Bachelor Dissertation

TOWARDS A NANOSCALE DEVICE SIMULATOR BASED ON THE NON-EQUILIBRIUM GREEN FUNCTION METHOD

Diego Neves de Lemos

Brasília, December 2019

UNIVERSIDADE DE BRASÍLIA

FACULTY OF TECHNOLOGY

UNIVERSIDADE DE BRASÍLIA

Faculty of Technology

Bachelor Dissertation

TOWARDS A NANOSCALE DEVICE SIMULATOR BASED ON THE NON-EQUILIBRIUM GREEN FUNCTION METHOD

Diego Neves de Lemos

Relatório submetido ao Departamento de Engenharia Elétrica como requisito parcial para obtenção do grau de Electrical Engineer

Banca Examinadora

Prof. Daniel C Café, ENE/UnB *Orientador*	—————————————
Prof. Stefan M Blawid, ENE/UnB *Co-orientador*	—————————————
Prof. Alexandre R S Romariz, ENE/UnB *Examinador interno*	—————————————
Prof. Justin W Wells, Dept. of Physics/NTNU *Examinador externo*	—————————————

ABSTRACT

Nanoscale electronic devices will complement and partially substitute silicon-based technology. Although they hold the promises to be more energy-efficient and to allow ultra-high-density integration at a lower cost, their performance is highly impacted by variations in the fabrication process. In addition to that, these emergent devices may have operating principles very different from those of the Si-technology, making new design approaches necessary. To interpret new quantum effects observed in nanodevices, optimize their geometry and decrease costs, we began the implementation of a new atomistic-simulator based on the non-equilibrium Green function formalism, which is the main theoretical tool for quantum transport simulations. In this work, the first steps towards this goal are presented using three examples. First, the impacts of different types of electrostatic boundary conditions over the self-consistent simulations of a short Si-resistor are discussed. In this case, fixing the values of the potential at the contacts makes the device heavily charged, breaking the expected charge neutrality. Second, a qualitative analysis of the current transport in a resonant-tunneling diode is given. Finally, the quantum effects present in two ultra-thin field-effect-transistors with multiple top gates are studied.

RESUMO

Dispositivos eletrônicos de nanoescala, baseados em materiais como nanotubos de carbono e grafeno, poderão substituir ou complementar a atual tecnologia primariamente baseada em Silício. Tais nanodispositivos prometem ser mais eficientes no consumo de energia, ser menores e terem menor custo de fabricação. Entretanto, devido aos princípios físicos que regem seu comportamento, o funcionamento desses dispositivos é altamente impactado por defeitos que ocorrem no processo de fabricação. Além disso, esses componentes ter um comportamento elétrico distinto dos atuais transistores, tornando necessária a criação de novas técnicas para projeto de circuitos. Para estudar novos efeitos físicos presentes em nanodispositivos, otimizar a geometria desses componentes e reduzir custos de fabricação, iniciou-se a implementação de um novo simulador baseado no método da função de Green, que é a principal ferramenta teórica para simulação de transporte quântico de cargas. Apresenta-se neste trabalho três exemplos para ilustrar os primeiros passos do projeto. No primeiro deles discute-se os impactos de diferentes tipos de condições de fronteira para a energia potencial dos elétrons sobre o transporte. Neste caso, verifica-se que se os valores da energia potencial na fronteira forem fixados, o dispositivo torna-se altamente carregado, violando a neutralidade de cargas esperada para o canal. No segundo exemplo faz-se uma descrição qualitativa do transporte de cargas em um diodo de tunelamento ressonante a partir de visualizações fornecidas pelo método. Por fim, estuda-se efeitos quânticos presentes em nanotransistores ultra-finos com múltiplas portas, que podem ser reconfigurados em tempo real para operar em diferentes modos.

SUMMARY

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Over the last decades, the semiconductor industry followed a path called Moore's Law. Proposed by Gordon Moore in the past century, this empirical law states that *the number of transistors on a microprocessor chip will double every two years.* As a consequence, the chip's performance also increases. To follow this rule was the main objective of the industry, that every two years released a roadmap to conduct the research efforts of the area. The principal method employed to achieve this goal was the transistor size reduction [1], allowing more components in the same area. Normally, it also allowed the circuits to go faster and consume less power. Figure 1.1 depicts how the number of transistors per chip evolved over the years.

The current technology is almost entirely based on silicon, a material with dozens of desired properties for this application. Now, silicon transistors are just a few nanometers long and billions of them can be integrated into a small chip. However, beyond this point, further reductions might become financially and technologically unfeasible. For smaller sizes, plenty of effects begin to degrade device performance. Effects like high leakage current, threshold voltage fluctuation, mobility degradation, tunneling effects, and other quantum effects [3]. Besides that, the computing demand began to be determined by mobile devices, which have very different characteristics from those of desktop machines. In this case, several types of very specific circuits are needed, and following Moore's law can become economically prohibitive [1].

Despite these facts, building even smaller devices and making use of those quantum effects are two of the most promising options to further advance electronics. In this case, the current silicon-based devices are no longer absolute. Several emerging revolutionary technologies have been proposed, which create new paradigms and opportunities. The most prominent are nanowire transistors, carbon nanotube field-effect transistors (CNTFETs), and graphene transistors [3]. These devices will probably reach the large-scale industry in the following decade, complementing and partially substituting silicon technology. Moreover, conceptual devices like single-electron transistors might one day cross the laboratory boundary.

On one hand, there are great promises. Nanowire transistors may provide higher carrier mobility due to the reduction of scattering and overcome some fabrication challenges. CNTFETs stand out by their great performance in metrics such as low power, low noise and high speed.
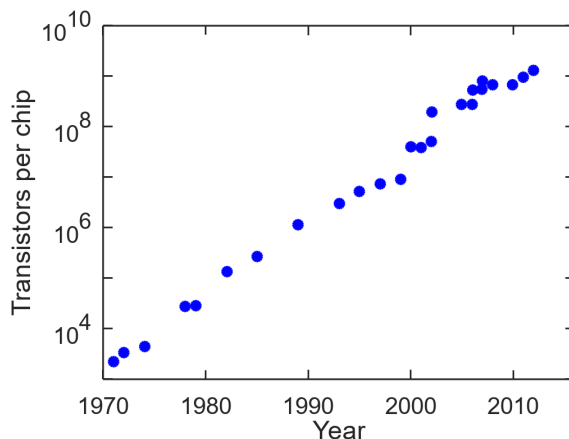
Figure 1.1: Number of transistors per chip in the last decades. Adapted from [1].

Graphene-based devices have similar properties as those of CNTFETs, but may be compatible with some conventional fabrication processes. On the other hand, there are difficult challenges. In all above devices, variations in the geometries and defects can have a large impact on circuit performance. Additionally, the new physics and behavior of these components are not completely understood and, because of that, the design becomes more difficult than in the previous technology.

In order to overcome these obstacles or minimize their impact, a new generation of technology computer aided design (TCAD) tools will be needed. Charge transport in nanodevices, as those introduced above, has to be described by quantum mechanics, because the hypothesis for the application of drift-diffusion equations are no longer satisfied. Therefore, evolution of nanoelectronics will demand quantum mechanical simulators capable of producing accurate results in a plausible execution time.

Among all methods to describe charge transport in nanodevices, the non-equilibrium Green function (NEGF) formalism stands out. It is a numerical effective way to describe devices from ballistic to diffusive regime. In particular, it can provide a rigorous description of the device as an open-system. Non-equilibrium simulations, where a voltage is applied to the electrodes, and systematic inclusion of scattering effects are two of the major advantages of this method. Besides that, it gives several theoretical tools to understand how electrons behave inside the device. Due to its great flexibility, NEGF has been employed to study a wide range of devices: CNTFETs [4, 5, 6], nanowires [7], quantum cascade lasers [8], molecular transistors [9], organic devices [10], photovoltaic cells [11], and so forth.

Problem Statement and Goals

Much remains to be explored in the nanoscale world. Nanodevices' physics is governed by rules inherently different from those used to describe classical devices. In addition to the prominent components cited previously, novel and disruptive devices can be envisioned based on a whole

variety of new quantum phenomena. Besides that, experimental studies of nanoscale devices can make known quantum effects that will demand an advancement of the theoretical and computational models. Accurate and efficient numerical simulations are needed to interpret experimental results and to allow device's design and optimization.

For example, experimental results showed that an asymmetric positioning of the gate in CNT-FETs can lead to a significant reduction in the leakage current (see [12]). Furthermore, the gate can electrically modulate the contact properties in these devices, allowing improvements in the electrical behavior of CNTFETs (see [5]). To simulate these effects, that are not seen in classical transistors, flexibility is needed to experiment with these new geometries and to implement theoretical models for describing the band-to-band tunneling, present in the first case, and the extended contacts, used in the second case.

There are several available tools based on NEGF formalism to efficiently simulate charge transport in nanoscale devices [13], but they are all restricted to predefined theoretical models and device geometries chosen by the developers. This fact makes it harder for other researchers to test their proposals for new methods and to study recently discovered phenomenons. Having the flexibility to alter the simulator code can facilitate the creation of models and techniques to study revolutionary devices.

In order to achieve this flexibility, we began the implementation of a new atomistic-simulator based on the NEGF formalism. This work shows the first steps towards this objective and illustrates some applications. This initial part of the project was focused on the use of NEGF to describe quantum phenomena in 1D ballistic devices, where inellastic scattering is not considered. Topics such as self-consistent solutions for more complex devices and inclusion of scattering events in the channel will be treated in a future work. In summary, the key objectives of this work are:

- Present the basic structure of the Non-equilibrium Green Function method for simulating 1D ballistic devices;

- Discuss how different types of boundary conditions for the electrostatic potential can impact the current transport;

- Provide a qualitative analysis of quantum effects that dictate current transport in a resonant-tunneling diode;

- Describe quantum effects present in ultrathin multiple-gate field-effect transistors;

Text Structure

Chapter 2 introduces the basic theoretical concepts of the NEGF method and gives a simulation example, to illustrate how the quantities provided by the formalism can be used to study quantum charge transport. After that, we present a brief explanation of the electrical behavior of resonant-tunneling diodes and reconfigurable field-effect transistors, that are discussed later. The simulation algorithm is presented in the chapter 3 together with a method to take band-to-band tunneling

into account. Ideal potential profiles for making qualitative studies are also presented. Chapter 4 presents and discusses simulations of four devices: a Si-nanoresistor, a resonant-tunneling diode and two nanotransistors with multiple gates. Finally, conclusions and future work are included in chapter 5.

Chapter 2

Concepts and Definitions

2.1 Non-equilibrium Green Function Formalism

The basic structure of a nanodevice is depicted in fig. 2.1. It is composed by three huge electrodes (source, drain and gate) and one nanoscale channel. The source and the drain are physically coupled to the channel, providing and removing electrons from it. The gate is only electrostatically coupled to the channel, that is, the gate can only influence the electrons in the channel. Essentially, a quantum transport simulation has two objectives: (1) calculate the current through the device for a given bias (i.e. flow of electrons from the source to the drain) and (2) the density of electrons at each position of the channel (i.e. how electrons are distributed). The second quantity is important to describe the device's capacitances.

Both quantities of interest can be obtained from the density matrix ρ. For a 1D channel described by a discrete lattice of N points, from x_1 to x_N, ρ is a NxN matrix with complex elements. The current I can be calculated from ρ using

$$I = -q \operatorname{Tr}(\rho J_{op}), \tag{2.1}$$

where q is the elementary charge and J_{op} is the current operator, that is also a NxN matrix given by

$$J_{op} = \frac{t}{\hbar N} \begin{pmatrix} 0 & -i & 0 & \cdots & 0 & 0 \\ +i & 0 & -i & \cdots & 0 & 0 \\ 0 & +i & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -i \\ 0 & 0 & 0 & \cdots & +i & 0 \end{pmatrix}, \tag{2.2}$$

with $t = \hbar^2/2ma^2$, where m is the electron effective-mass and a is the distance between points (see fig. 2.1). The carrier density $n(x)$, which is the number of electrons per volume at x_j, for $1 \leq j \leq N$, can be directly obtained from the diagonal elements of ρ by

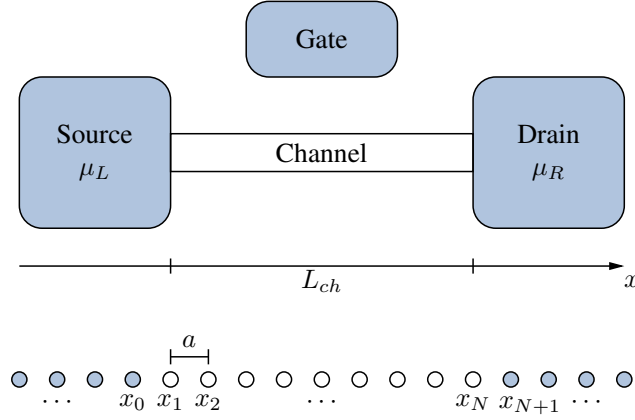$$n(x_j) = \frac{1}{\Omega} \rho_{jj}, \tag{2.3}$$

Figure 2.1: Usual device structure and discretization. Source and drain Fermi levels are μ_L and μ_R, respectively.

where Ω is a volume constant resulting from discretization. In fact, ρ can be interpreted as a generalization of the electron density [2].

There are two situations of interest. When no voltage is applied between source and drain and, consequently, $\mu_L = \mu_R$ the device is said to be in equilibrium. In this case, electrons injected into the channel are in thermal equilibrium with both contacts and no net current flow between the electrodes. When a voltage V is applied (referenced to the source), the Fermi levels difference is $\mu_L - \mu_R = qV$ and the device is said to be in non-equilibrium. As a result, it is no longer clear how the electrons in the channel will be occupied.

The Non-equilibrium Green function method can handle both situations exactly, giving a explicit form for ρ, whereas other methods generally need to employ approximations in order to make non-equilibrium calculations. This is possible because NEGF provides a well-defined way to differentiate which electrons in the channel are in equilibrium with the source, and are occupied using μ_L, and which are in equilibrium with the drain, and are occupied using μ_R. The calculation of the density matrix by NEGF is explained in the following paragraphs.

In the discrete case, the retarded Green function $G(E)$ gives an NxN matrix for each energy E. It can be deduced from the Schrödinger equation and is given by

$$G(E) = \left[(E + i\eta)\bar{I} - H - \Sigma_L(E) - \Sigma_R(E)\right]^{-1} \tag{2.4}$$

in the time-independent scenario [14]. Here, $\bar{I}$ is the NxN identity matrix, η is a infinitesimal positive number included to ensure convergence, H is the Hamiltonian matrix that describes the channel's physics and $\Sigma_{L/R}(E)$ are the so-called self-energy matrices, which describe all the influence of the electrodes over the channel.

The matrix H used in this work is the discrete version (sometimes called tight-binding version) of the effective-mass Hamiltonian operator $\hat{H} \equiv -\frac{\hbar}{2m}\frac{d^2}{dx^2} + U(x)$, where $U(x)$ is the electrons potential energy ($\hat{H}$ maps a function $\psi(x)$ into another function following $\hat{H}\psi(x) = -\frac{\hbar}{2m}\frac{d^2\psi(x)}{dx^2} + U(x)\psi(x)$). The gate electrode impacts the charge transport by changing $U(x)$. Denoting $U(x_j)$

as U_j, H is given by

$$
H = \begin{pmatrix}
2t + U_1 & -t & 0 & \cdots & 0 \\
-t & 2t + U_2 & -t & \cdots & 0 \\
0 & -t & 2t + U_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 2t + U_N
\end{pmatrix},
\tag{2.5}
$$

with $t = \hbar^2/2ma^2$ as already introduced. In quantum mechanics, H is the operator related to the total energy of a particle and is also related to the electron evolution through the Schrödinger equation [15]. The Hamilonian expressed in (2.5) only describes the channel. Electrons in the contacts are described by matrices much larger than that, even semi-infinite in some cases.

