Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 202 (2005) 150–180
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Subband decomposition approach for the simulation of quantum electron transport in nanostructures

E. Polizzi [a,*], N. Ben Abdallah [b]

[a] Department of Computer Sciences, Purdue University, 250 N. University street, West-Lafayette, IN 47907-2066, USA
[b] Laboratoire MIP, UMR CNRS 5640, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex 4, France

Received 19 January 2004; received in revised form 16 June 2004; accepted 3 July 2004
Available online 17 August 2004

Abstract

The modeling of ballistic quantum transport in ultimate size semiconductor devices usually involves a self-consistent solution between the Schrödinger and the Poisson equations. In the 2D or 3D real space, this procedure requires huge computer resources to obtain the I–V characteristics. The general approach proposed in this article relies on the decomposition of the wave function on subband eigenfunctions, which account for the confinement of the electrons in the whole structure. The method can be applied to study large 2D and 3D real systems with a drastic reduction of the numerical cost, since the dimension of the transport problem for the Schrödinger equation is now reduced in real space. The results obtained for the 2D nanoscale MOSFETs show the efficiency of the algorithm and allow to estimate the effects of the coupling between the subbands. The asymptotic approach of the subband decomposition is also presented for devices showing a strong confinement for the electron gas as the 3D electron waveguide devices.

Keywords: Schrödinger–Poisson system; NEGF; Mode decomposition; Nanoscale MOSFETs; Electron waveguide devices; Quantum ballistic transport; Open nanoscale semiconductor devices

1. Introduction

Nowadays, a great challenge consists in the down-scaling of electronic components at the nanometer scale with the aim to obtain high speed and high functionality devices. In this task, modeling and numerical simulations play an important role in the determination of the limit size of the nanoscale MOSFETs [1–4] as well as the design of new devices whose electron transport properties are mainly based on quantum

* Corresponding author. Tel.: +1-765-494-6012; fax: +1-765-494-0739.
E-mail addresses: epolizzi@purdue.edu (E. Polizzi), naoufel@mip.ups-tlse.fr (N. Ben Abdallah).

0021-9991/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2004.07.003

effects, such as the electron waveguide devices [5–12]. In these ultimate size devices, the electron transport is expected to operate in a quantum ballistic regime in the active region outside of the electron reservoirs [13–15]. The model used in this article is defined by a self-consistent process between the calculation of the electron density and the space charge effects using the Poisson's equation. In the device, the electrons are in a mixed state with given statistics. The elementary states depend on a continuous index and are solutions of the stationary Schrödinger equation with open boundary conditions at the contacts between the device and the electron reservoirs. At each iteration step of the self-consistent process, the calculation of the electron density involves the computation of a large number of wave functions.

An alternative approach consists to solve the non-equilibrium Green's function equations (NEGF) self-consistently with the Poisson's equation, and we refer [16,17] where this method is applied, respectively, for the one gate and the two gates nanoMOSFETs. The Schrödinger–Poisson and NEGF-Poisson approaches are formally equivalent, since the self-energy functions which are involved within the NEGF formalism can also be considered as open boundary conditions for the Schrödinger equation [18,19]. However, the two approaches show some important differences: (i) in their numerical treatment [20], (ii) in their possibility to be extended to scattering problems (see [21] for the NEGF formalism, and [22] for the wave function formalism), (iii) in the derivation of open boundary conditions associated with complex structures for the reservoirs (beyond the semi-infinite waveguides and the parabolic dispersion relation assumption). All these points will be developed in a future article.

In a previous article [23], our parallel finite element simulator NESSIE was presented to simulate the electron transport in 3D nanodevices using a full self-consistent quantum model. In the 2D and 3D systems, this self-consistent approach usually requires huge computer resources due to the large size and the number of the linear systems arising from the transport problem. A quasi-3D model has been proposed to overcome these difficulties, and it has been successfully applied to electron waveguide devices showing a strong confinement for the electron gas. However, many other semiconductor devices involve heterostructures where the electron gas is confined in one or many directions of the device. This article proposes to extend the functionalities of NESSIE to handle arbitrary structures using a general subband decomposition approach of the Schrödinger equation in the confined directions, while reducing the dimension of the transport problem in real space. Also, the derivation of the quasi-full dimensional model is presented as an asymptotic approach within this general framework.

Prior works on the decomposition by modes of the transport problem have been presented in the early 90s in [24–29]. A subdivision of the quantum point contact device by parallel slices along the transport direction was considered, while either the mode matching method or the recursive Green's function method [5,16,30,31] was essentially applied to solve the problem. In the mode matching method, the wave function is expanded in terms of exponentials, which is correct only if one assumes a stepwise potential between the slices. The original problem using the mode matching technique is then approximated at the physics level. The recursive Green's function approach is a numerical procedure often used to solve a tight-binding type model (i.e. a finite difference discretization in real space). However, this method has some numerical limitations for large structures, since it requires to invert a matrix at each slice (where the size is proportional to the number of discretization points in the slice) in a recursive way along the transport direction. Our approach consists rather to discretize the variational form of the problem using the finite element method, and to solve the obtained large sparse symmetric linear systems with a preconditioned Krylov subspace iterative method. In addition, the potential is computed within a fully self-consistent quantum model, while many authors compute the potential within a semi-classical self-consistent Thomas–Fermi approach (this approximation is only used as initial guess in our Schrödinger/Poisson solver). More recently in [17], the authors compare the results of their full 2D self-consistent quantum simulation on a double gate MOSFET, with those obtained using a space mode representation model which is based on a subband decomposition with no-coupling effects between the modes. However, we will show in this article that the effects of the coupling between the modes become important in the I–V characteristics for non-uniform structures.

Section 2 of this article consists in an overview of the full transport model in the real space, where the nature of boundary conditions in the device is discussed. The subband decomposition of the Schrödinger equation in the confined directions for the electrons is presented in Section 3. This approach will be applied to simulate two particular quantum devices: the nanoscale MOSFETs in Section 4 and the electron waveguide devices in Section 5.

2. Overview of the full model

Considering Ω a 3D box in the real space, the quantum devices nature is related to the boundary conditions for the wave function on $\partial\Omega$ in the directions x, y and z (see Fig. 1).

