} \mathbf{J}_n^h \cdot \mathbf{q}_i^h \, dK + \int_K \nabla\varphi_n^h \cdot \mathbf{q}_i^h \, dK = 0 \qquad \forall i = 1, 2, 3 \tag{6.8}$$

Operating the intagration we obtain the sequent formula for the current density components:

$$[\mathbf{J_n}]_i = -\mathbb{H}_K \left(q\mu_n n_i exp\left(\frac{\varphi - \varphi_n}{V_{th}}\right) \right) \frac{\partial\varphi_n^h}{\partial x_i} \quad i = 1...d \quad \forall K \in \tau_h \tag{6.9}$$

where $\mathbb{H}_K(f)$ is the armonic average on the elment K of the function f.

Although resolve the armonic average with a comlete 3D integration may be expensive in calculation time and propably not necessary. One approximation of this integral would be pass from a 3D integration to 1D integration along one edge of the element K.

$$\left(\frac{\int_K f^{-1} \, dK}{|K|} \right)^{-1} \simeq \left(\frac{\int_{e*} f^{-1} \, de}{|e^*|} \right)^{-1} \tag{6.10}$$

The approximation (6.10) is valid if we consider the correct edge.

Consider a quantity defined on the verteces:

$$\Phi := \varphi - \varphi_n \tag{6.11}$$

which is the difference between the electrostatic potential and the quasi fermi potential level. Now for every element consider two vertices: $\mathbf{x}_m$ s.t. $\Phi(\mathbf{x}_m) = \Phi_m := min_K(\Phi)$ and $\mathbf{x}_M$ s.t. $\Phi(\mathbf{x}_M) = \Phi_M := max_K(\Phi)$. Obviously it exists only one edge which connects these two points and on this one we perform the 1D integration (6.10). First of all as we reduce the dimension is feasible to represent $\sigma(\mathbf{x})$ in a easier mode as follows:

$$\sigma_n(s) = q\mu_n n_i exp\left(\Phi_m + (\Phi_M - \Phi_m)\frac{s - s_m}{|e^*|}\right) \tag{6.12}$$

where $s \in [s_m, s_M]$ is the parameter refered to the edge e^* s.t. $\sigma_n(s_m) = \sigma_n(\mathbf{x}_m)$ and $\sigma_n(s_M) = \sigma_n(\mathbf{x}_M)$. We can easily resolve (6.10) with the substitution of variable:

$$\eta := \frac{s - s_m}{|e^*|}$$

this lead us to the sequent steps of integration:

$$\int_{e^*} \sigma_n^{-1} \, de = |e^*| \int_0^1 \frac{exp\left(-\Phi_m - (\Phi_M - \Phi_m)\eta\right)}{q\mu_n n_i} \, d\eta$$

$$= |e^*| \frac{exp(-\Phi_m)}{q\mu_n n_i} \frac{exp(\Phi_m - \Phi_M) - 1}{\Phi_m - \Phi_M}$$

$$= |e^*| \frac{exp(-\Phi_m)}{q\mu_n n_i} \frac{1}{\mathbf{B}(\Phi_m - \Phi_M)}$$

finally we obtain:

$$\int_K \sigma_n^{-1} \, dK \simeq q\mu_n n_i exp(\Phi_m)\mathbf{B}(\Phi_m - \Phi_M) \tag{6.13}$$

Similar results may be obtained repeating the integration and considering s_M as start point:

$$\int_K \sigma_n^{-1} \, dK \simeq q\mu_n n_i exp(\Phi_M)\mathbf{B}(\Phi_M - \Phi_m) \tag{6.14}$$

Numerical results (qua sarebbe carino fare un po' di test con una parte o l'altra della formula per mettere in crisi) shows that the best choice is e linea combination of these approzimations as follows:

$$\mathbf{J_n}^K = -q\mu_n \left[\frac{n_{min}\mathbf{B}(-\Delta\Phi_{max}) + n_{max}\mathbf{B}(\Delta\Phi_{max})}{2}\right] \nabla\varphi_n^h \tag{6.15}$$

This approach is the natrual extension of the *Sharfetter − Gummel* formula for the 1D case, indeed it's possible demonstrate the equivalence assuming a monodimensionale domain.

(a) *FEMOS*   (b)   (c) *SDEVICE*
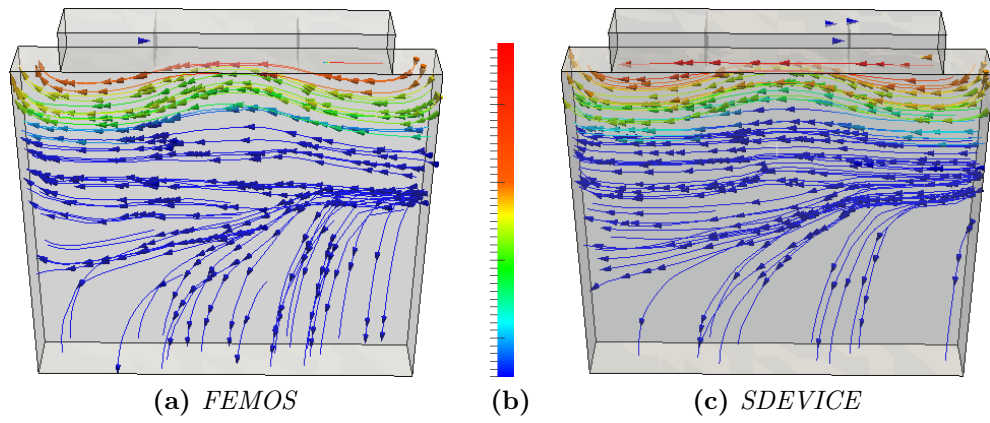
Figure 6.1: Electron current density Vgate = 2.0 [V].

Bibliography

[Deu74] P. Deuflhard. 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.

[GS74] R. Gusmeroli and A. S. Spinelli. Accurate boundary integrals calculation in semiconductor device simulation. *Numer. Math*, 22:289–315, 1974.

[IE83] Babuška I. and Osborn J. E. Generlized finite element methods: Their performance and their relation to mixed methods. *Numer. Anal.*, 20:510–536, 1983.

[PTVF07] William H. Press, Saul A. Teukoisky, 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.

[YT09] Taur Yuan and H. Ning Tak. *Fundamentals of Modern VLSI Devices*. Cambridge University Press, 2009.