One of the most interesting advantages of the NEGF formalism is to lump all the influence of the electrodes into the self-energy matrices $\Sigma_{L/R}$, giving an exact model of the whole system with matrices of a reasonable size. For an end-bonded coupling, where the coupling of the source (drain) with the channel occurs only between the points x_0 and x_1 (x_N and x_{N+1}), $\Sigma_{L/R}$ are given by

$$
\Sigma_L(E) = \begin{pmatrix}
-t\exp(ik_L(E)) & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix}
\text{ and } \Sigma_R(E) = \begin{pmatrix}
0 & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & -t\exp(ik_R(E))
\end{pmatrix}, \tag{2.6}
$$

where $k_{L/R}(E)$ satisfy the dispersion relations in the contacts, i.e., $E = U_0 + 2t(1 - cos(k_L(E)a))$ and $E = U_{N+1} + 2t(1 - cos(k_R(E)a))$. To these relations be good approximations of a parabolic band, t has to be greater than the energy range of interest. Here the electrodes were considered to have a semi-infinite size. It can be interpreted that the self-energy matrices modify the channel's Hamiltonian to account for the coupling. Deduction of the expressions in 2.6 can be found in [14, 16, 17].

For both equilibrium and non-equilibrium situations, ρ can be calculated using

$$
\rho = \int_{-\infty}^{\infty} \frac{dE}{2\pi} \left[f(E - \mu_L)A_L(E) + f(E - \mu_R)A_R(E) \right], \tag{2.7}
$$

where $f(E - \mu) = (1 + \exp((E - \mu)/kT))^{-1}$ is the Fermi function and

$$
A_{L/R}(E) = G^\dagger \Gamma_{L/R} G, \tag{2.8}
$$

with $\Gamma_{L/R}(E) = i(\Sigma_{L/R} - \Sigma_{L/R}^\dagger)$, are called spectral functions ($M^\dagger$ denotes the transpose conjugate of M). Thus, in non-equilibrium, ρ is the sum of separate portions, one in equilibrium with the source, occupied by $f(E - \mu_L)$, and other in equilibrium with the drain, occupied by $f(E - \mu_R)$. Since the diagonal elements of ρ are related with the density of carriers, the diagonal elements of $A_{L/R}/2\pi$ can be interpreted as some kind of density of states. These states are occupied by the Fermi function to find the number of electrons. Therefore, the partial and total local densities of states can be defined as

$$\text{LDOS}_L(x_i, E) = \frac{A_{L,ii}(E)}{2\pi}, \quad \text{LDOS}_R(x_i, E) = \frac{A_{R,ii}(E)}{2\pi} \tag{2.9}$$

and

$$\text{LDOS}(x_i, E) = \text{LDOS}_L(x_i, E) + \text{LDOS}_R(x_i, E). \tag{2.10}$$

The local densities of states show where the electrons can be in the channel and which energies they can have. In particular, the partial densities give information about states occupied by electrons from each contact separately. It is also interesting to define the local densities of charge as

$$\text{LDOC}_L(x_i, E) = f(E - \mu_L)\,\text{LDOS}_L(x_i, E), \quad \text{LDOC}_R(x_i, E) = f(E - \mu_R)\,\text{LDOS}_R(x_i, E) \tag{2.11}$$

and

$$\text{LDOC}(x_i, E) = \text{LDOC}_L(x_i, E) + \text{LDOC}_R(x_i, E). \tag{2.12}$$

LDOC gives where the electrons actually are and which energies they have. The understanding of the transport can be improved by the information contained in these quantities provided by the NEGF method.

Another way to calculate the current through a ballistic nanodevice is by means of the transmission coefficient $T(E)$. It is given by

$$T(E) = \text{Tr}(\Gamma_L A_R) \tag{2.13}$$

and is interpreted as the probability of an electron injected into the channel with energy E be transmitted to the other side. This gives information about what energies can carry current. The current can be calculated from $T(E)$ using

$$I = \frac{q}{h} \int_{-\infty}^{\infty} T(E)[f(E - \mu_L) - f(E - \mu_R)]\, dE. \tag{2.14}$$

In summary, for a given potential energy, NEGF provides a direct way to obtain the density matrix, from which all quantities of interest can be obtained. The process is the same for equilibrium and non-equilibrium situations. The influence of the huge contacts over the channel is completely described by the self-energy matrices. In addition to ρ, the method provides the local densities of states and charges and the transmission coefficient. These quantities can improve the understanding of transport phenomena.

2.2 Poisson Equation and the Self-consistent Solution

For a given U, the density matrix can be calculated. However, U is usually not known a priori, and has also to be calculated. The electrostatic potential is determined by the Poisson equation

$$\nabla^2 U = \frac{q}{\varepsilon}(N_d - n), \tag{2.15}$$

where N_d is the donor doping density and ε is the electric permittivity. The electrodes' potentials are boundary conditions for 2.15. Accordingly, U depends of n, that is determined using U. To
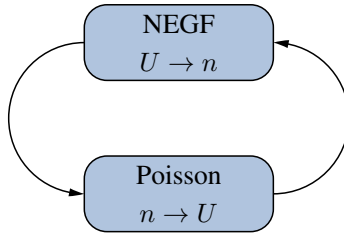
Figure 2.2: Iterative process in the self-consistent approach.

solve this problem, the so-called self-consistent method is applied. Fig. 2.2 illustrates the iterative process in the self-consistent approach.

Basically, self-consistence begins with an approximation for U (or for n) and a corresponding n (or U) is calculated using NEGF (or Poisson). Then, this process is repeated until this quantity stop to change. It means that the final solution for U gives a density that, if put in the Poisson equation, will give the same U again. In other words, the solution of both equations are consistent. The convergence of this method depends largely on the initial approximation, hence a good choice has to be made.

A very common scenario involves devices composed by a 3D channel with symmetry in two directions (charge density must vary only in one coordinate), the Poisson equation can often be simplified to obtain the equation

$$\frac{d^2U(x)}{dx^2} = \frac{q}{\varepsilon}(N_d(x) - n(x)). \tag{2.16}$$

In this case, there are two possible types of boundary conditions: (1) Dirichlet conditions, when $U(x_0)$ and $U(x_{N+1})$ are specified, and (2) Neumann conditions, when $U'(x_0)$ and $U'(x_{N+1})$ are given. At first glance, Dirichlet conditions seem more intuitive, since the electrical potentials at the contacts are controlled by external sources. However, Neumann conditions are the choice of most researchers, because they ensure charge neutrality within the channel [2, 18].

2.3 Example

In this section, an example taken from [2] is given to illustrate the concepts presented in the previous ones. The device proposed by Datta is a short 3D silicon channel, 30 nm long, attached to two semi-infinite contacts. A high doping density is considered, especially near the electrodes. Fig. 2.3 depicts the device's structure and table 2.1 gives the simulation parameters. Simulations were made for voltages from 0 V (equilibrium) to 0.25 V (non-equilibrium). Datta considered that the transport occurs only in the x-direction and that electrons are free in y- and z-directions. Thus, employing separation of variables in the 3D Schrödinger equation, an 1D approach can be used. However, simulations' results have to be summed over all transverse modes [17].
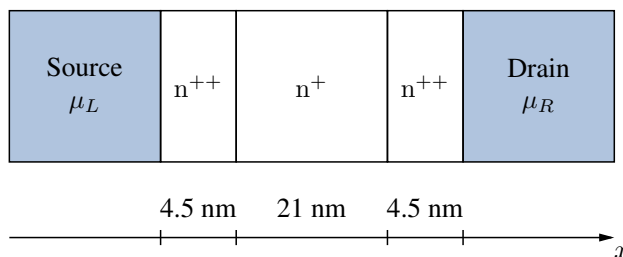
9

Figure 2.3: Short Si-resistor

Effective-mass	$m = 0.25m_0$
Temperature	$T = 290$ K
Permittivity	$\varepsilon = 10\varepsilon_0$
Source Fermi Level	$\mu_L = 0.318$ eV
n^{++} doping	$N_d = 10 \cdot 10^{25}$ m^{-3}
n^+ doping	$N_d = 5 \cdot 10^{25}$ m^{-3}
Number of x points	$N = 100$
Boundary Conditions	Neumann with $U' = 0$

Table 2.1: List of simulation simulation parameters and choices for the nanoresistor of fig. 2.3.

Fig. 2.4 illustrates the self-consistent energy diagrams and carrier densities for all biases, together with the transmission coefficient and the I-V characteristics. It can be noted in fig. 2.4a that, due to the choice of Neumann boundary conditions, the voltage drop inside the channel is just a fraction of the applied bias. This effect can be related to some kind of contact resistance. With these conditions, the values $U(x_0)$ and $U(x_{N+1})$ can be adjusted in order to maintain charge neutrality (fig. 2.4b). It can be seen in fig. 2.4c that electrons injected with energy below $\max(U(x))$ a very low probability to be transmitted, hence these carriers are reflected. For energies above $\max(U(x))$, $T(E) \approx 1$. The tiny oscillations seen in $T(E)$ are due to resonances with the eigenstates of the channel. As expected for a resistor, a linear current-voltage curve was obtained (fig. 2.4d).

The local densities of states and of charges are shown in figs. 2.5 and 2.6, respectively. For LDOS, the bright regions indicate that many electrons can have those energies and be in those positions. Thus, the black region below the conductance band is the Energy gap. For LDOC, the bright regions have a bigger populations of electrons than the dark ones. In both figs., images (c) and (d) are related to states or electrons associated with the source, whereas images (e) and (f) are related to states or electrons associated with the drain. In these figs., resonances with the channel's eigenenergies lead to oscillations between higher and lower densities. It should be noted that in some cases (e.g. fig. 2.5d and fig. 2.6d) there are states or electrons with energies under $\max(U(x))$ that reach the other side by tunneling. But, when the barrier thickness is not small
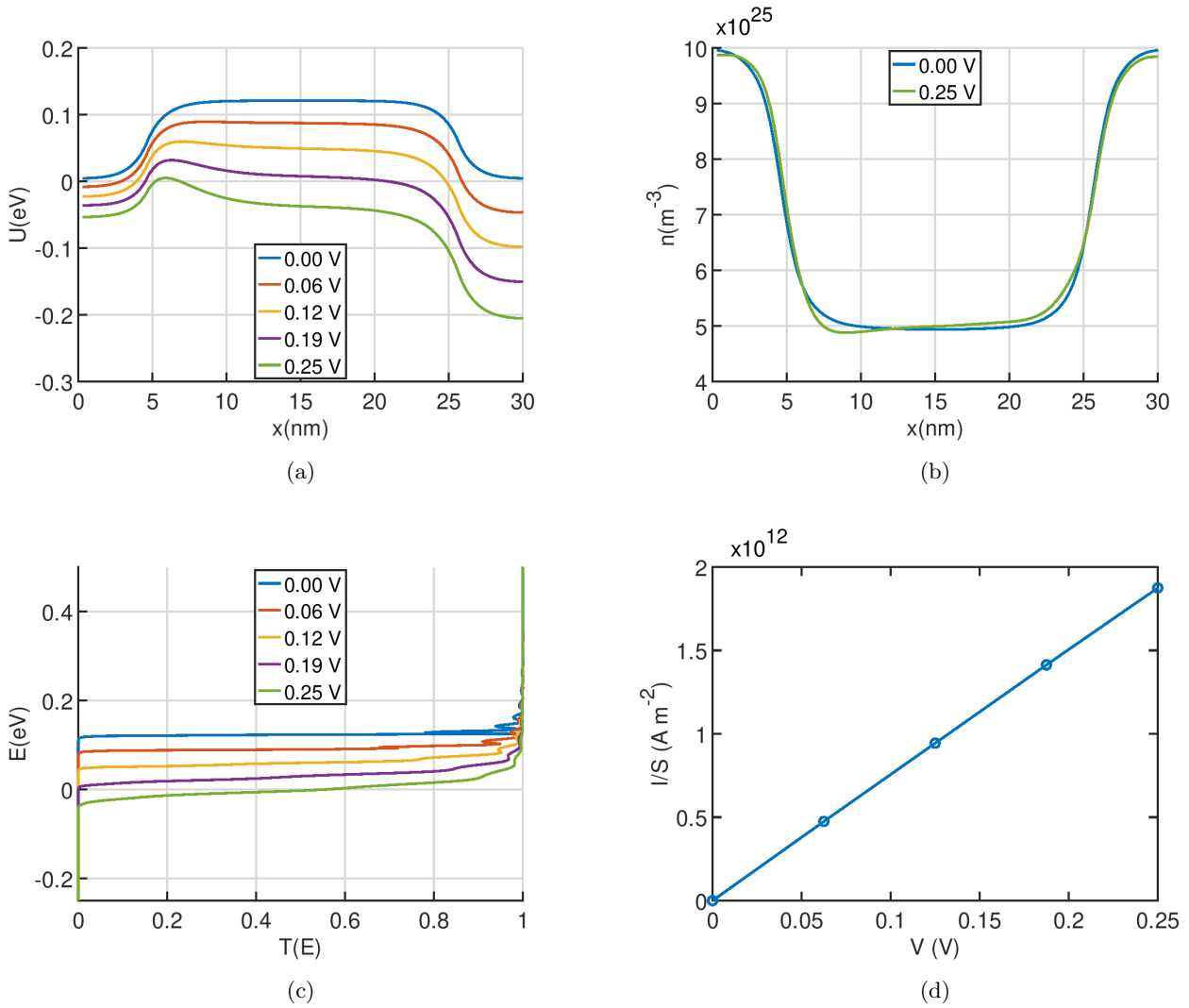
Figure 2.4: Non-equilibrium self-consistent simulations for Datta's [2] example (fig. 2.3). (a) Conductance Band, (b) Electrons concentration, (c) Transmission coefficient and (d) I-V characteristic.

enough to allow tunneling, electrons accumulate (see, for example, the bright region in fig. 2.6d for $x < 5$ nm).