In the usual cases, one can define three kinds of boundary conditions such as the infinite, the confined and the open boundary conditions. The infinite boundary conditions are related to a translation invariance of the potential, where the wave function is defined as plane waves. The open boundary conditions take into account the motion of the electrons in the structures and they are usually associated to the contact of the device with an external electron reservoir. Finally, the confined boundary conditions are defined when the probability to find an electron outside the device is equal to zero (this yields a quantization of energy levels). The first column of Table 1 lists different device configurations as function of the nature of the boundary conditions on each direction.

Denoting e_1 the confined direction(s) of the electron gas, e_2 the transport direction(s) where the motion of the electrons takes place, and e_3 the infinite direction(s), we write

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$$

with $e_1 \in \Omega_1$, $e_2 \in \Omega_2$ and $e_3 \in \Omega_3$. Due to the translation invariance of the problem in the e_3 direction (the existence of a e_3 infinite direction leads to $\Omega_3 =]-\infty,+\infty[$), the full wave function on Ω can be separated as

$$\Psi_E(e_1, e_2, e_3) = \mathrm{e}^{\mathrm{i}k_3 e_3}\psi_\varepsilon(e_1, e_2), \tag{1}$$

where the additional exponential part for the wave function corresponds to a plane wave associated to the wave vector k_3. The expression of the energy is given by

$$E = \varepsilon + \frac{(\hbar k_3)^2}{2m^*} \tag{2}$$

with m^* is the isotropic effective mass, and ε is the energy defined in the reduced domain $\Omega_{1,2} = \Omega_1 \times \Omega_2$. Finally, the wave function ψ_ε is solution of the following Schrödinger equation in $\Omega_{1,2}$

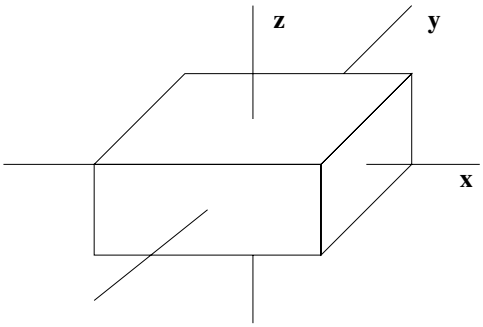

Fig. 1. A 3D box in the real space.

Table 1
List of different device configurations as function of the nature of boundary conditions on each direction

	Direction x	Direction y	Direction z	e_1	e_2	
Free electrons	Infinite	Infinite	Infinite			$\Psi_E = e^{ik_x x} e^{ik_y y} e^{ik_z z}$
1D devices	Open	Infinite	Infinite		x	$\Psi_E = e^{ik_y y} e^{ik_z z} \varphi(x)$
2D devices	Open	Infinite	Open		x, z	$\Psi_E = e^{ik_y y} \varphi(x,z)$
3D devices	Open	Open	Open		x, y, z	$\Psi_E(x,y,z) \equiv \varphi(x,y,z)$
0D quantum dots	Confined	Confined	Confined	x, y, z		$\Psi_E = \delta(E - E_i)\chi_i$ and $\{E_i, \chi_i(x,y,z)\}$
1D quantum wires	Infinite	Confined	Confined	y, z		$\Psi_E = e^{ik_x x} \chi_i$ and $\{E_i, \chi_i(y,z)\}$
2D electron gas	Infinite	Infinite	Confined	z		$\Psi_E = e^{ik_x x} e^{ik_y y} \chi_i$ and $\{E_i, \chi_i(z)\}$
3D waveguide devices I	Open	Open	Confined	z	x, y	$\Psi_E = \sum_i \varphi^i(x,y)\chi_i$ and $\{E_i(x,y), \chi_i(z;x,y)\}$
3D waveguide devices II	Open	Confined	Confined	y, z	x	$\Psi_E = \sum_i \varphi^i(x)\chi_i$ and $\{E_i(x), \chi_i(y,z;x)\}$
2D nanoscale MOSFETs	Open	Infinite	Confined	z	x	$\Psi_E = e^{ik_y y} \sum_i \varphi^i(x)\chi_i$ and $\{E_i(x), \chi_i(z;x)\}$

The wave functions Ψ_E are expressed within the subband decomposition approach using e_1 as the confined directions while e_2 is related to the transport directions. We denote by $\varphi$the wave function solution of the Schrödinger equation in the transport directions and $\{E_i(e_2), \chi_i(e_1;e_2)\}$ the eigenpairs of the Schrödinger equation in the confined directions.

$$-\frac{\hbar^2}{2m^*}\Delta_{1,2}\psi_\varepsilon(e_1,e_2) + U(e_1,e_2)\psi_\varepsilon(e_1,e_2) = \varepsilon\psi_\varepsilon(e_1,e_2), \tag{3}$$

where U is the energy potential, and we note $\Delta_{1,2} = \partial^2/\partial e_1^2 + \partial^2/\partial e_2^2$.

One way to solve this equation is to use the Quantum Transmitting Boundary Method in the $\Omega_{1,2}$ domain [32]. This method consists to compute the variational form of the problem using a finite element method. For this purpose, suitable open boundary conditions are derived for the contacts region with the reservoirs which are considered as semi-infinite leads (see Appendix A).

The full real space numerical treatment appears very time consuming because of both the size of problem (the number of discretization points in $\Omega_{1,2}$ may be high) and the large number of Schrödinger equations which are involved in the calculation of the electron density (integration over the energies). The subband decomposition approach presented in the next section is expected to significantly improve the computation times while reducing the size of the transport problem in real space.

3. Subband decomposition of the Schrödinger equation in the confined direction

In practice, the method requires that one can subdivide the structures into a large number of Ω_1 slices (1D or 2D) along the transport e_2 direction in Ω_2 then perpendicular to the current flow (see Fig. 2 for example).