The quantities shown in this sections, especially $T(E)$, LDOS and LDOC, are powerful tools to interpret and explain the charge transport in ballistic nanostructures. Transmission allow to distinguish which energies can (or cannot) contribute to the current. The local densities allow to determine and visualize where quantum phenomena occur. Therefore, these quantities will be used extensively in the following chapters.
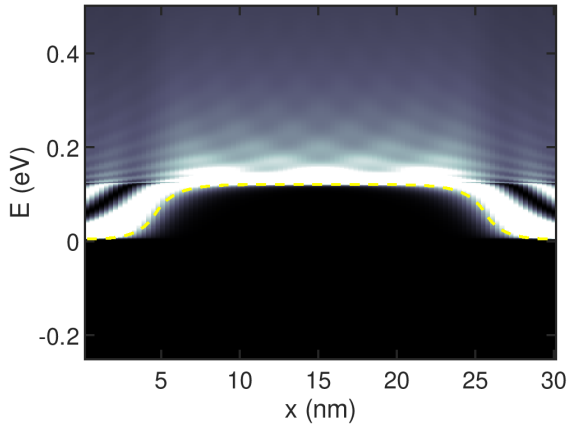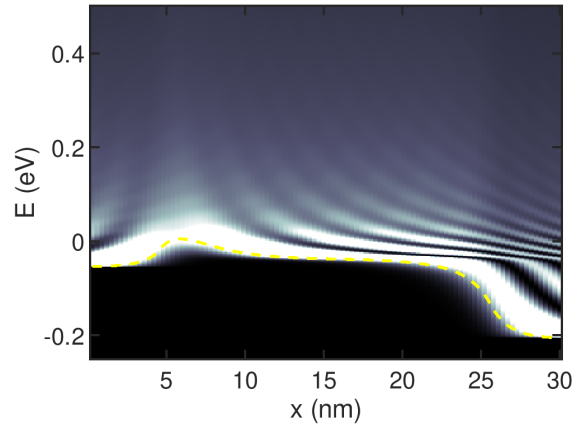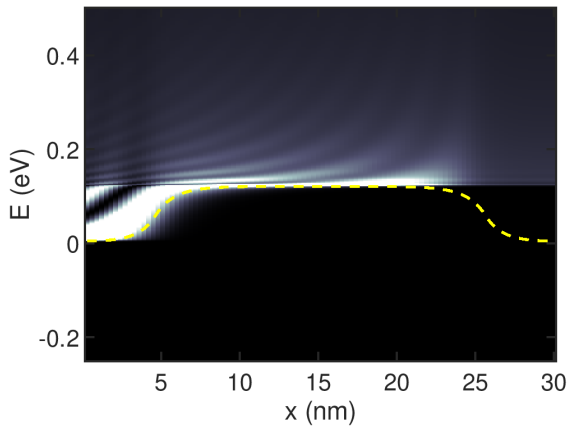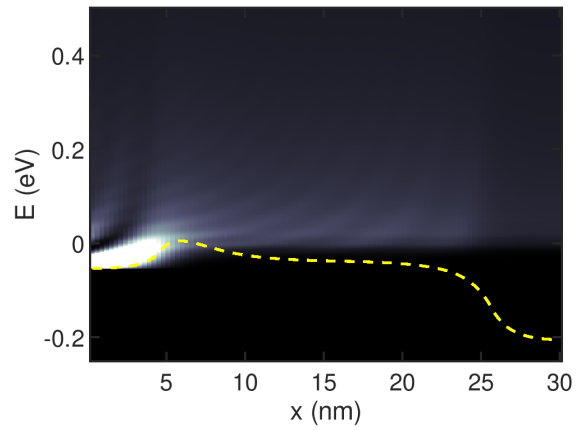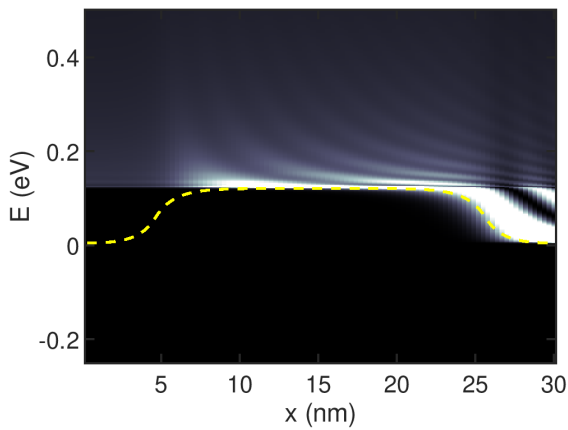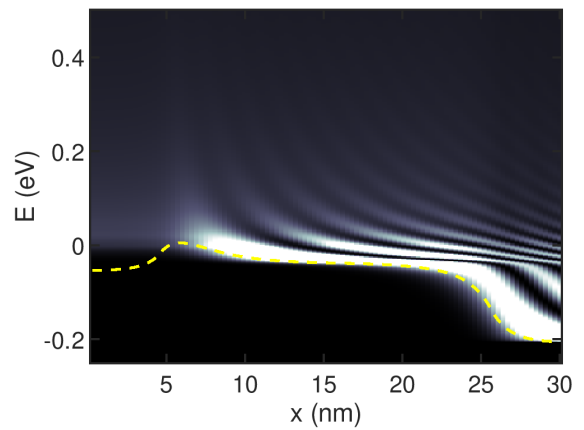
Figure 2.5: Local densities of states for (a - total, c - left, e - right) $V = 0$ V and (b - total, d - left ,f - right) $V = 0.25$ V. (a)
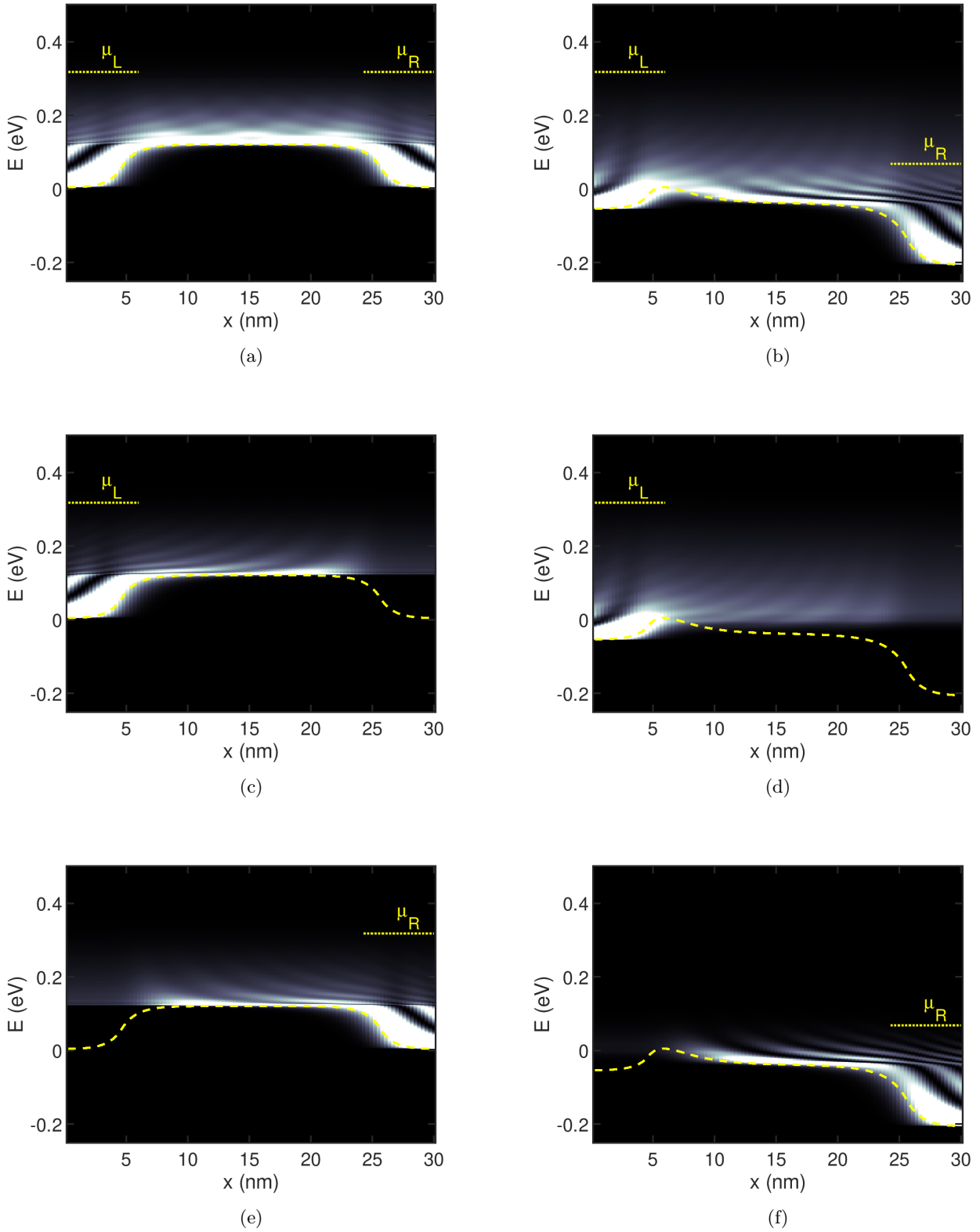
Figure 2.6: Local densities of carriers for (a - total, c - left, e - right) $V = 0$ V and (b - total, d - left ,f - right) $V = 0.25$ V. (a)
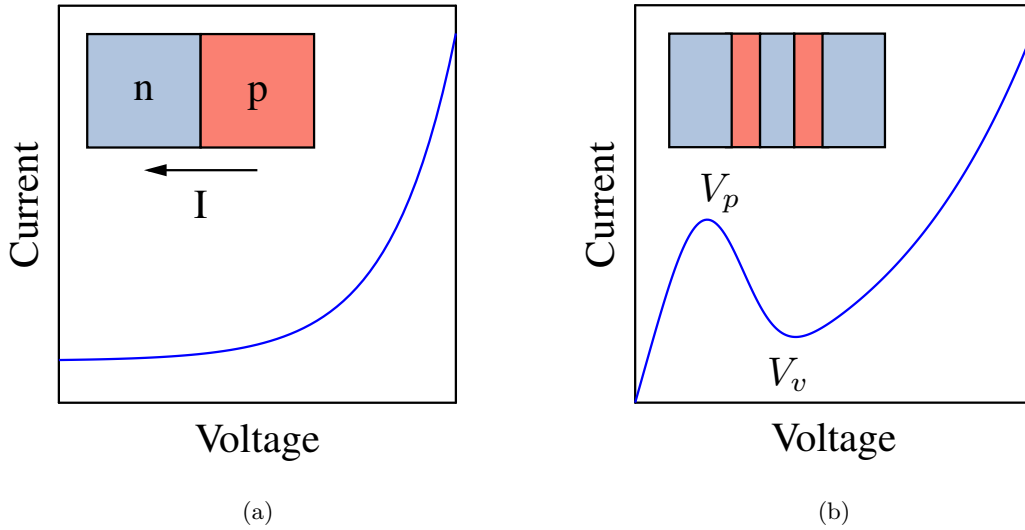
Figure 2.7: Current-voltage characteristics and basic structures of a (a) pn-diode and a (b) resonant-tunneling diode. Possible materials for a RTD are GaAs, for blue region, and $Al_xGa_{1-x}As$, for the other region.

2.4 Devices of Interest

In addition to the nanoresistor already introduced, other three devices will be studied in this work: a resonant-tunneling diode (RTD), a reconfigurable field-effect transistor with three top gates (3GRFET) and a multi-mode field-effect transistor with two top gates (2GmmFET). Quantum effects are crucial in the operation of these devices. This section briefly introduces how they work.

2.4.1 Resonant-tunneling Diodes

Diodes are two-terminal devices with a non-linear current-voltage characteristic. Basically, they have a very low current for low bias and an exponentially-increasing current for high voltages. A very common type of diode is composed by a pn-junction, i.e., two regions with different doping type put together [19]. Fig. 2.7a illustrates the I-V curve for this device and its channel structure.

For a pn-diode, I always increase with V. Thus, it has a positive differential-resistance, defined by

$$R_d(V) = \left(\frac{dI(V)}{dV}\right)^{-1}, \tag{2.17}$$

for all V. An ordinary resistor behaves in a similar way, having a positive and constant R_d.

On the other hand, RTDs have a very different electric behavior. Composed by an heterostructure (i.e. layers of different materials), RTDs possess one or more regions of negative differential-resistance in their I-V characteristics [16, 20]. That is, for some V, current decreases

14

with a voltage increase. This "strange" behavior, allied with the small size of this device, can be useful to build high-frequency oscillators.

RTD operation is based on the energy level of a quasi-bound state that is localized in the central material layer. Diode's parameters are chosen in order to put this level above μ_L in equilibrium. Electrons injected into the channel can only be transmitted if they tunnel to this state. The energy level of this state can be controlled by V. A current peak occurs when V is enough to make a match between μ_L and the state's energy (this match is the resonance). Then, when V exceeds V_p, this match is broken and the current reaches a valley for $V = V_v$. After that, other mechanisms begin to contribute to the transport and current increases again.
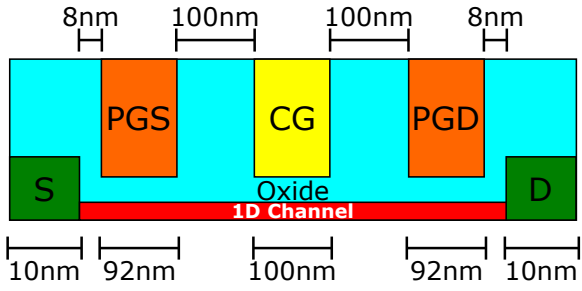
2.4.2 Reconfigurable and Multi-mode FETs

Transistors are multi-terminal devices. Current flow between two electrodes (source and drain) and the other terminals, called gates, are used to control the channel's conductance. The gate terminal changes the electrostatic potential within the channel. Traditional FETs have one gate and a chemically doped channel. However, devices with multiple gates can be interesting, especially electrostically doped devices.

Fig. 2.8a illustrates one of these devices, proposed in [21], that is composed by an 1D undoped channel and three top gates, where two of them (PGS and PGD) are programming gates (this device is called 3GRFET). These gates shift the energy bands at the edges, leading to an electrostatic n- or p-type doping in portions near the source and drain. Thus, with ambipolar contacts (i.e. contacts that can provide both electrons and holes), device can be reconfigured at run time to operate as an n-type or p-type transistor by changing $V_{PGS/D}$. Fig. 2.8b shows transfer characteristics for both configurations.

In addition to polarity configuration, the programming gates can be used to change the switching mechanism from thermal injection to band-to-band (btb) tunneling. The first mechanism suffers from higher leakage currents due to a physical limit, that implies a maximum slope for the transfer curves in 2.8b at the transition between off and on states [19]. On the other hand, this limit can be surpassed by the second mechanism, making it appropriate for low-power applications. However, the current delivered at on state of this regime is significant lower than that achieved with thermal injection. The device in 2.8a can operate as a tunnel FET (transport based in btb tunneling) by applying opposite biases at PGS and PGD. Therefore, the three gates geometry also allows a trade-off between high-performance (thermal injection with higher currents) and low-power (btb tunneling with a lower leakage currents).

Fig. 2.9a depicts an optimized version of the 3GRFET, created to better explore the energy/performance trade-off. This device is called multi-mode FET (abbreviated for 2GmmFET). With just two gates and unipolar contacts, this transistor is simpler, shorter and has a lower leakage current than the 3GRFET. However, it can no longer alternate between n-type and p-type operation. Fig. 2.9b shows the transfer curves for both high-performance and low-power modes. Quantum effects that determine the carrier transport in both of these devices will be further
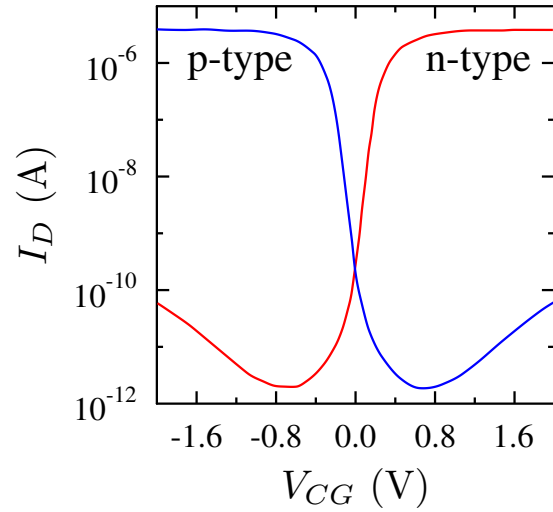
(a)   (b)

Figure 2.8: Reconfigurable FET (a) geometry and (b) transfer characteristics for $|V_{DS}| = 1$ V. n-type or p-type configurations are achieved by making $V_{PGS/D} = +1.2$ V or -1.2 V, respectively.
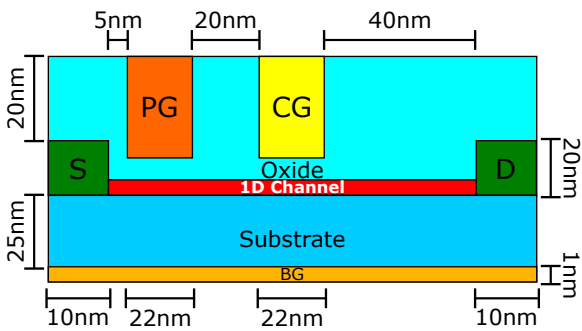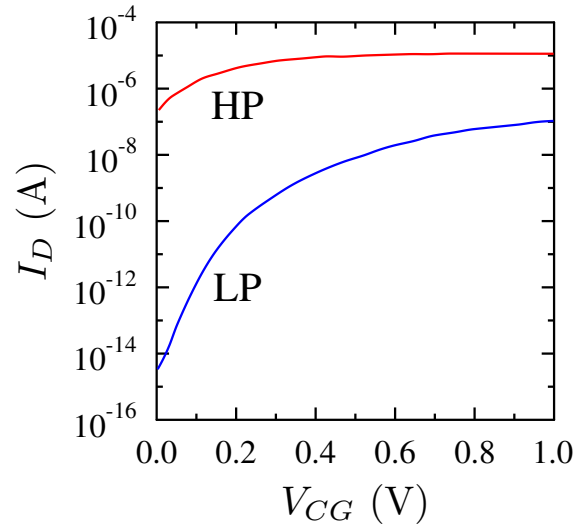
(a)   (b)

Figure 2.9: Multi-mode FET (a) geometry and (b) transfer characteristics. For $V_{PG} = +1$ V, transport is based in a thermal mechanism and device operates in a high-performance mode. For $V_{PG} = -1$ V, transport is based in band-to-band tunneling and device operates in a low-power mode.

explored in the results chapter.