We assume that the potential energy $U(e_1,e_2)$ is given on $\Omega_{1,2}$. Let $\chi_i(e_1;e_2)$ be the normalized eigenfunction solving the eigenvalue problems on Ω_1 in the e_1 confined direction which depend on the e_2 transport direction

$$-\frac{\hbar^2}{2m^*}\Delta_1\chi_i(e_1;e_2) + U(e_1,e_2)\chi_i(e_1;e_2) = E_i(e_2)\chi_i(e_1;e_2), \tag{4}$$

such as $\forall e_2$

$$\begin{cases} \int_{\Omega_1} \chi_i(e_1;e_2)\overline{\chi}_j(e_1;e_2)\,\mathrm{d}e_1 = \delta_{i,j}, \\ \chi_i(e_1;e_2) = 0 \quad \text{if } e_1 \in \partial\Omega_1, \end{cases} \tag{5}$$

therefore the reduced wave function ψ_ε can be expanded in the multi-mode basis χ_i

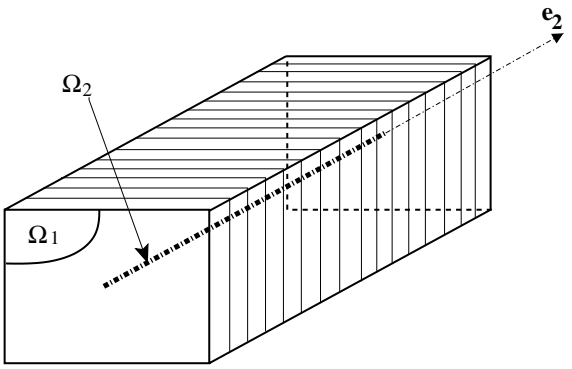

Fig. 2. In this 3D box, the domain $\Omega_{1,2}$ is subdivided into Ω_1 2D slices (confined domain associated to the confined e_1 directions) along the 1D transport e_2 direction.

$$\psi_\varepsilon(e_1,e_2)=\sum_{i=1}^{\infty}\varphi_\varepsilon^i(e_2)\chi_i(e_1;e_2). \tag{6}$$

The wave functions φ_ε^i satisfies the following coupled Schrödinger equation on Ω_2

$$-\Delta_2\varphi_\varepsilon^i(e_2)-2\sum_{j=1}^{\infty}\vec{c}_{ij}^{\,1}(e_2)\cdot\nabla_2\varphi_\varepsilon^j(e_2)-\sum_{j=1}^{\infty}c_{ij}^2(e_2)\varphi_\varepsilon^j(e_2)=\frac{2m^*}{\hbar^2}(\varepsilon-E_i(e_2))\varphi_\varepsilon^i(e_2), \tag{7}$$

where $\vec{c}_{ij}^{\,1}$ and c_{ij}^2 are extra-diagonal coefficients for $i\neq j$ given by

$$\vec{c}_{ij}^{\,1}=\int_{\Omega_1}\chi_i(e_1;e_2)\nabla_2\chi_j(e_1;e_2)\,\mathrm{d}e_1, \tag{8}$$

$$c_{ij}^2=\int_{\Omega_1}\chi_i(e_1;e_2)\Delta_2\chi_j(e_1;e_2)\,\mathrm{d}e_1. \tag{9}$$

One can show that the coefficient $\vec{c}_{ij}^{\,1}$ is equal to zero if $i=j$, and the expression of the c_{ij}^2 coefficients can also be written like

$$c_{ij}^2=\mathrm{div}_{e_2}(\vec{c}_{ij}^{\,1})-\int_{\Omega_1}\nabla_2\chi_i(e_1;e_2).\nabla_2\chi_j(e_1;e_2)\,\mathrm{d}e_1. \tag{10}$$

The subband decomposition approach can be applied to different devices configurations for which the expression of the wave functions are summarized in Table 1. For particular structures such as electron waveguide devices and nanoscale MOSFETS (in the last three rows of the table), one note that in addition to 1D (resp. 2D) parallel slices for a 2D (resp. 3D) structure, the subdivision of the devices along the transport direction can also be concerned with 1D slices into 3D structure. The expressions of the boundary conditions of the transport Schrödinger equation (7) are derived from the full problem and will be discussed for these particular types of devices in Appendix B.

To summarize, the wave function $\Psi_E(e_1,e_2,e_3)$ is obtained through the subband representation defined by Eqs. (1), (4)–(9). The non-infinite part $\psi_\varepsilon(e_1,e_2)$ is computed by solving eigenvalue problems in the e_1 confined directions and an open coupled Schrödinger equation along the e_2 transport directions. The overall problem is thus equivalent to solve only one Schrödinger equation in $\Omega_{1,2}$. However, the large number of the wave functions Ψ_E required to compute the electron density, involves the same basis functions $\chi_i(e_1;e_2)$ which is computed only once in the subband decomposition (for a given potential). Also, the cou-

pled Schrödinger equation (7) has to be solved numerically only for a finite and relevant number of coupling modes which depends on the energy level. Since the size of the matrix used to solve the wave function $\varphi_\varepsilon(e_2)$ will be then much smaller than the one used to solve directly the wave function $\psi_\varepsilon(e_1,e_2)$, the subband decomposition approach is then expected to require less computer resources than the full dimensional one.

4. The subband decomposition approach applied to the nanoscale MOSFETs

The most basic element in the design of a large scale integrated circuit is the transistor. The challenges of continued down-scaling of the metal oxide semiconductor field effect transistor (MOSFET) provide opportunities for nanoelectronic to improve the speed and efficiency of semiconductor materials. Our typical device of interest are very small Si/SiO_2 structures with very short channel(s) lengths where quantum transport is assumed to be ballistic (see for example Fig. 3 for the MOSFET with one and double gates).

The I–V characteristics of the devices are obtained by a self-consistent Schrödinger–Poisson model, using open boundary conditions for the scattering states at the source/drain interfaces (x_s and x_d) of the reduced domain $[x_s,x_d]\times[0,L_z]$.

In this approach, the positive doping profile regions at the source and drain contacts are equivalent to small electron reservoirs, in which we assume that the potential does not depend on the transport direction. Moreover, the model is invariant in the y direction (infinite boundary conditions) and the problem is then studied in the x,z domain. The presence of positive and negative doping profiles in the device may also be taken into account in the model. However, the holes transport due to a possible negative doping profile will be not considered (the value of this negative doping profile is very small compared to the positive one). Due to the anisotropic effects in silicon, we note m_t^* the transverse effective mass and m_l^* the longitudinal one. Also, three different electron configurations appear in the bandstructure within the parabolic approximation (Fig. 4).

Denoting $n_{x,y,z} \equiv n_{m_x^*,m_y^*,m_z^*}$ and $\vec{J}_{x,y,z} \equiv \vec{J}_{m_x^*,m_y^*,m_z^*}$, respectively, the electron and the current densities of one of this configuration, where m_x^* (respectively m_y^* and m_z^*) corresponds to the effective mass in the x direction (respectively y and z direction), then the electron and the current densities in the device are given by

$$n(x,z) = 2(n_{t,l,t}(x,z) + n_{l,t,t}(x,z) + n_{t,t,l}(x,z)), \tag{11}$$

and

$$\vec{J}(x,z) = 2\big(\vec{J}_{t,l,t}(x,z) + \vec{J}_{l,t,t}(x,z) + \vec{J}_{t,t,l}(x,z)\big), \tag{12}$$

where the factor 2 is due to the symmetry of each valleys (see Fig. 4).

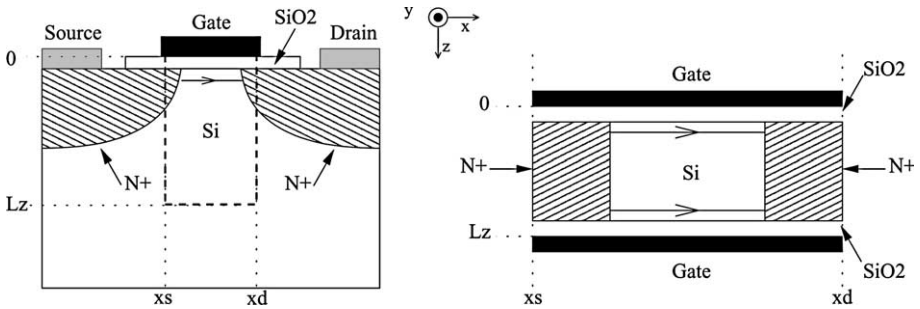

Fig. 3. Schematic view of the MOSFET with one gate (on the left) and double gates (reduced domain on the right). The electrons motion takes place in one channel in the simple gate MOSFET case, and in two channels in the double gate MOSFET case.

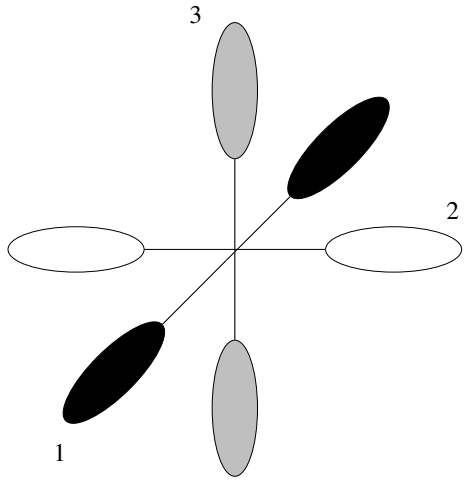

Fig. 4. Constant energy surface for the first conduction band in silicon (six ellipsoids and three different configurations for the electrons due to the symmetry properties).

A 2D Schrödinger–Poisson solver was implemented to solve this problem but the simulation time is high. Since the electron gas is confined in the z direction both in the reservoirs and along the channel(s) in the middle of the devices, it is suitable to use the subband decomposition. Therefore, the wave functions are expanded in a multi-mode basis which depends on the x transport direction. The problem consists now to solve many 1D eigenvalue problems in the z confined direction and a 1D coupled Schrödinger equation in the x transport direction.

In the following, we present a detailed study of the model used to solve the electron density for a particular effective mass configuration (m_x^*, m_y^*, m_z^*) as well as the most important steps of its numerical implementation. Finally, the results of the numerical simulations of the double-gate transistor will be discussed as well as the influence of the coupling modes within the subband decomposition approach.

4.1. Description of the model

As mentioned in Sections 2 and 3, we propose to apply the subband decomposition approach to our particular case of interest. The confined direction e_1 is now given by z, the transport direction e_2 is x and the infinite direction is y. In the 3D domain, the full Schrödinger equation is given by

$$(H + U(x,z))\Psi_E(x,y,z) = E\Psi_E(x,y,z), \tag{13}$$

where the Hamiltonian H is defined by

$$H = -\frac{\hbar^2}{2}\left(\frac{1}{m_x^*(z)}\Delta_x + \frac{1}{m_y^*(z)}\Delta_y\right) - \frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_z^*(z)}\frac{\partial}{\partial z}\right), \tag{14}$$

where the effective mass depends on the z direction due to the transition between the oxide and the silicon.

The expressions (1) and (2) of the wave function Ψ_E and the energy E in the full domain can be now written as

$$\Psi_E(x,y,z) = \mathrm{e}^{\mathrm{i}k_y y}\psi_\varepsilon(x,z) \qquad \text{with } E = \varepsilon + \frac{(\hbar k_y)^2}{2\overline{m}_y^*}, \tag{15}$$

where $\psi_\varepsilon(x,z)$ is solution of the 2D Schrödinger equation (3) with the open boundary conditions described in Appendix A. Also, in this case, a plane wave is associated to the translation invariance of the y direction and $\overline{m}_y^*$ is assumed to be a mean value of $m_y^*(z)$ on z.

For a given energy ε, the wave function ψ_ε can be written using the subband decomposition

$$\psi_\varepsilon(x,z) = \sum_i \varphi_\varepsilon^i(x)\chi_i(z;x), \tag{16}$$

where the $\chi_i(z;x)$ are the normalized eigenfunctions associated to the eigenvalues $E_i(x)$ solving the following 1D Schrödinger equation in the z confined direction in the $]0,L_z[$ domain

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_z^*(z)}\frac{\partial}{\partial z}\chi_j(z;x)\right) + U(x,z)\chi_j(z;x) = E_j(x)\chi_j(z;x). \tag{17}$$

Finally, the wave functions $\varphi_\varepsilon^i(x)$ satisfies the following coupled 1D Schrödinger equation in the $]x_s,x_d[$ domain

$$-\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi_\varepsilon^i(x) - 2\sum_{j=1}^{\infty} c_{ij}^1(x)\frac{\mathrm{d}}{\mathrm{d}x}\varphi_\varepsilon^j(x) - \sum_{j=1}^{\infty}\left(c_{ij}^2(x) + \frac{2}{\hbar^2}c_{ij}^0(x)(\varepsilon - E_j(x))\right)\varphi_\varepsilon^j(x) = 0, \tag{18}$$

where the coefficients c_{ij}^1, c_{ij}^2, and c_{ij}^0 are given by

$$c_{ij}^2(x) = \int_0^{L_z} \chi_i(z;x)\frac{\partial^2}{\partial x^2}\chi_j(z;x)\,\mathrm{d}z, \tag{19}$$

$$c_{ij}^1(x) = \int_0^{L_z} \chi_i(z;x)\frac{\partial}{\partial x}\chi_j(z;x)\,\mathrm{d}z, \tag{20}$$

$$c_{ij}^0(x) = \int_0^{L_z} m_x^*(z)\chi_i(z;x)\chi_j(z;x)\,\mathrm{d}z. \tag{21}$$