Chapter 3

Methodology & Tools

One of the main objectives of this work is to illustrate how the tools provided by the NEGF method can give us a new and powerful way to understand ballistic devices. To achieve this goal, we implemented some necessary techniques to study two new conceptual electronic devices, reconfigurable and multi-mode transistors, introduced in section 2.4. Our implementation had Datta's work [2] as a starting point, but in order to describe the nanoFETs, new techniques have been employed.

In this chapter, we initially present in section 3.1 the basic structure of the NEGF implementation presented in [2], which describes the ballistic non-equilibrium electron transport in a nanoresistor with high doping-level. The following sections introduce methods needed to describe effects present in the devices of interest. The inclusion of a valence-band and the description of band-to-band tunneling effect are discussed in 3.2 and some idealized approaches are proposed in 3.3.

3.1 Algorithm

Algorithm 1 presents the basic structure of Datta's implementation. As already stated, the charge transport in a nanodevice is simulated by the self-consistent solution of the Green function and Poisson equation. To start the iterative process, we must provide four vectors: a good initial guess for the electron potential energy, the donor doping profile, the set of points given by energy coordinate discretization and the set of applied bias values. We also have to give five numbers as inputs: the source Fermi level, the hopping energy and the distance between sites, the number of points used in the spatial discretization and the tolerance, which determines when the self-consistent process finishes. As results, the real potential profile and the density matrix are obtained for all applied voltages.

Algorithm 1: Self-consistent scheme.

Data: initial U, N_d, $E = \{E_k\}$, $V = \{V_i\}$, μ_L, t, a, N, tol

Result: final U and ρ

1 error $\leftarrow$ 10 ;

2 $\rho \leftarrow$ zero NxN matrix ;

3 foreach *bias* $V_i \in V$ **do**

4   `%%% Self-consistence %%%`;

5   **while** *error > tol* **do**

6   `%%% transport equation %%%`;

7   **foreach** *energy point* $E_k \in E$ **do**

8 $H \leftarrow$ `Hamiltonian`(U, t, N) ;

9   calculate $f(E_k - \mu_L)$ and $f(E_k - \mu_L + V_i)$;

10   find $\Sigma_{L/R}(E_k)$ using 2.6;

11 $\Gamma_{L/R}(E_k) \leftarrow i(\Sigma_{L/R} - \Sigma_{L/R}^\dagger)$;

12   obtain $G(E_k)$ applying 2.4;

13   calculate $A_{L/R}(E_k)$ with 2.8;

14   accumulate to find ρ from 2.7;

15   **end**

16 $n \leftarrow (1/a)\cdot$`diag`(ρ)

17   `%%% Poisson equation %%%`;

18 $U \leftarrow$ `solverPoisson`$(U, N_d - n)$;

19   error $\leftarrow$ stopping_criteria();

20   **end**

21   return U and ρ;

22 end

In the simulation process, for each bias of interest one self-consistent calculation is made. Generally, we go from a small to a large bias and the resultant potential profile of a previous calculation is used to begin the next. Given a bias and a guess for U, we find the density matrix from NEGF and after that we use the diagonal of ρ to obtain the density of electrons, from which we calculate a new U using the Poisson equation. This process repeats itself until we have a error smaller than the specified tolerance. Once we have ρ, obtained using the correct potential energy, we can obtain all quantities of interest, e.g. current.

3.1.1 Numerical integration

The density matrix is given by a matrix integral for energies from $-\infty$ to ∞ (see equation 2.7). To evaluate this integral, we use our knowledge about the integrand to change these limits. For $E \gg \mu_L$, the Fermi functions $f(E - \mu_L)$ and $f(E - \mu_R)$ are very close to zero. Moreover, for E lower than the lowest eigenenergy, the spectral functions $A_{L/R}(E)$ become negligible. Thus, ρ

can be calculated by the proper integral

$$\rho = \int_{E_{min}}^{E_{max}} \left(f(E - \mu_L) A_L(E) + f(E - \mu_R) A_R(E) \right) \frac{dE}{2\pi}, \tag{3.1}$$

where E_{max} is a few kT above μ_L and E_{min} is a sufficient low value.

A numerical integration scheme is used to obtain ρ. In this work we used a simple method that uses an uniform set of points $\{E_k\}$, i.e. $E_{k+1} - E_k = \triangle E$ for $\forall k$, distributed between E_{min} and E_{max}. For each energy point, we evaluate the integrand and then we calculate

$$\rho \approx \sum_k \left(f(E_k - \mu_L) A_L(E_k) + f(E_k - \mu_R) A_R(E_k) \right) \frac{\triangle E}{2\pi}. \tag{3.2}$$

This method has simplicity as the main advantage and is a good option for a first implementation. Its major drawback is the requirement of a huge number of energy points to get an accurate result. For each E_k a very large matrix has to be inverted to find $G(E_k)$. This step is the most time consuming calculation of the method. Thus, this scheme can become impracticable for systems with very thin localized states.

In practical simulators, an adaptive integration scheme together with a contour integration method are applied [22]. The first method distributes more energy points in regions where the integrand varies rapidly and less when it varies slowly, needing a quantity of points much smaller to achieve the same accuracy. The second method is based in complex integration theory to change the integral in 2.7 by an integral over a complex contour, in which the integrand is a smooth function, also resulting in a lower quantity of needed points.

3.1.2 Numerical Solution of the Poisson equation

Once the density matrix was calculated using the integration method discussed in the former section, the Poisson equation (eq. 2.16) can be used to find a new approximation of $U(x)$. Following a simple discretization scheme, the discrete version of the equation is given by

$$U_{j+1} - 2U_j + U_{j-1} = \frac{qa^2}{\epsilon} (N_{Dj} - n_j), \quad j = 1, \dots, N, \tag{3.3}$$

for $x = x_j$. In particular, the equations for $j = 1$ and N are

$$U_2 - 2U_1 + U_0 - \beta (N_{D1} - n_1) = 0$$

and

$$U_{N+1} - 2U_N + U_{N-1} - \beta (N_{DN} - n_N) = 0.$$

The values U_0 and U_{N+1} are related to points located outside the channel, thus they determine the boundary conditions for the Poisson equation. For Dirichlet boundary conditions, these values are given and have to stay constant during the process. For zero Neumann boundary conditions, when the derivative of U is null at $j = 1$ and $j = N$, the equalities $U_1 - U_0 = 0$ and $U_{N+1} - U_N = 0$ are imposed. In this case, the new system of equations is

$$U_2 - U_1 - \beta \left(N_{D1} - n_1 \right) = 0,$$
$$U_{j+1} - 2U_j + U_{j-1} - \beta \left(N_{Dj} - n_j \right) = 0, \quad j = 2, \ldots, N-1$$
$$-U_N + U_{N-1} - \beta \left(N_{DN} - n_N \right) = 0,$$

Generally, this is a non-linear system of equations. One of the better known methods to solve problems like this one is the Newton-Raphson method. Given an initial approximation $U^{(0)}$ for the potential energy, a new vector $U^{(1)} = U^{(0)} + \triangle U^{(0)}$ can be calculated, where the elements of $\triangle U^{(0)}$ obey the linear system of equations

(3.4)
$$\begin{pmatrix} M_1^{(0)} & -1 & 0 & \cdots & 0 \\ -1 & M_2^{(0)} & -1 & \cdots & 0 \\ 0 & -1 & M_3^{(0)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & M_N^{(0)} \end{pmatrix} \begin{pmatrix} \triangle U_1^{(0)} \\ \triangle U_2^{(0)} \\ \triangle U_3^{(0)} \\ \vdots \\ \triangle U_N^{(0)} \end{pmatrix} = \begin{pmatrix} R_1^{(0)} \\ R_2^{(0)} \\ R_3^{(0)} \\ \vdots \\ R_N^{(0)} \end{pmatrix}$$

with
$$M_j^{(0)} = 2 - \beta \left. \frac{dn_j}{dU_j} \right|_{U_j^{(0)}} \quad \text{and} \quad R_j^{(0)} = U_{j+1}^{(0)} - 2U_j^{(0)} + U_{j-1}^{(0)} - \beta(N_{Dj} - n_j). \quad (3.5)$$

After finding $U^{(1)}$ more iterations could be made to find better approximations of U. However, during the self-consistent method, the Poisson equation will be solved for an unreal electron density. Therefore, a very high accuracy at this point does not make sense. It should be noted that the charge density derivative in M_j is of crucial importance. The electron density varies exponentially with U_j, and, if this term is not considered, an instability will appear in the iterative process [23].

Another possible technique is the relaxation method, that probably is the simplest method to solve the Poisson equation. In order to apply it, we just need to rewrite equation 2.16 as

$$U_j = \frac{1}{2} \left(U_{j+1} + U_{j-1} \right) - \frac{\beta}{2} \left(N_{Dj} - n_j \right). \quad (3.6)$$

A simple interpretation follows from the above equation. The potential energy at point j is given by two contributions: (1) the average of the potential at the nearest neighbors and (2) the charge density at point j. Therefore, starting from an initial approximation for U, a new value for U_j can be calculated using 3.6. Some iterations must be made to find a new value reasonably different from the initial approximation [24].

Imposing boundary conditions is straightforward. To use Dirichlet conditions, we just need to keep U_0 and U_{N+1} constant, that is, do not calculate new values for them. To use Neumann

conditions, we need to keep the differences $U_1 - U_0$ and $U_{N+1} - U_N$ constant. This can be made by updating U_0 and U_{N+1} by U_1 and U_N after each iteration.

This method stands out by its simplicity and flexibility. It can be rapidly implemented to calculate U in different situations. However, relaxation method is generally slow and is not very accurate. To make the convergence faster, we can use a method called overrelaxation, where the new potential U_{new} is given by an weighted average between the old value U_{old} and the value U_{calc}, calculated from 3.6, i.e.

$$U_{new} = \omega U_{calc} + (1 - \omega)U_{old}, \tag{3.7}$$

with $\omega > 1$. In this way, U_{new} tends to change faster. On the other hand, for $\omega < 1$ (under-relaxation or damping), a better accuracy can be achieved, but more iterations will be needed. Therefore, the value of ω represents a trade-off between accuracy and speed.

3.2 Valence band and band-to-band tunneling

A straightforward way to include a valence band and band-to-band tunneling was proposed in [25]. In this method, the transport equation is solved for an energy-dependent effective potential $U_{eff}(E, x)$, that is constructed from the channel potential. Denoting the conduction and the valence potentials by $U_c(x)$ and $U_v(x)$, respectively, the mid-gap potential $U_n(x)$ is defined as

$$U_n(x) = \frac{U_c(x) + U_v(x)}{2}. \tag{3.8}$$

A particle injected in the channel with energy E is considered an electron for x such that $E > U_n(x)$ and a hole for x in which $E < U_n(x)$. This differentiation leads to the definition of U_{eff} as

$$U_{eff}(E, x) = \begin{cases} U_c(x), & E \geq U_n(x), \\ 2E - U_v(x), & E < U_n(x). \end{cases} \tag{3.9}$$

Fig. 3.1 exemplifies the form of U_{eff} for three different injection energies. The effective potential is only used in the transport equation (NEGF), but the Poisson equation is not modified for self-consistent simulations. In addition to this modification, an energy dependent effective-mass is used in [25]. However, in this work the effective-mass was considered constant for simplicity.

3.3 Ideal band diagrams

The self-consistent method can suffer of many convergence problems [23]. Fortunately, for some devices a qualitative analysis can be made using ideal band diagrams, i.e., simplified potential profiles that approximate the real ones. In this case, since U is given, only the transport equation have to be solved. This approach provides an easy way to gain a better understanding about devices operation. Here, ideal models are given for two of the three devices introduced in 2.4.

An ideal band diagram for resonant-tunneling diodes was proposed in [16]. In equilibrium (fig. 3.2a), it consists of a constant value for each part of the channel, forming two potential
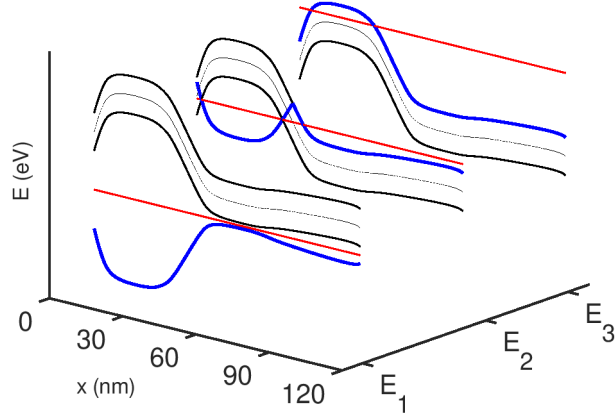
Figure 3.1: Energy-dependent effective potential profile for inclusion of btb tunneling. The black lines are the conductance and valence bands, the blue lines are U_{eff} and the red lines indicate the injection energies.
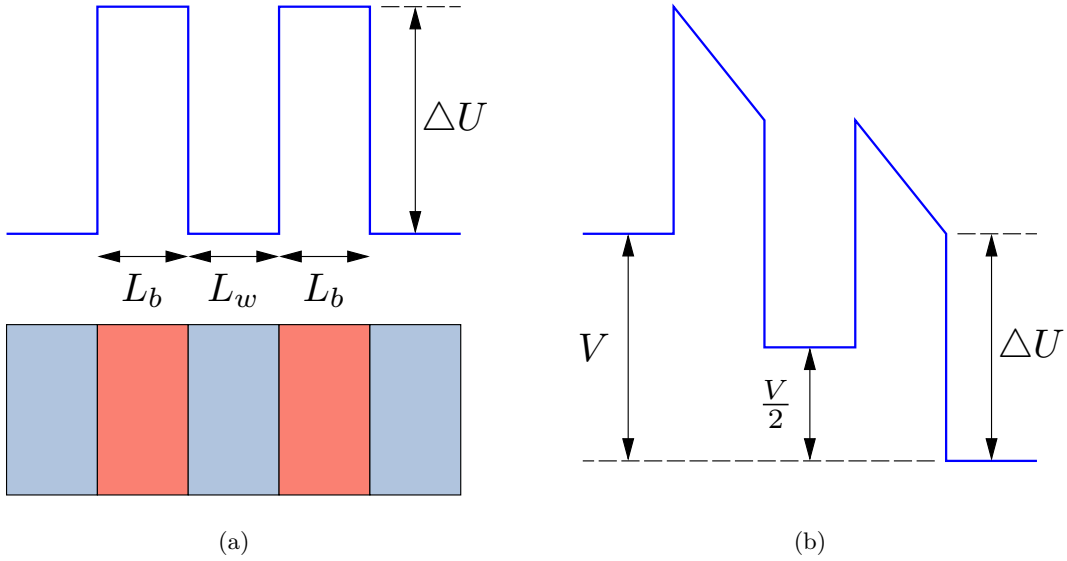

Figure 3.2: Ideal energy diagram for a resonant-tunneling diode. (a) Equilibrium case and structure. (b) Non-equilibrium case for a bias V.

barriers, with width L_b and height $\triangle U$, and one quantum well, with width L_w. When a voltage is applied (fig. 3.2b), half of the bias drops in each barrier. This simplified profile provides a good approximation if V is small. For a device far from equilibrium, most of the bias drops outside the barriers, differing from the model proposed here.