The source/drain open boundary conditions for φ_ϵ^i are derived from the projection of the 2D quantum transmitting boundary conditions on each mode i. We assume that the c_{ij}^0 coefficients are neglected if $i \neq j$ ($c_{ii}^0 \gg c_{ij}^0$), which is satisfied if the electron gas is assumed mainly located in the silicon at the contacts. Therefore, the obtained boundary conditions are similar to the 1D open boundary conditions described in Appendix B. Namely for a contact located at a point x_p, we have

$$\alpha_p\frac{\partial}{\partial x}\varphi^i(x)_{|_{x_p}} = ik_{x,i}^p(-2a_i^p + \varphi^i(x_p)) \quad \text{if } i \leqslant I_p(\varepsilon), \tag{22}$$

$$\alpha_p\frac{\partial}{\partial x}\varphi^i(x)_{|_{x_p}} = -k_{x,i}^p\varphi^i(x_p) \quad \text{if } i > I_p(\varepsilon), \tag{23}$$

where the contact p (≡s or d) is equivalent to the source if $x_p = x_s$ and $\alpha_p = -1$ or to the drain if $x_p = x_d$ and $\alpha_p = 1$, $k_{x,i}^p$ is the wave vector in the x direction associated to the transverse mode i of the contact p, a_i^p is the amplitude of the incoming wave from the contact p on the mode i, and I_p is the number of propagating modes of the contact p given by

$$I_p = \sup_i\{\varepsilon > E_i(x_p)\}. \tag{24}$$

In the sequel, we denote by $\varphi_{p0,i0,kx}^i$ the scattering state solution of the Schrödinger equation (18) associated to one incoming wave from the contact p_0, the transverse mode i_0, and with the wave vector k_x. The amplitude coefficient is then equal to

$$a_i^p = \delta_{x_p,x_{p_0}}\delta_{i,i_0}, \tag{25}$$

and the reduced energy is now given by

$$\varepsilon(p_0, i_0, k_x) = E_{i_0}(x_{p_0}) + \frac{(\hbar k_{x,i_0}^{p_0})^2}{2\overline{m}_x^*}, \tag{26}$$

where $\overline{m}_x^*$ denotes the average of the effective mass m_x in the z direction

$$\overline{m}_x^*[\chi_{i_0}^{p_0}] = \int_0^{L_z} m_x^*(z)|\chi_{i_0}(x_{p_0}, z)|^2 \, \mathrm{d}z \equiv c_{i0i0}^0(x_{p_0}). \tag{27}$$

Similarly, the expression of the effective mass $\overline{m}_y^*$ in (15) can be chosen in order to select the region where the electron gas is localized in the 2D domain. One possible choice is to replace the m_y^* anisotropic effective mass by its mean value in the x,z directions using as a weight, the wave function ψ_ε itself

$$\overline{m}_y^*[\psi_\varepsilon] = \frac{\int_{x_s}^{x_d} \int_0^{L_z} m_y^*(z)|\psi_\varepsilon(x,z)|^2 \, \mathrm{d}x \, \mathrm{d}z}{\int_{x_s}^{x_d} \int_0^{L_z} |\psi_\varepsilon(x,z)|^2 \, \mathrm{d}x \, \mathrm{d}z}. \tag{28}$$

Another possible choice would be to replace similarly the $1/m_y^*$ coefficients by its mean value.

To summarize, for a given energy E, the associated full wave function can be written as

$$\Psi_E(x,y,z) = \mathrm{e}^{\mathrm{i}k_y y}\psi_{p_0,i_0,k_x}(x,z) \tag{29}$$

with

$$\psi_{p_0,i_0,k_x}(x,z) = \sum_i \varphi_{p_0,i_0,k_x}^i(x)\chi_i(z;x). \tag{30}$$

For a given potential $U(x,z)$ in the $[x_s,x_d]\times[0,L_z]$ region, the electron and current densities in a (m_x^*, m_y^*, m_z^*) effective mass configuration correspond to a statistical mixture of these scattering states

$$n_{x,y,z}(x,z) = 2\sum_{p_0}\sum_{i_0}^{\infty}\int_0^{+\infty} |\psi_{p_0,i_0,k_x}(x,z)|^2 \left(\int_{-\infty}^{+\infty} f_{\mathrm{FD}}\left(E(p_0,i_0,k_x,k_y) - \mu_{p_0}\right)\frac{\mathrm{d}k_y}{2\pi}\right)\frac{\mathrm{d}k_x}{2\pi}, \tag{31}$$

where 2 is for the spin factor, f_{FD} is the Fermi Dirac distribution and $\mu_{p_0} \equiv \mu(x_{p_0})$ is the chemical potential associated to the source if $x_{p_0} = x_s$ or to the drain if $x_{p_0} = x_d$. In the latter expression the integral over the k_y wave vector is related to the plane waves in the y direction while the sum over the contacts, the modes as well as the integral over the wave vector k_x are concerned only with the wave functions ψ_{p_0,i_0,k_x}.

Similarly, the current density is given by

$$\begin{aligned}\vec{J}_{x,y,z}(x,z) = 2\sum_{p_0}\sum_{i_0}^{\infty}\int_0^{+\infty} \frac{q\hbar}{\overline{m}_x^*}\mathrm{Im}\{\overline{\psi}_{p_0,i_0,k_x}(x,z)\nabla\psi_{p_0,i_0,k_x}(x,z)\} \\ \times\left(\int_{-\infty}^{+\infty} f_{\mathrm{FD}}\left(E(p_0,i_0,k_x,k_y) - \mu_{p_0}\right)\frac{\mathrm{d}k_y}{2\pi}\right)\frac{\mathrm{d}k_x}{2\pi}.\end{aligned} \tag{32}$$

At equilibrium (with no applied bias voltage between the source and the drain contacts) the chemical potentials are all equal to a single value μ. Out of equilibrium, denoting v_{p_0} the applied bias at the contact p_0, we obtain

$$\mu_{p_0} = \mu - qv_{p_0}, \tag{33}$$

and the net current is non-vanishing.

Finally, the current intensity I in the transport direction is defined by

$$I = \int_0^{L_z} \vec{J}(x,z)\cdot\vec{x}\,\mathrm{d}z. \tag{34}$$

4.2. Finite element discretization

To summarize the previous sections, in the Schrödinger–Poisson coupled system, we essentially need to solve three kinds of equations: (i) a 1D eigenvalue Schrödinger equation in the z direction (17), (ii) a 1D coupled open Schrödinger equation in the x direction (18) and (iii) a 2D Poisson's equation in the x,z region. A natural choice for the mesh is then a non-uniform grid $[x_s,x_d]\times[0,L_z]$ composed by $N_x\times N_y$ nodes. In this case, the only non-standard finite element discretization is concerned with the equation (ii) which is described in detail in the following.

Let M be the number of modes, then the weak variational formulation of the 1D coupled problem (18) with open boundary conditions (22) and (23) is as follows.

Find $\varphi_\varepsilon \equiv (\varphi_\varepsilon^1 \dots \varphi_\varepsilon^M) \in (H^1([x_s,x_d]))^M$, such as for all arbitrary test functions $\phi_\varepsilon \equiv (\phi_\varepsilon^1 \dots \phi_\varepsilon^M) \in (H^1([x_s,x_d]))^M$,

$$a(\varphi_\varepsilon,\phi_\varepsilon) = L(\phi_\varepsilon), \tag{35}$$

where

$$\begin{aligned} a(\varphi_\varepsilon,\phi_\varepsilon) = {} & \sum_{i=1}^{M}\int_{x_s}^{x_d} \nabla\varphi_\varepsilon^i \nabla\phi_\varepsilon^i \,\mathrm{d}x - 2\sum_{i=1}^{M}\sum_{j=1}^{M}\int_{x_s}^{x_d} c_{ij}^1(x)\frac{\mathrm{d}}{\mathrm{d}x}\varphi_\varepsilon^j\phi_\varepsilon^i \,\mathrm{d}x \\ & - \sum_{i=1}^{M}\sum_{j=1}^{M}\int_{x_s}^{x_d}\left(c_{ij}^2(x)+\frac{2}{\hbar^2}c_{ij}^0(x)\big(\varepsilon - E_j(x)\big)\right)\varphi_\varepsilon^j\phi_\varepsilon^i \,\mathrm{d}x - \mathrm{i}\sum_{i=1}^{I_s(E)} k_i^s(E)\varphi_\varepsilon^i(x_s)\phi_\varepsilon^i(x_s) \\ & - \mathrm{i}\sum_{i=1}^{I_d(E)} k_i^d(E)\varphi_\varepsilon^i(x_d)\phi_\varepsilon^i(x_d) + \sum_{i=I_s(E)+1}^{M} k_i^s(E)\varphi_\varepsilon^i(x_s)\phi_\varepsilon^i(x_s) + \sum_{i=I_d(E)+1}^{M} k_i^d(E)\varphi_\varepsilon^i(x_d)\phi_\varepsilon^i(x_d), \end{aligned} \tag{36}$$

and

$$L(\phi_\varepsilon) = -2\mathrm{i}\sum_{i=1}^{I_s(E)} a_i^s k_i^s(E)\phi_\varepsilon^i(x_s) - 2\mathrm{i}\sum_{i=1}^{I_d(E)} a_i^d k_i^d(E)\phi_\varepsilon^i(x_d), \tag{37}$$

with $k_i^p \equiv k_{x,i}^p$.

We use the P^1 finite element method to solve problem. We note $u^i \equiv (u_1^i,\dots,u_{Nx}^i)$ the vector of the N_x (unknown) nodal values of φ_ε^i corresponding to a given mesh by

$$\varphi_\varepsilon^i(x) = \sum_{n=1}^{N_x} u_n^i \omega_n(x), \tag{38}$$

where ω_n is the P^1 basis function of the n node with $\omega_n(x_{n'}) = \delta_{n,n'}$ ($x_{n'}$ is the position of the n' node). We can make similar expansions for the test function ϕ^i. The system to be solved is

$$\begin{pmatrix} A^{11} & A^{12} & \dots & A^{1M} \\ A^{21} & A^{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ A^{M1} & \dots & \dots & A^{MM} \end{pmatrix}\begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^M \end{pmatrix} = \begin{pmatrix} L^1 \\ L^2 \\ \vdots \\ L^M \end{pmatrix}, \tag{39}$$

where A^{ii} are symmetric, complex valued and tridiagonal $(N_x\times N_x)$ matrices, A^{ij} $(i\neq j)$ are symmetric, real and tridiagonal $(N_x\times N_x)$ matrices, and L^i are complex $(N_x\times 1)$ vectors. The elements of these matrix are defined as follows:

$$(A^{ii}_{n,n'})_{1\leqslant n,n'\leqslant N_x} = \int_{x_s}^{x_d} \nabla\omega_n \nabla\omega_{n'}\,\mathrm{d}x - \int_{x_s}^{x_d} \left(c^2_{ii}(x) + \frac{2}{\hbar^2}c^0_{ii}(x)(\varepsilon - E_i(x))\right)\omega_n\omega_{n'}\mathrm{d}x - \mathrm{i}k^s_i(E)\delta_{n,1}\delta_{n',1}\delta_{i\leqslant I_s(E)} - \mathrm{i}k^d_i(E)\delta_{n,N_x}\delta_{n',N_x}\delta_{i\leqslant I_d(E)} + k^s_i(E)\delta_{n,1}\delta_{n',1}\delta_{i>I_s(E)} + k^d_i(E)\varphi^i(x_d)\delta_{n,N_x}\delta_{n\prime,N_x}\delta_{i>I_d(E)}, \tag{40}$$

$$(A^{ij}_{n,n'})^{i\neq j}_{1\leqslant n,n'\leqslant N_x} = -2\int_{x_s}^{x_d} c^1_{ij}(x)\frac{\mathrm{d}}{\mathrm{d}x}\omega_n\omega_{n'}\,\mathrm{d}x - \int_{x_s}^{x_d}\left(c^2_{ij}(x) + \frac{2}{\hbar^2}c^0_{ij}(x)\big(\varepsilon - E_j(x)\big)\right)\omega_n\omega_{n'}\,\mathrm{d}x, \tag{41}$$

$$(L^i_n)_{1\leqslant n\leqslant N_x} = -2\mathrm{i}a^s_i k^s_i(E)\delta_{n,1}\delta_{i\leqslant I_s(E)} - 2\mathrm{i}a^d_i k^d_i(E)\delta_{n,N_x}\delta_{i\leqslant I_d(E)}, \tag{42}$$

where $\delta_{i\leqslant I_{p(E)}}$ and $\delta_{i>I_{p(E)}}$ denote the Heavyside function.