In order to compare the switching process in the multi-mode FETs for both HP and LP modes, we propose a similar approach to approximate the energy bands of this device. Fig. 3.3 gives a
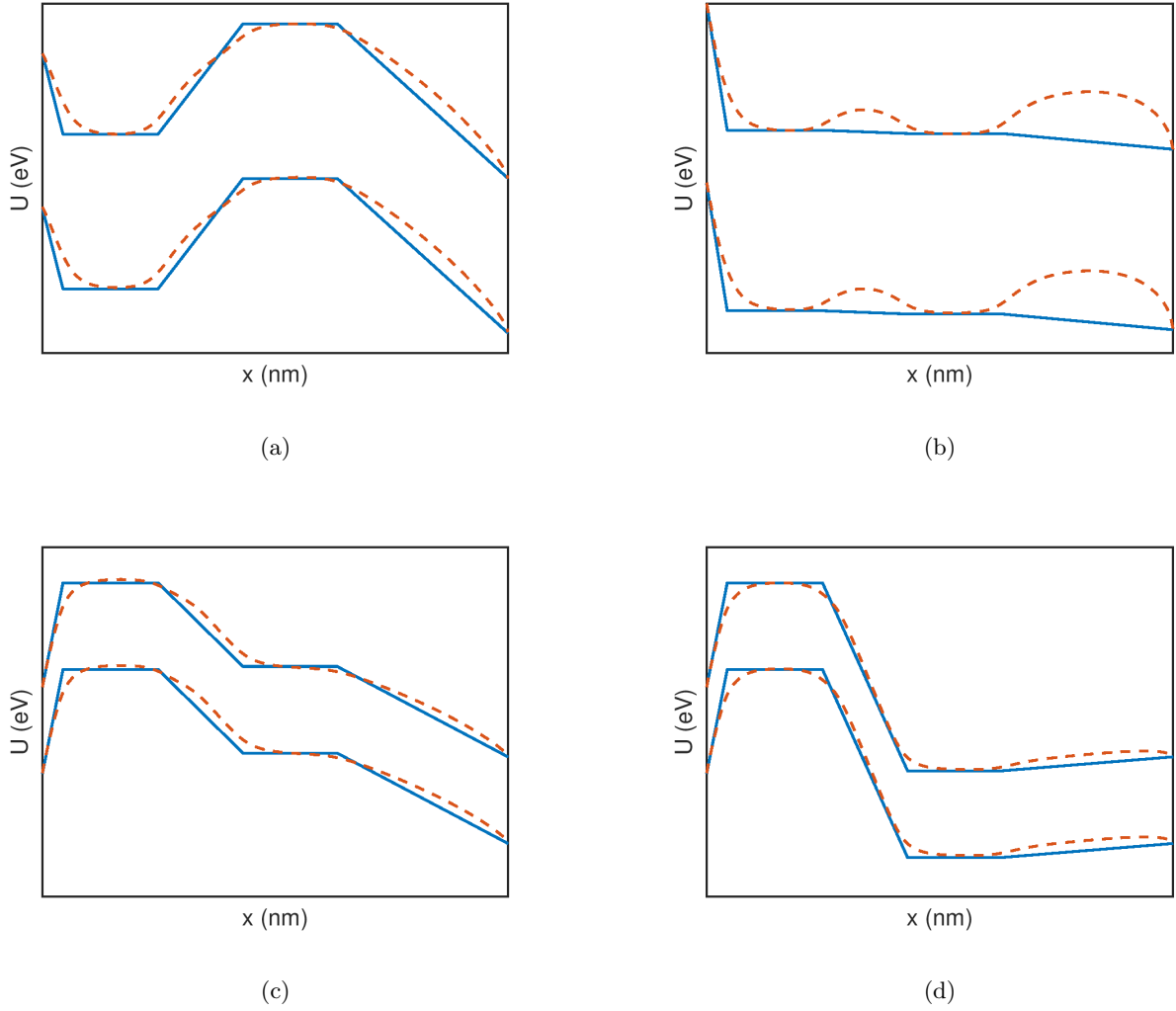
Figure 3.3: Ideal energy diagrams for 2GmmFETs (blue solid-line for ideal and traced-line for real). (a) HP mode in off state, (b) HP mode in on state, (c) LP mode in off state and (d) LP mode in on state.

comparison between the ideal and the real diagrams, whose calculation is described elsewhere [26]. Basically, the ideal profile is constant under the gates and the other parts are connected by straight lines. During the simulations, the value of U under the control gate is varied according to V_{GS}. Since the charge density is related to the potential curvature via the Poisson equation, 2.16, these simplifications lead to perfect neutrality where U is constant and to peaks of density where the lines connect. This behavior does not reflect reality. However, the proposed model is enough to study the quantum effects that determine current transport in a 2GmmFET.

Chapter 4

Results and Discussions

This chapter presents and explains the main results obtained in this work. First, the example of ballistic transport in a nanoresistor, discussed in section 2.3, was the base for a computational experiment, which have the purpose of clarifying what are the impacts of different types of boundary conditions for the Poisson equation.

Second, we studied the transport of charge in a resonant tunneling diode employing the ideal potential profile introduced in 3.3. The operation of this type of device is based in tunneling to a fundamental quasi-bound state formed in the channel. The local density of states and the transmission coefficient obtained using NEGF were used to present a qualitative explanation of the current flow.

Finally, we discuss some crucial quantum effects that determine the transport of charge in ultrathin transistors with multiple top gates. Two classes of devices were simulated. The first class is composed of reconfigurable transistors with three gates, which can be switched dynamically between an n-type and a p-type operation. The second class has multi-mode transistors, which give the possibility of a trade-off between high performance and low power.

4.1 Impacts of Different Types of Electrostatic Boundary Conditions

In this section, we present different self-consistent simulations of the ballistic short Si-resistor introduced in 2.3. Now Dirichlet boundary conditions (fixing $U(x_0) = 0$ eV and $U(x_{N+1}) = -V$ eV) were employed to solve Poisson equation instead of the Neumann conditions (that fix $U'(x_0) = U'(x_{N+1}) = 0$) proposed in [2]. This second type is usually used in the literature for 1D simulations with the objective of ensuring charge neutrality in the channel. However, this choice normally makes the voltage drop within the channel be smaller than the applied bias, indicating some kind of contacts resistance that is not completely understood. On the other hand, Dirichlet conditions remove this effect, but do not mathematically guarantee charge neutrality. With these simulations, we hope to clarify what are the impacts of this choice for the current transport.
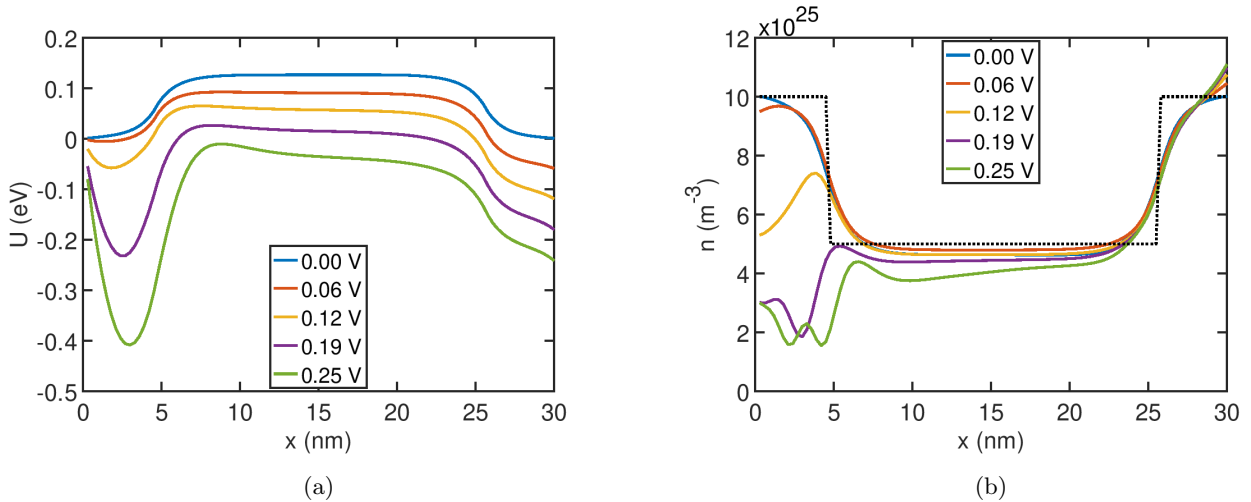
Figure 4.1: Non-equilibrium self-consistent simulations using Dirichlet boundary conditions, with $U_0 = 0$ eV and $U_{N+1} = -V$ eV. (a) Conductance Band, (b) Electron density (traced-line indicates donor density).

Effective Mass	$m = 0.25m_0$
Temperature	$T = 290$ K
Permittivity	$\varepsilon = 10\varepsilon_0$
Source Fermi Level	$\mu_L = 0.318$ eV
Number of x points	$N = 100$
Number of E points	$N_E = 501$
error and tolerance	$\max(\delta U)/\max(|U|) < 0.001$

Table 4.1: List of simulation parameters for Dirichlet boundary conditions

The self-consistent potential energy and the electron density for V from $0V$ to 0.25 V are shown in fig. 4.1. The local densities of charges for $V = 0.25$ V are shown in fig. 4.2 to explain the changes seen in the former results. Finally, fig. 4.3 gives a comparison between the results obtained with Neumann (already shown in 2.3) and with Dirichlet conditions. The relaxation method was used to make these simulations. The simulations' parameters are listed in table 4.1, where δU denotes the absolute changes in the potential energy between two consecutive iterations.

It can be seen in fig. 4.1a that $\max(U)$ decreased with the increase in V. With U_0 held in 0 eV, this caused a reduction in the number of electrons trapped in the region near the source, decreasing the electron density (fig. 4.1b). As a consequence, a sharp positive curvature appeared in U. At a certain point, the bias was enough to pull down $\max(U)$ below 0 eV. In this case, the electrons injected by the source can no longer be confined near the electrode (fig. 4.2) and hence the device became very negatively charged at that region. It should also be noted the presence of small oscillations in the electron density for $x < 8$ nm and $V = 0.25$ V. They are caused by the occupation of a quasi-bound state formed in the potential well (fig. 4.2c).
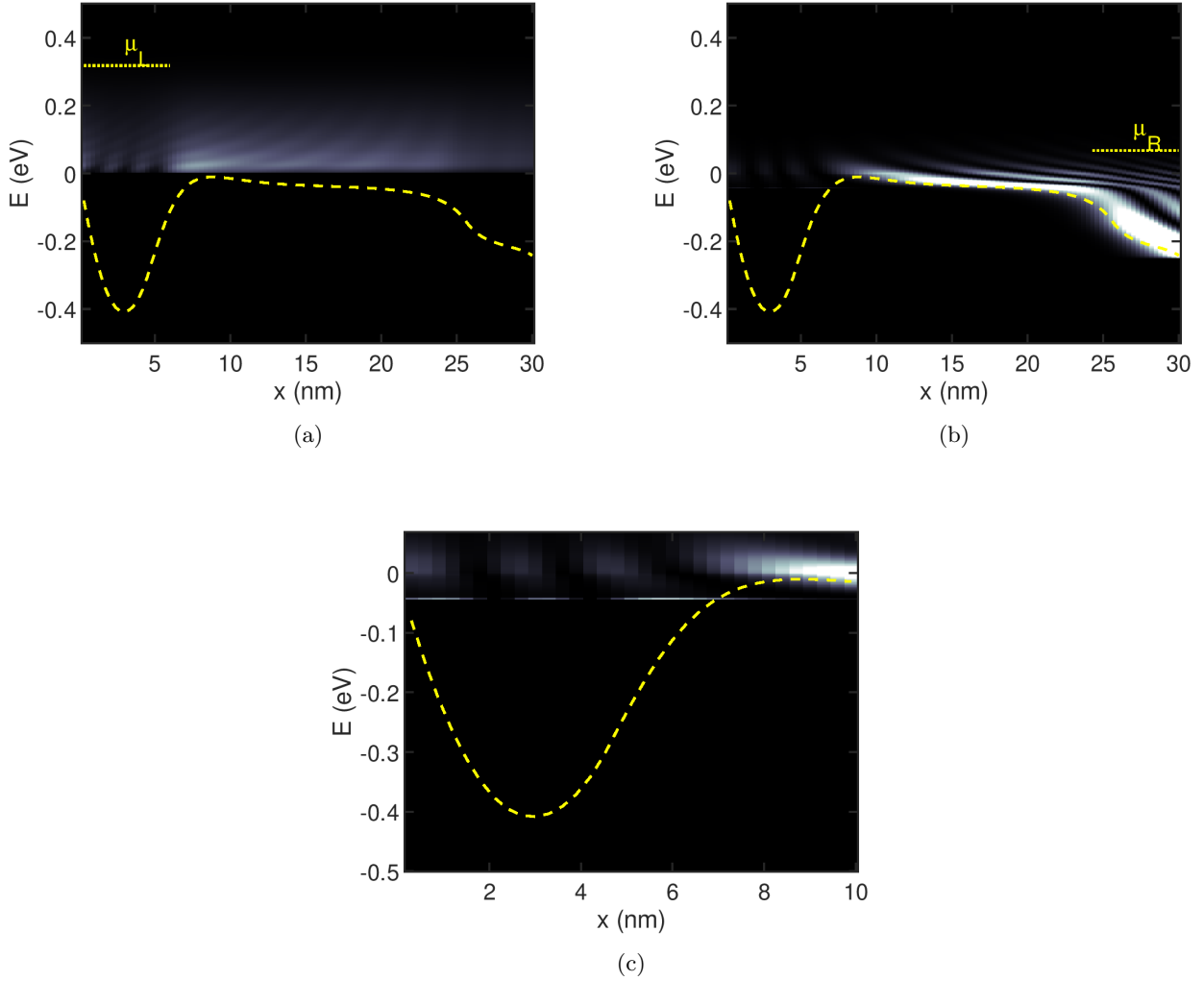
Figure 4.2: Partial local densities of charge for Dirichlet boundary conditions with $V = 0.25$ V. Electrons injected by (a) the source (left electrode) and (b) the drain (right electrode). A zoom of the potential well in (b) is shown in (c).

The contrast of the potential energy and the carrier density between both types of boundary conditions is apparent in fig. 4.3a and fig. 4.3b. However, until U_0 became the maximum value of U, the current is almost the same for both cases (fig. 4.3c). This effect can be explained by the high charge density in the channel, that partially shielded the interior region from the influence of the boundary conditions. Consequently, for $V < 0.25$ the position and the height of $\max(U)$ remained approximately unchanged comparing the two types of conditions.

Contrary to Neumann conditions, Dirichlet boundary conditions impose that all voltage drop happens within the channel. However, they make the device heavily charged, breaking the expected charge neutrality. This can lead to convergence difficulties and can invalidate some hypothesis, due to the high electric field. Even though the I-V characteristic is almost the same for both options up to a certain bias voltage, Neumann boundary conditions proved to be more
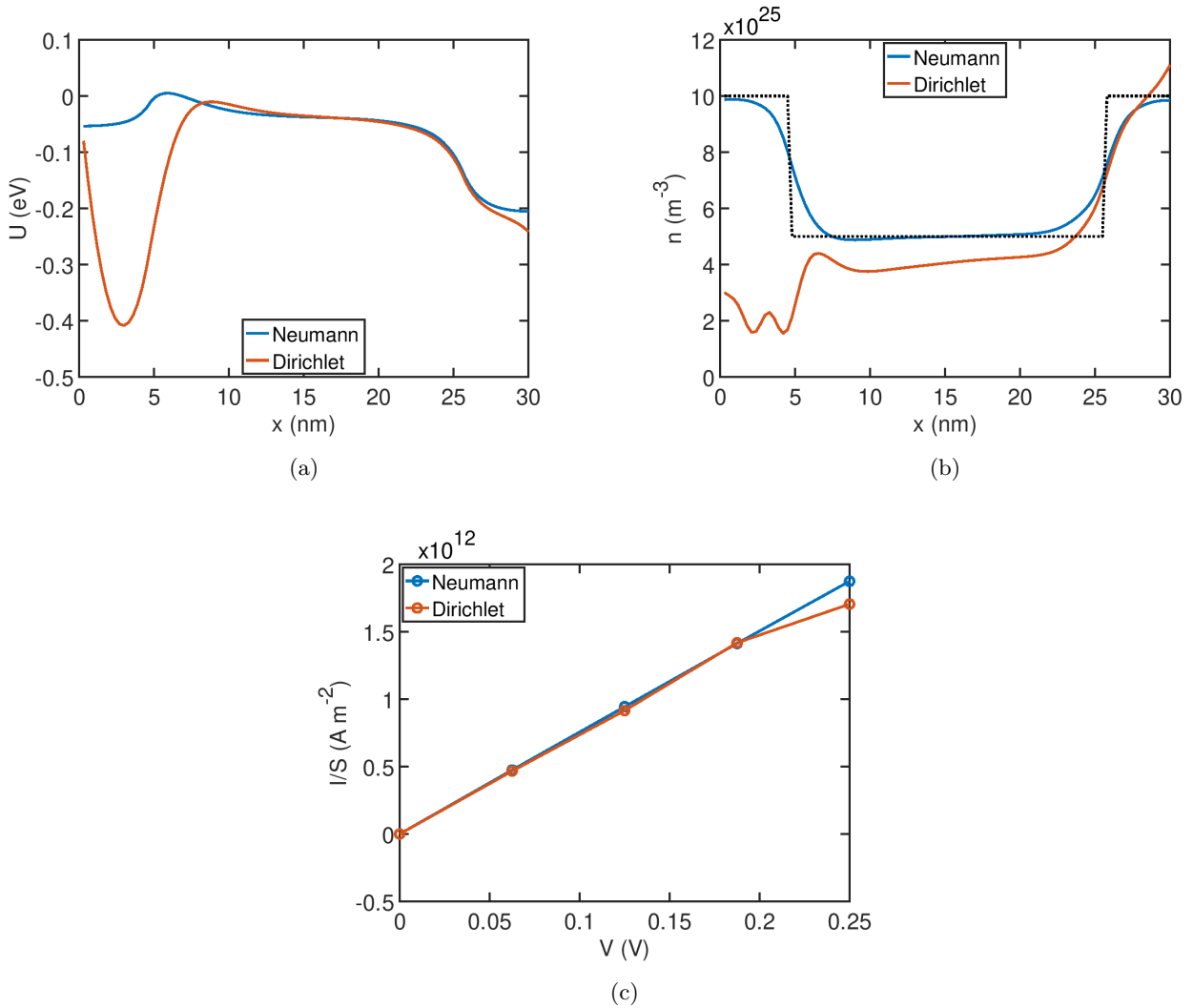
(a)

(b)

(c)

Figure 4.3: Comparison of non-equilibrium self-consistent simulations using Neumann and Dirichlet boundary conditions for $V = 0.25$ V. (a) Conductance Band, (b) Electrons concentration (traced-line indicates donor density) and (c) I-V characteristics.

appropriated, because they guarantee an almost neutral-channel as expected and do not cause some convergence difficulties caused by Dirichlet conditions.