Finally, we note that the construction of the system (39) has been done when varying in (40)–(42), the index i over the mode after the index n over the nodes. The matrix can also be set up by the variation of the index i before the variation on the index n. In this approach, the obtained matrix is tridiagonal by blocks where each $M\times M$ block is dense.

4.3. Numerical implementation of the Schrödinger–Poisson system

The numerical Schrödinger–Poisson algorithm is summarized by the five following steps:

(1) For a given potential energy $U(x,z)$ in the $[x_s,x_d]\times[0,L_z]$ region, we solve the generalized eigenvalues problem (obtained after the finite element discretization of (17)) in the confined direction z for all the nodes along the transport direction with coordinate x_n, $n = 1\ldots N_x$, using the direct method from LAPACK [33]. Therefore, we obtain N_x sets of eigenfunctions $\{\chi_i(z)\}$ and eigenvalues $\{E_i\}$ (see Fig. 5). These functions have to be normalized in the $L^2[0,L_z]$ space. The coefficients c^2_{ij}, c^1_{ij} and c^0_{ij} are also calculated using (19)–(21).

(2) For a given energy ε, the wave vectors k^s_i and k^d_i associated to the incoming scattering states in the domain in the x transport direction, are calculated using (26). We construct the $(M\times M)$ block matrix (39) composed by the $N\times N$ tridiagonal matrices $A^{i,j}$ defined in (40) and (41). For a given incoming wave in the contact p_0 ($p_0\equiv$ s,d), and in its transverse mode i_0, the coefficients a^s_i and a^d_i are given by (25). The complex symmetric linear system (39) is solved by direct methods or by iterative ones as the quasi-minimal residual (QMR) procedure with SSOR or ILUT preconditioner [34].

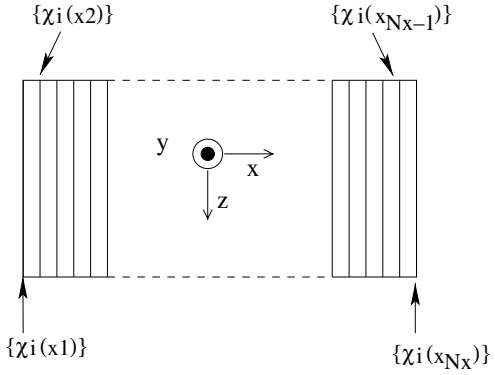

Fig. 5. A eigenvalue problem is solved on each slices of the device.

(3) The electron and the current densities are calculated, respectively, using (11), (31), and (12), (32), by iterations of the previous step on the incoming wave features (contact, mode and wave vector). Otherwise, the calculations of the electron and current densities require to solve a large number of independent complex symmetric linear sparse systems. The upper limit onthe integration over the energy is fixed at $\mu_p + 4k_B T$ for the contact p.
(4) The Poisson's equation is solved in the x,z domain using Dirichlet boundary conditions for the interfaces with the gates, and Neumann boundary conditions on the source and drain interfaces. The obtained real symmetric positive definite sparse system is solved using the preconditioned conjugate gradient method (with incomplete Cholesky factorization as preconditioner).
(5) Repeat three times all the four previous steps, respectively, for the different effective mass configurations summarized in (11) and (12). Note that the eigenvalues problems in the first step are to be solved only twice, respectively, for the m_l^* and the m_t^* configurations in the z variable.

We propose to point out some remarks related to the numerical implementation of this algorithm:

- The most time consuming part of the previous algorithm consists in the calculation of the wave functions. Since both the generalized eigenvalues problems involved in the step (1) and the linear systems involved in the step (2) can be solved in an independent way, a parallel version of the code was developed using MPI directives. This allows to approximately divide the calculation time by the number of processors.
- Implicit numerical schemes have to be used to solve the highly non-linear coupled Schrödinger–Poisson system. Moreover, the Newton–Raphson method is not appropriate since the density does not depend locally on the potential. Therefore, for a given potential V^n at the step n, we propose to implicit the scheme as follows:

$$-\nabla\left(\epsilon_r(z)\nabla V^{n+1}\right) = \frac{q}{\epsilon_0}\left(n_D(z) - n(x,z)\frac{T[V^{n+1}]}{T[V^n]}\right), \tag{43}$$

where T is a functional of V, ϵ_0 is the vacuum permittivity, ϵ_r is the relative dielectric constant, q is the free electron charge and n_D is the positive doping profile. Because of the exponential behavior of the electron density in function of the potential V (see below the Thomas–Fermi approximation), a suitable choice for T is then given by $T[V] = \exp(q\beta V)$ (with $\beta = 1/k_B T$). The linearization of this coupled system leads to the Gummel iterative scheme [35] where for a given potential V^n at the step n, the new potential V^{n+1} is now given by

$$-\nabla\left(\epsilon_r(z)\nabla V^{n+1}\right) + \frac{q}{\epsilon_0}n(x,z)\frac{V^{n+1}}{V_{ref}} = \frac{q}{\epsilon_0}\left(n_D(z) - n(x,z)\left(1 - \frac{V^n}{V_{ref}}\right)\right) \tag{44}$$

with $V_{ref} = k_B T/q$.
- At equilibrium and in order to obtain a suitable initial guess to begin the simulations, we use the Thomas–Fermi/Poisson semi-classical approximation:

$$n_{x,y,z}(x,z) = 2\int\int\int_{-\infty}^{+\infty} f_{FD}(E-\mu)\,\frac{dk_x\,dk_y\,dk_z}{8\pi^3}, \tag{45}$$

where the wave functions are plane waves in the whole domain and the energy depends on the local potential

$$E = U(x,z) + \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_x^*(z)} + \frac{k_y^2}{m_y^*(z)} + \frac{k_z^2}{m_z^*(z)}\right). \tag{46}$$