4.2   Analysis of Current Transport in a Resonant-Tunneling Diode

In this section, we present an analysis of the current transport in a resonant-tunneling diode using the NEGF tools. As mentioned in section 2.4, current transport in a RTD is dictated by the alignment of a quasi-bound state in the central region with the source Fermi level. Therefore, the local density of states provides a very useful visualization of these mechanism, improving the understanding of the device's operation. Here we employed the ideal potential profile introduced in 3.3 to solve the transport equation, without the need of a self-consistent solution. Both Fermi
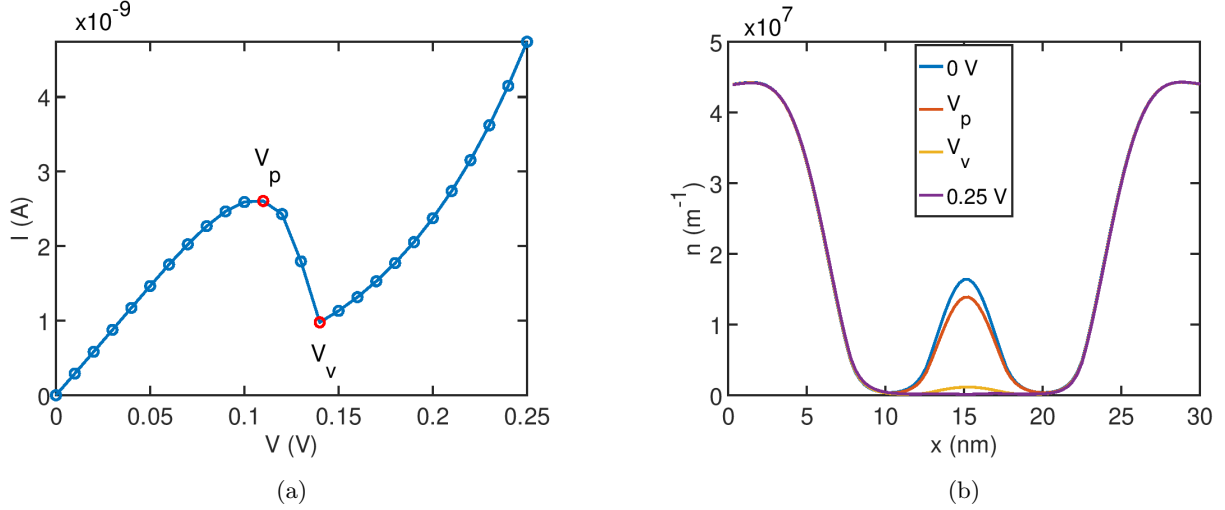
Figure 4.4: (a) Current-voltage characteristic and (b) electron density for chosen voltages for the simulated resonant-tunneling diode.

Effective Mass	$m = 0.10m_0$
Temperature	$T = 290$ K
Source Fermi Level	$\mu_L = 0$ eV
Drain Fermi Level	$\mu_R = -V$ eV
Number of x points	$N = 100$
Number of E points	$N_E = 10000$
Step in x	$a = 0.3$ nm

Table 4.2: List of simulation parameters for the resonant-tunneling diode.

levels were aligned with the conduction band, i.e. $\mu_L = 0$ eV and $\mu_R = -V$ eV.

Fig. 4.4a shows the obtained I-V curve, which has a behavior similar to what was expected. In summary, current increase until a peak for the applied voltage of V_p and decreases after that, reaching a valley for $V = V_v$. For $V > V_v$, current increases again. Fig 4.4b illustrates the electron density for chosen bias voltages. For these voltages, we present the local density of states in fig. 4.5 and the transmission coefficient in fig. 4.6. The parameters used are presented in tab. 4.2.

In the equilibrium case (fig. 4.5a), the fundamental state is found between the barriers for $E \approx 0.7$ eV. As a result, a sharp peak of transmission appears for this energy (fig. 4.6a). With $\mu_L = \mu_R$, this state is equally occupied by electrons injected from both contacts, creating a local maximum in the carrier density, but the net current carried by these particles is zero.

With a small applied voltage, the fundamental state began to be mainly occupied by electrons from the source, since its energy distanced from μ_R and approximated from μ_L. The current reaches a local maximum when this energy level is almost aligned with μ_L and reasonably far from μ_R. At this point, the transmission peak is inside the range of energies which can carry current (fig. 4.6b). As before, we can see a high density of electrons in the central region of the
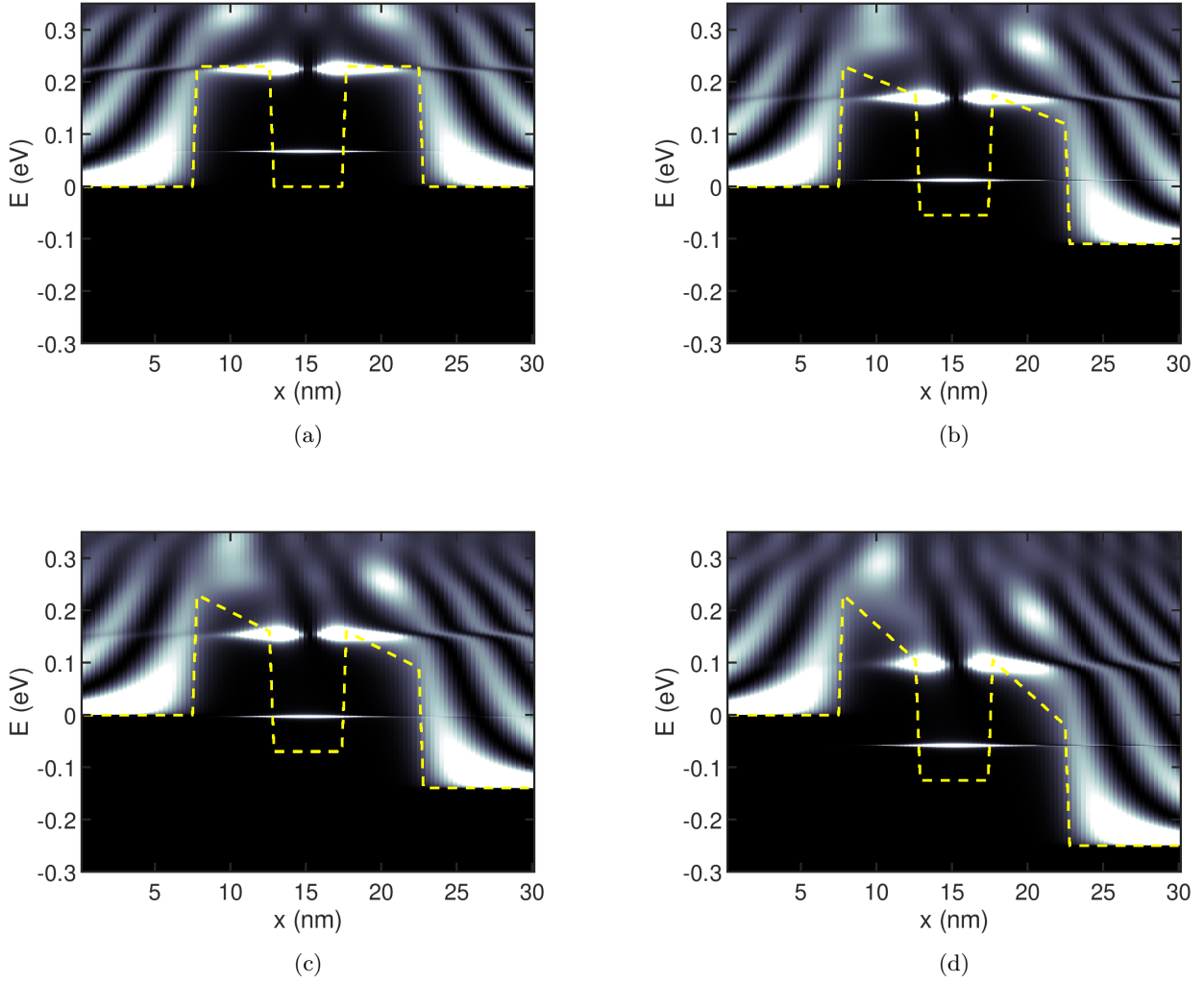
Figure 4.5: Local density of states of the resonant-tunneling diode for selected voltages. (a) 0 V, (b) $V_p = 0.11$ V, (c) $V_v = 0.14$ V and (d) 0.25 V.

channel, but now these charges contribute to the transport.

For $V > V_p$, the fundamental state cannot be occupied by electrons from the source, because the source cannot inject for energies below 0 eV. In this case, the central state is occupied by a few electrons injected from the drain, but they can not tunnel to the other side (fig. 4.5c). Because of that, the transmission coefficient does not present a peak for this level anymore (fig. 4.6c). The small current is due to a weak occupation of the second state, but this is not enough to be seen in the electron density.

For $V > V_v$, the second resonant got closer to μ_L, and a very high current started to flow 4.4a. This can be easily seen in the transmission coefficient for $V = 0.25$ V (fig. 4.6d), which has a large portion inside the range of energies that carry current. Despite that, this state is not occupied by a large number of electrons at this bias, since it is filled only by few electrons from the source.
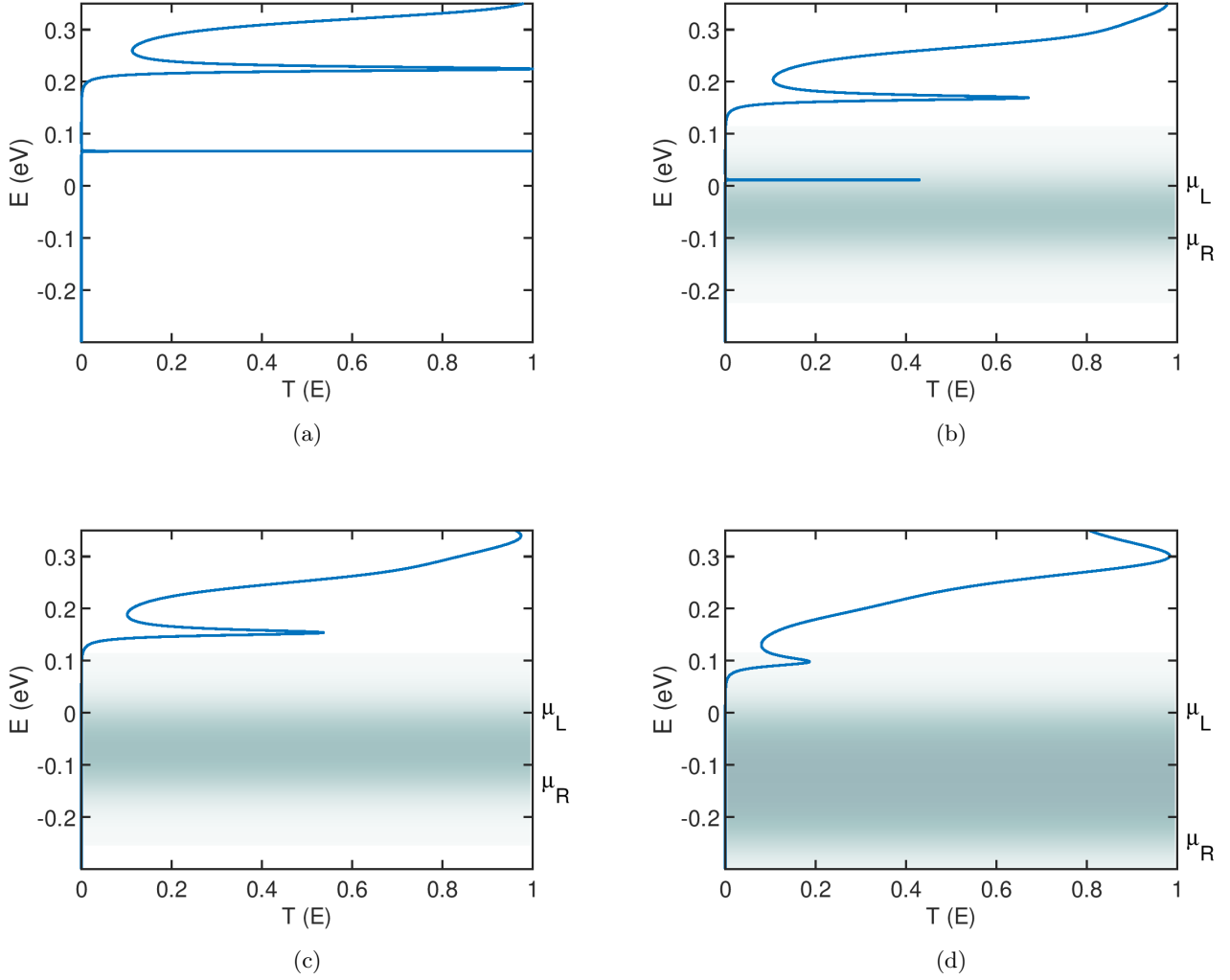
Figure 4.6: Transmission coefficient for the resonant-tunneling diode for selected voltages. (a) 0 V, (b) $V_p = 0.11$ V, (c) $V_v = 0.14$ V and (d) 0.25 V. Current-carrying energies are indicated by the blue regions.

4.3 Quantum Tunneling in Ultrathin Transistors with Multiple Gates

In this section we employed the non-equilibrium Green function method to discuss the operation of the nanoFETs introduced in section 2.4, in which intra-band and inter-band (band-to-band) tunneling play an essential role. To account for this second phenomenon, the effective potential approach presented in section 3.2 was used. Both conduction and valence particles were considered to have equal effective-masses. For these simulations, the transport equation was solved for given energy diagrams, whose calculation has been described in [21, 26]. Other simulations, using the ideal potentials proposed in 3.3, were made to investigate the switching process in the multi-mode 2G nanoFET.
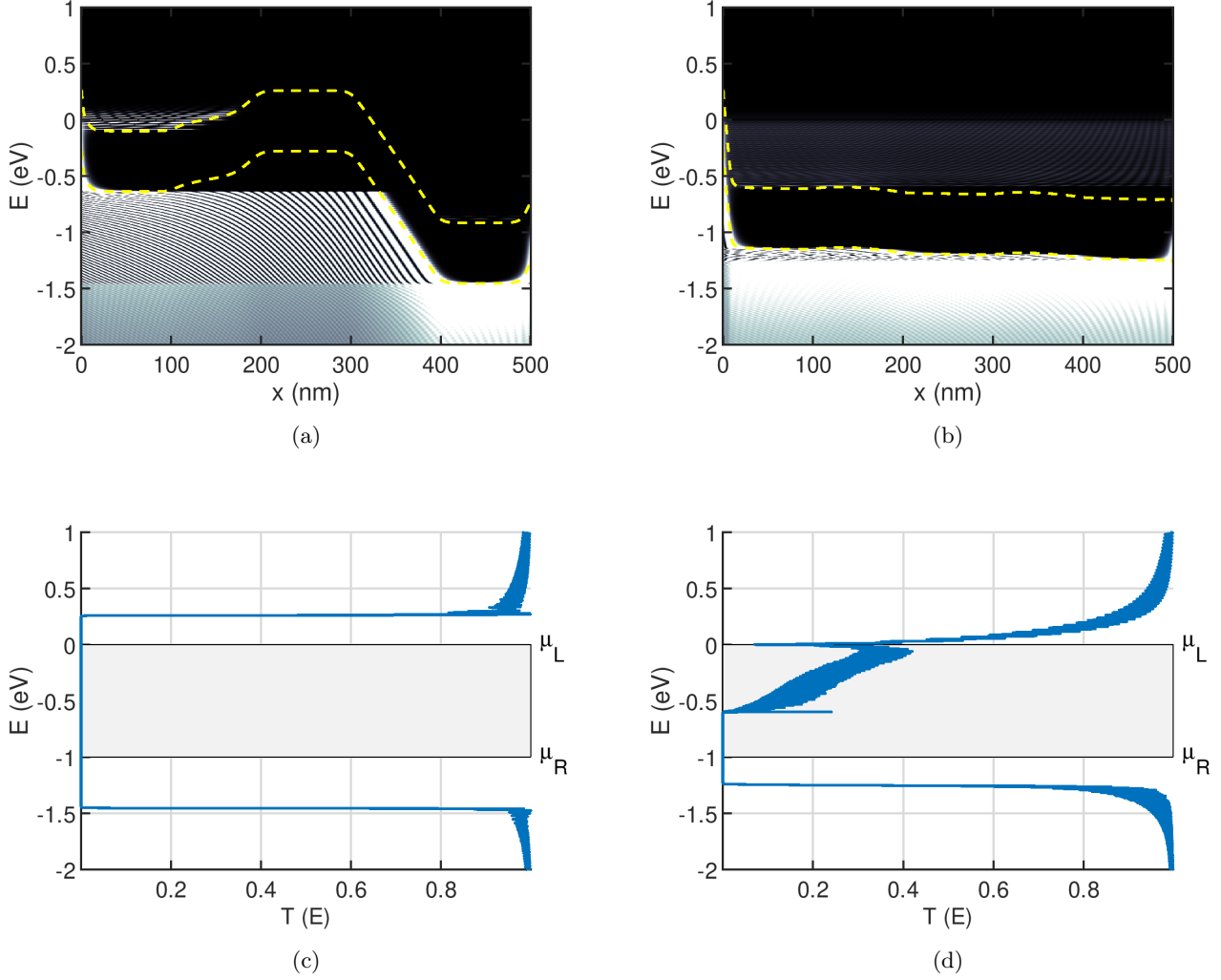
(a)

(b)

(c)

(d)

Figure 4.7: Local densities of carriers and transmission coefficients for 3GR n-FET: (a, c) off-state, (b,d) on-state.

Polarity control in a reconfigurable nanoFET with three top gates

The 3G nanoFET, whose structure is depicted in figure 2.8a, can be dynamically configured as an n-FET or a p-FET by applying a bias of $V_{PGS} = +1.2$ V or -1.2 V to both program gates, respectively (see sec. 2.4). The local densities of charges and energy diagrams for n-type operation, with $V_{DS} = 1$ V, are shown in fig. 4.7. Fermi levels μ_L and μ_R are placed in the middle of the channel semiconductor bandgap. In this case, PGS decreases the thickness of the Schottky barrier to allow tunneling as the main injection mechanism. For the off-state (fig. 4.7a), when $V_{CG} = 0$ V, the injected electrons from the left electrode are trapped in the well formed between PGS and CG. These charges will contribute for the carrier concentration, but will not influence current, because they cannot tunnel through the large potential barrier at the center of the channel. Moreover, the huge bandgap barrier between CG and PGD does not allow btb tunneling for electrons injected in the valence band.

In the on-state (fig. 4.7b), when $V_{CG} = 1$ V, the control gate shifts the energy bands at the center in order to pull down the barrier. At the same time, the Schottky barrier becomes even thinner, increasing the tunneling probability. Now electrons injected by the source in energies between 0 eV and -0.6 eV create a current flow. The transmission coefficients for both states are shown in figs. 4.7c and 4.7d . For the off-state, we can see that only states with energy above the central potential barrier allow transmission, but they are not occupied by electrons. For the on-state, transmission is allowed in the current-carrying energies.

The local densities of charges and the transmission coefficients for p-FET configuration are shown in fig. 4.8. In figs. 4.8a and 4.8b, holes are located in the weakly occupied states between μ_L and the valence band profile. Arguments analogous to those used above explain the transport of charge in this configuration, but now the conduction through the valence band is caused by holes. The symmetric characteristics for n- and p-FET arise from the equal effective-mass used for the two bands.

Power control in a multi-mode nanoFET with two top gates

Fig. 2.9a depicts a nanoFET with two top gates, which can operate in a high-performance mode (HP) or in a low-power mode (LP). The off and on local densities of carriers and the energy diagrams for HP, with $V_{DS} = 0.5$ V, are shown in fig. 4.9. In the HP mode, chosen by setting $V_{PG} = +1V$, the transistor operates in a similar way as a 3G nanoFET in n-type configuration. In other words, CG controls if the injected electrons are trapped between PG and CG (fig. 4.9a), blocking current flow, or if they are allowed to transmit through the channel (see fig. 4.9b). The first situation wherein no transmission is possible at current-carrying energies can be seen in fig. 4.9c. The second situation can be seen in fig. 4.9d. In this case, electrons can reach the other side.

In the LP mode, the device operation is dictated by another switching mechanism. Fig. 4.10 shows the local densities of carriers and the energy diagrams for this mode. The application of a bias $V_{PG} = -1$ V creates a well that traps electrons injected in the valence band. If $V_{CG} = 0$ V, those electrons stay confined between PG and CG and the transistor is off (fig. 4.10a). On the other hand, if $V_{CG} = 1$ V, the bandgap barrier becomes thinner, allowing the trapped charges to tunnel from the valence band to the conduction band. This turns the transistor on. The formation of these states responsible for btb tunneling can be appreciated comparing the figs. 4.10c and 4.10d, in which they appear as small peaks between μ_L and μ_R.

The carriers effective-mass significantly influences the transmission probability for the LP device at energies for which tunneling is possible. To exemplify, $LDOC$ and $T(E)$ are shown in fig. 4.11 for two different effective-masses. As we can see, $T(E)$ for $E > max(U_c)$ or $< min(U_v)$ is almost unitary for the two masses. However, at energies where tunneling is necessary for transmission, the result for $m = 0.01m_0$ presents much bigger peaks than that for $m = 0.05m_0$. The reason is that the transmission probability scales with $\exp\left(-\sqrt{m/m0}\right)$ and, for these energies, the carriers must tunnel twice. From electrode to the channel, through the Schottky barrier, and through the btb tunnel barrier.
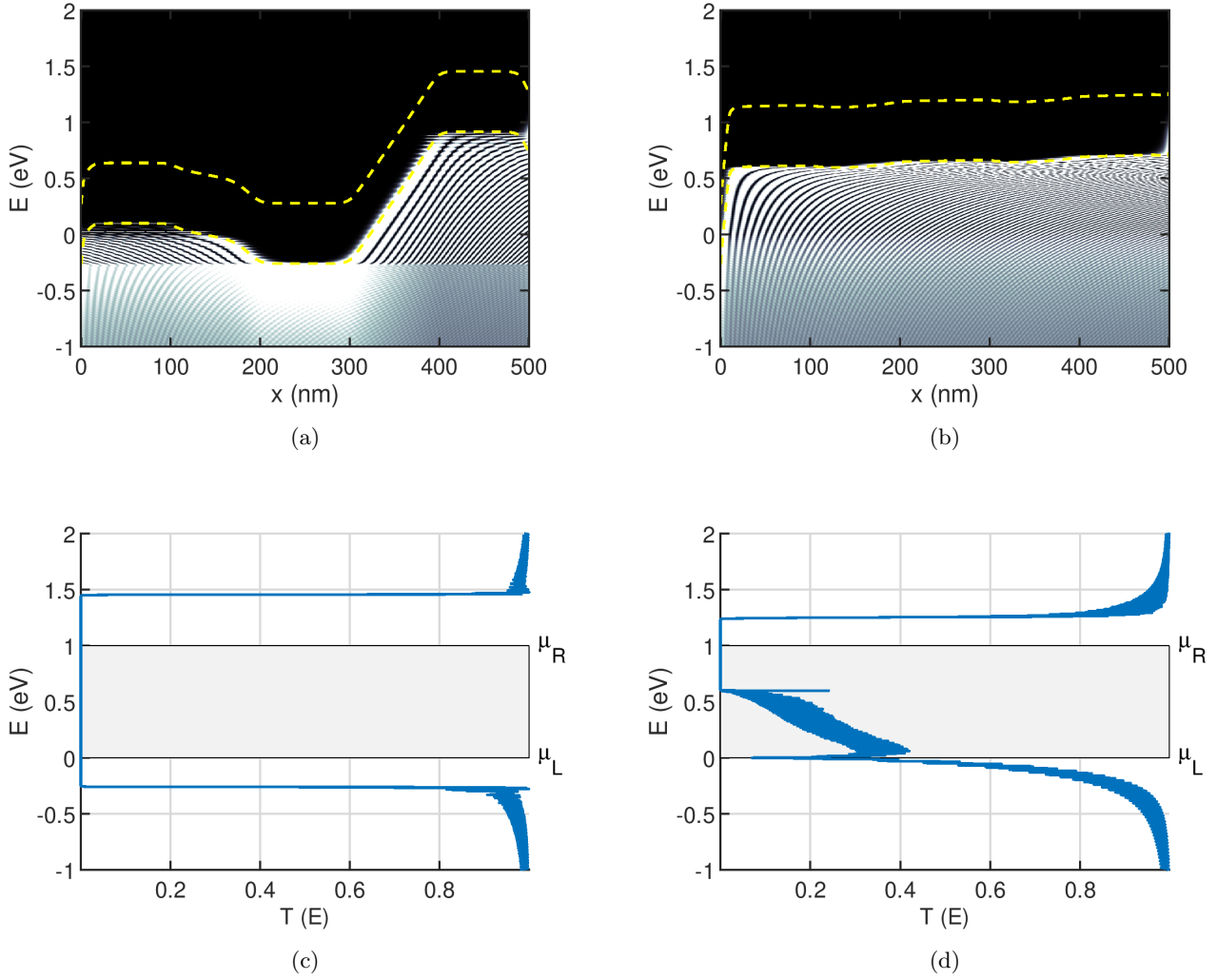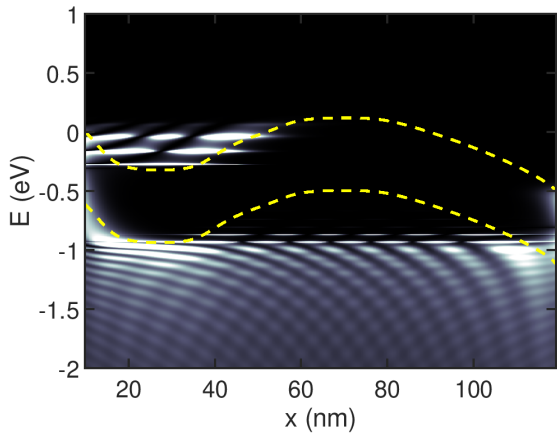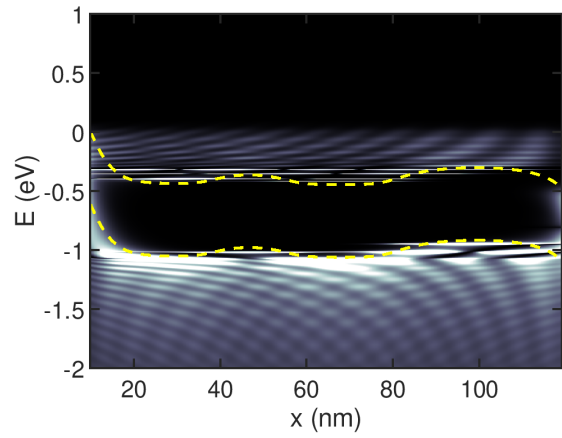
Figure 4.8: Local densities of carriers and transmission coefficients for 3GR p-FET: (a, c) off-state, (b,d) on-state.

The ideal potential profile introduced in 3.3 was used to study the on-off switching of both HP and LP modes and to further explore the effective-mass influence. Fig. 4.12 shows transfer characteristics comparisons between HP and LP with $m = 0.06 m_0$ and LP with $m = 0.01 m_0$ and $m = 0.06$. These results have the same behavior of those obtained by self-consistent solutions, despite all simplifications. The HP mode presents a much bigger on-current than the LP mode. On the other hand, for the LP mode when switching off the transistor, the current decreases nearly 10^8 times. In other words, HP can deliver a higher current, but LP provides a much more efficient on-off switching. As we can see in fig. 4.12b, the characteristics of LP become closer to those of HP with the effective-mass decrease.
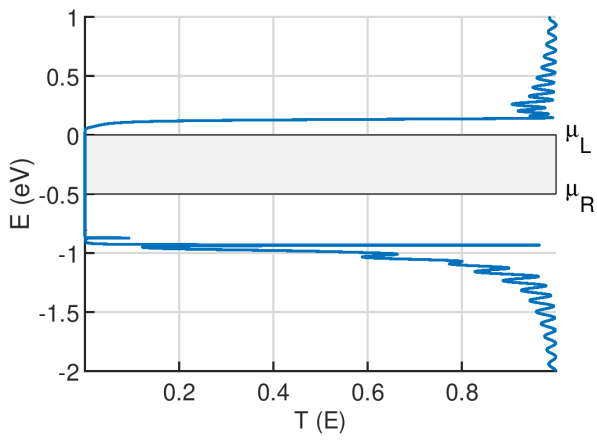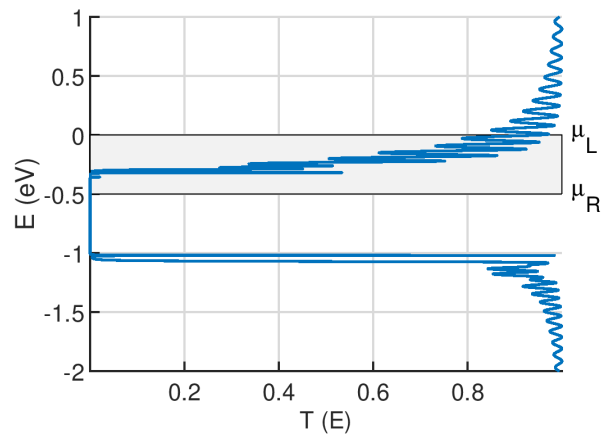
Figure 4.9: Local densities of carriers and transmission coefficients for 2G HP-mode: (a, c) off-state, (b,d) on-state.
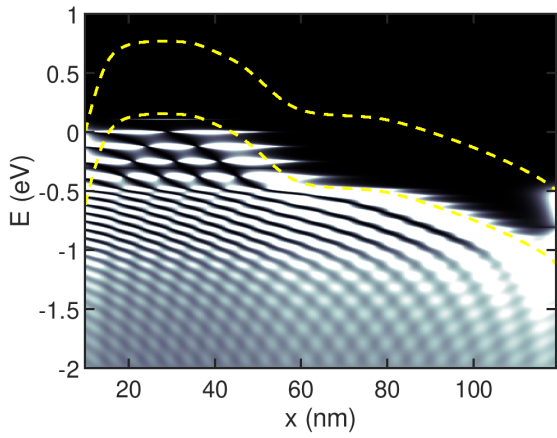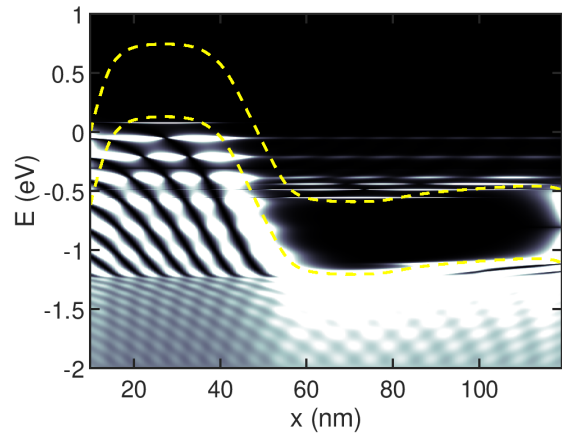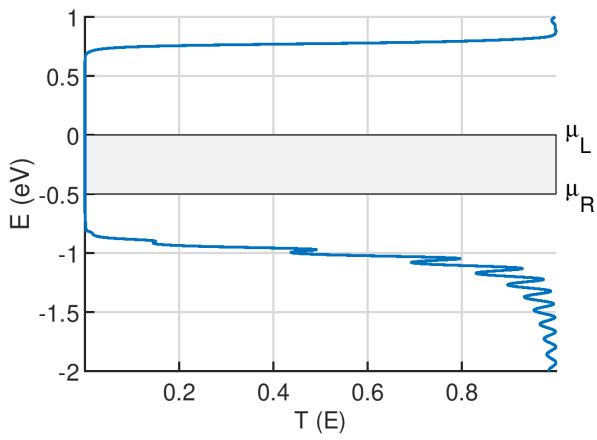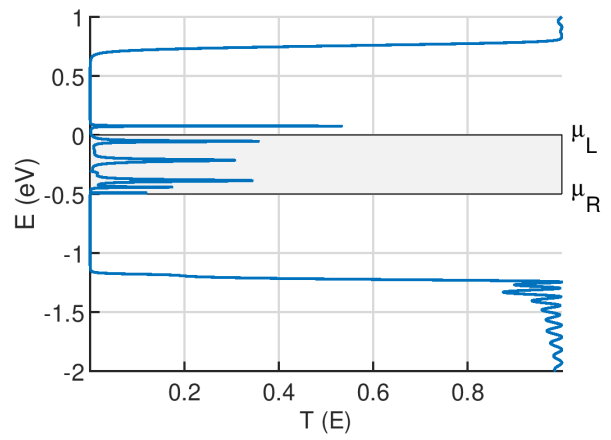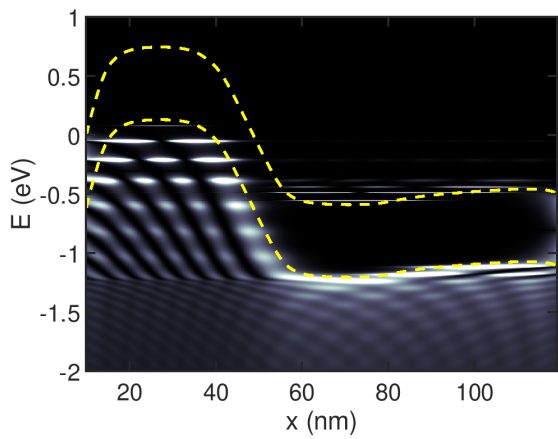
(a)

(b)

(c)
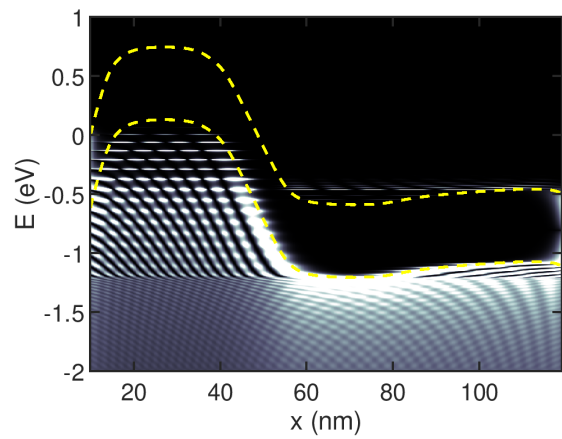

(d)

Figure 4.10: Local densities of carriers and transmission coefficients for 2G LP-mode: (a, c) off-state, (b,d) on-state.
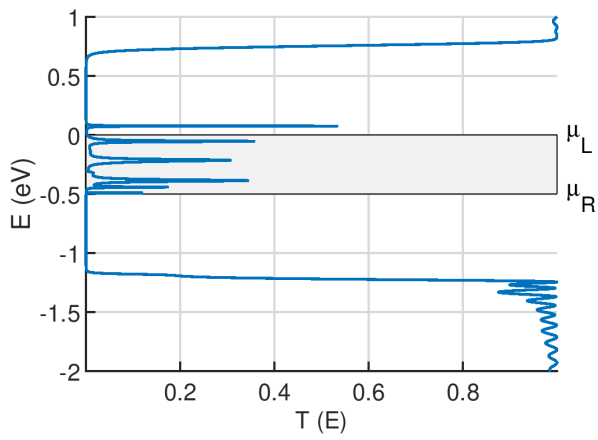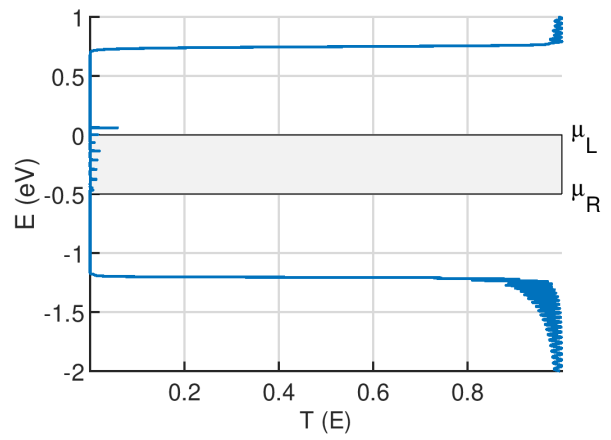
Figure 4.11: Local densities of carriers and transmission coefficient for different effective-masses. (a,c) $m = 0.01m_0$ and (b,d) $m = 0.05m_0$.
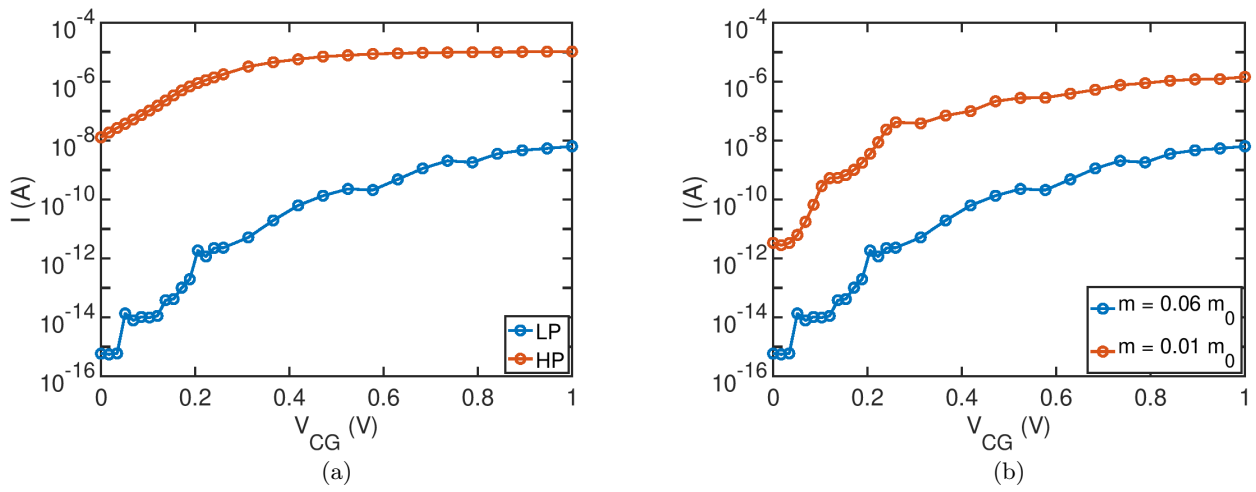
Figure 4.12: Transfer characteristics for the multi-mode nanoFET calculated with idealized energy diagrams. (a) Comparison between LP and HP for $m = 0.06m_0$. (b) Comparison between $m = 0.01m_0$ and $m = 0.06m_0$ for LP mode.

Chapter 5

Conclusions

There is no doubt that nanoscale devices will be an important part of the semiconductor industry, complementing and partially substituting Si-based technology. These devices hold the promise to be more energy-efficient and to allow higher integration at a lower cost. On the other hand, their performance is very sensitive to variations in the fabrication process. Therefore, a new generation of TCAD tools based on quantum mechanical simulations is needed to conduct the evolution of the nanoelectronics field.

In the present work, we presented the first steps toward the implementation of a new atomistic-simulator based on the non-equilibrium Green function method. This formalism provides a mathematically effective way to simulate current transport in nanoscale devices. Simulations of three ballistic nanodevices were discussed. First, we made self-consistent simulations of a short Si-resistor to study the impacts of using different types of boundary conditions for the Poisson equation. As a result, we found that Dirichlet boundary conditions can make the channel heavily charged, breaking the expected charge neutrality. In the second example, we applied the NEGF method to describe the current transport in a resonant-tunneling diode. NEGF provided a powerful visualization of the transmission mechanisms. Finally, we studied quantum effects present in ultra-thin field-effect transistors that offer a trade-off between energy and performance and can be reconfigured at run-time to operate as n-/p-type. In conclusion, we consider that the above examples illustrated some of the basic possibilities of the NEGF formalism when applied to simulations of ballistic devices.

Our future research will focus on implementing methods to decrease the time spent in the calculation of the matrix Green function and models to describe inelastic scattering and extended contacts. To achieve the first goal, an iterative process to obtain $G(E)$, called Dyson Equation, will be applied [14]. It can significantly reduce the time needed to calculate $G(E)$ when the matrix has a huge number of elements. Additionally, an adaptive integration scheme will be implemented together with a contour integration method to calculate the density matrix with more accuracy using less energy points [22]. In this way, the matrix inversion in (2.4) can be made less times during the simulations. Inelastic scattering events will be implemented as described in [27, 17] to expand the range of devices that can be simulated. For example, such effects can be crucial in

self-consistent simulations of resonant-tunneling diodes [16]. Finally, models for extended contacts will be studied in order to simulate effects as those discussed in [5]. Both scattering and extended contacts are described via the self-energy matrices with a more complex form.

Bibliography

[1] WALDROP, B. Y. M. M.; LAW, O. F. M. S.; INTERESTING, M. More Than Will Soon Abandon Its Pursuit. *Nature*, v. 530, p. 145, 2016.

[2] DATTA, S. Nanoscale device modeling: the Green's function method. *SUPERLATTICES AND MICROSTRUCTURES*, ACADEMIC PRESS LTD, 28, n. 4, p. 253–278, OCT 2000.

[3] KIM, Y.-B. Challenges for Nanoscale MOSFETs and Emerging Nanoelectronics. *Transactions on Electrical and Electronic Materials*, v. 11, n. 3, p. 93–105, 2010. ISSN 1229-7607.

[4] FEDIAI, A. et al. Towards an optimal contact metal for CNTFETs. *The Royal Society of Chemistry*, v. 8, p. 10240–10251, 2016.

[5] CUMMINGS, A. W.; LÉONARD, F. Enhanced Performance of Short-Channel Carbon Nanotube Field-Effect Transistors Due to Gate-Modulated Electrical Contacts. *ACS Nano*, v. 6, n. 5, p. 4494–4499, 2012.

[6] NEOPHYTOU, N.; AHMED, S.; KLIMECK, G. Non-equilibrium Green's function (NEGF) simulation of metallic carbon nanotubes including vacancy defects. *Journal of Computational Electronics*, v. 6, n. 1-3, p. 317–320, sep 2007. ISSN 15698025.

[7] SARANGAPANI, P. et al. Non-equilibrium Green's function predictions of band tails and band gap narrowing in III-V semiconductors and nanodevices. apr 2019. Disponível em: <http://arxiv.org/abs/1904.07458>.

[8] KUBIS, T. et al. Theory of nonequilibrium quantum transport and energy dissipation in terahertz quantum cascade lasers. *Physical Review B - Condensed Matter and Materials Physics*, v. 79, n. 19, p. 1–10, 2009. ISSN 10980121.

[9] XUE, Y.; DATTA, S.; RATNER, M. A. First-principles based matrix Green's function approach to molecular electronic devices: General formalism. *Chemical Physics*, v. 281, n. 2-3, p. 151–170, aug 2002. ISSN 03010104.

[10] PECCHIA, A.; Di Carlo, A. *Atomistic theory of transport in organic and inorganic nanostructures.* aug 2004. 1497–1561 p.

[11] AEBERHARD, U. Theory and simulation of quantum photovoltaic devices based on the nonequilibrium Green's function formalism. *Journal of Computational Electronics*, v. 10, n. 4, p. 394–413, dec 2011. ISSN 15698025.

[12] SRIMANI, T. et al. Asymmetric gating for reducing leakage current in carbon nanotube field-effect transistors Asymmetric gating for reducing leakage current in carbon nanotube field-effect transistors. AIP Publishing LLC, v. 063107, n. June, 2019.

[13] NANOHUB. https://nanohub.org/. Accessed: 2019-12-08.

[14] RYNDYK, D. *Theory of Quantum Transport at Nanoscale: An Introduction.* [S.l.]: Springer International Publishing, 2015. (Springer Series in Solid-State Sciences). ISBN 9783319240886.

[15] GRIFFITHS, D. *Introduction to Quantum Mechanics.* [S.l.]: Pearson Prentice Hall, 2005. (Pearson international edition). ISBN 9780131118928.

[16] FERRY, D.; GOODNICK, S.; BIRD, J. *Transport in Nanostructures.* [S.l.]: Cambridge University Press, 2009. (Transport in Nanostructures). ISBN 9780521877480.

[17] DATTA, S. *Quantum Transport: Atom to Transistor.* [S.l.]: Cambridge University Press, 2013. ISBN 9781107632134.

[18] HAŁDAŚ, G.; KOLEK, A.; TRALLE, I. Modeling of mid-infrared quantum cascade laser by means of nonequilibrium Green's functions. *IEEE Journal of Quantum Electronics*, v. 47, n. 6, p. 878–885, 2011. ISSN 00189197.

[19] STREETMAN, B.; BANERJEE, S. *Solid State Electronic Devices, Global Edition.* [S.l.]: Pearson Education Limited, 2015. ISBN 9781292060767.

[20] KAUSHIK, B. K.; DASGUPTA, S.; PAL, P. K. Introduction to Nanoelectronics. In: . [S.l.]: CRC Press, 2017.

[21] MOURA, R. et al. Reconfigurable NanoFETs: Performance Projections for Multiple-Top-Gate Architectures. *IEEE Transactions on Nanotechnology*, v. 17, n. 3, p. 467–474, 2018. ISSN 1536125X.

[22] BEVAN, K. H. *First Principles Non-Equilibrium Green's Function Modeling of Vacuum and Oxide Barrier Tunneling.* Tese (Doutorado) — Purdue University, West Lafayette, Indiana, 8 2008.

[23] POURFATH, M. *The Non-Equilibrium Green's Function Method for Nanoscale Device Simulation.* [S.l.]: Springer, 2014. ISBN 9783709117996.

[24] ROBERTSON, D. G. *Relaxation Methods for Partial Differential Equations: Applications to Electrostatics.* [S.l.], 2010.

[25] CLAUS, M. et al. COOS: A wave-function based Schrödinger-Poisson solver for ballistic nanotube transistors. *Journal of Computational Electronics*, v. 13, n. 3, p. 689–700, 2014. ISSN 15728137.

[26] BLAWID, S. et al. Performance Projections for a Reconfigurable Tunnel NanoFET. *IEEE Journal of the Electron Devices Society*, v. 5, n. 6, p. 473–479, 2017. ISSN 21686734.

[27] ANANTRAM, M. P.; LUNDSTROM, M. S.; NIKONOV, D. E. Modeling of nanoscale devices. *Proceedings of the IEEE*, v. 96, n. 9, p. 1511–1550, sep 2008. ISSN 00189219.