In this case the coupled system can be solved using a Newton–Raphson method. However, a better choice for the initial guess consists to account for the confinement of the electron gas in the z direction using the $\chi_i(z;x)$ functions [36]. The electron density of this hybrid model is written as

$$n_{x,y,z}(x,z) = 2\sum_i^{\infty} \int_{-\infty}^{+\infty} |\chi_i(z;x)|^2 \left(\int_{-\infty}^{+\infty} f_{\mathrm{FD}}(E-\mu)\frac{\mathrm{d}k_y}{2\pi} \right) \frac{\mathrm{d}k_x}{2\pi} \tag{47}$$

with

$$E = E_i(x,z) + \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_x^*(z)} + \frac{k_z^2}{m_z^*(z)} \right). \tag{48}$$

Another alternative approach to this latter is to solve 1D self-consistent Schrödinger–Poisson problems for each slice (see for example [37,38]).

These two hybrid approaches (quantum in the confined direction and semi-classical in the transport one) represent a good approximation of the quantum model at equilibrium only with the assumption of a slow variation for the potential in the x transport direction.

- We propose now to point out some problems that one may meet with the calculation of the functions $\chi_i(z;x)$ in the simulation of the double gates MOSFET (see Fig. 3). Indeed, if the thickness of the silicon layer is large enough ($\sim$10 nm or more) then the first two modes $\chi_1(z;x)$ and $\chi_2(z;x)$ have very close energies.

In this case, the numerical procedure used to solve independently the eigenvalue problems in the z direction (see Fig. 5) does not insure the regularity of the functions χ_1 and χ_2 on x. Indeed, in this case the eigenvalues can be considered as double, so that only the vector space spanned by χ_1 and χ_2 is smooth in x. Since we diagonalize each x slice independently of the neighbouring one, the resulting wave functions may not be smooth in x (see Fig. 6). In order to have the correct values for χ_1 and χ_2, we proceed as follows.

Let $\chi_1(x_j) \equiv \chi_1(z;x_j)$ and $\chi_2(x_j) \equiv \chi_2(z;x_j)$ at the x_j coordinate, and let $\widehat{\chi_1}(x_{j+1})$ and $\widehat{\chi_2}(x_{j+1})$ be the result of the diagonalization procedure at x_{j+1}. The eigenfunctions χ_1 and χ_2 at $x = x_{j+1}$ are then obtained by "rotating" appropriately $\widehat{\chi_1}(x_{j+1})$ and $\widehat{\chi_2}(x_{j+1})$. Namely,

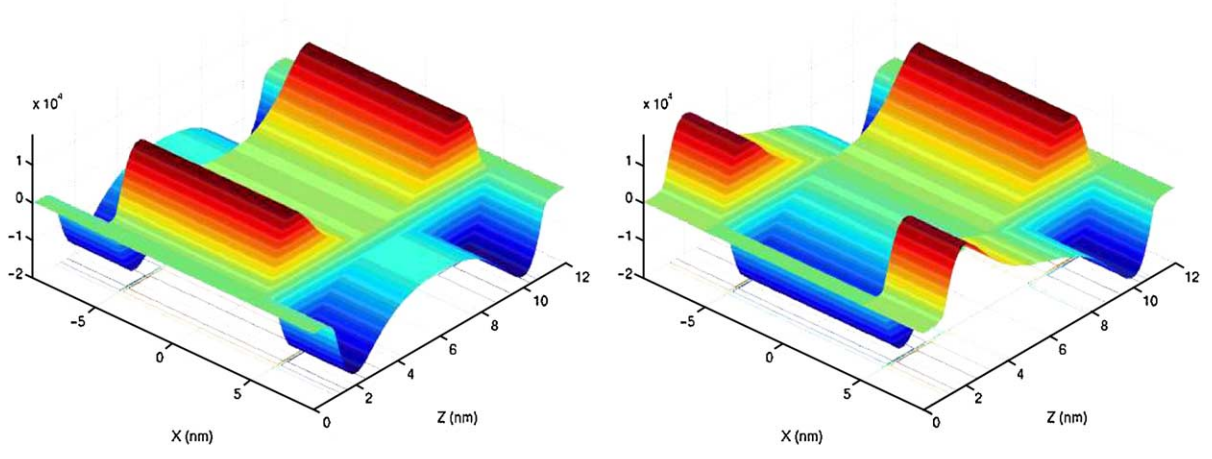

Fig. 6. Variations of the first two eigenfunctions $\chi_1(z;x)$ (on the left) and $\chi_2(z;x)$ (on the right) along the x transport direction. In this case (we take 10 nm both for the thickness and the length of the channel), we note that the eigenfunctions show a non-regular behavior in the middle of the device.

$$\begin{pmatrix} \chi_1(x_{j+1}) \\ \chi_2(x_{j+1}) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \widehat{\chi}_1(x_{j+1}) \\ \widehat{\chi}_2(x_{j+1}) \end{pmatrix}, \tag{49}$$

where θ is the unknown rotation angle. Since $\chi_1(x_{j+1})$ and $\chi_2(x_{j+1})$ must be very close to $\chi_1(x_j)$ and $\chi_2(x_j)$, we naturally impose the condition

$$\int_0^{L_z} \chi_1(x_{j+1})\chi_2(x_j)\,\mathrm{d}z = 0, \tag{50}$$

which defines θ by

$$\tan(\theta) = -\frac{\int_0^{L_z} \widehat{\chi}_1(x_j)\chi_2(x_j)\,\mathrm{d}z}{\int_0^{L_z} \widehat{\chi}_2(x_j)\chi_2(x_j)\,\mathrm{d}z}. \tag{51}$$

Finally, a last point concerns the sign of both the scalar product of $\chi_1(x_{j+1})$ with $\chi_1(x_j)$ and the scalar product of $\chi_2(x_{j+1})$ with $\chi_2(x_j)$ on $L^2[0,L_z]$, which must be positive. Fig. 7 illustrates now the variations of the first two eigenfunctions presented in Fig. 6 after the above numerical treatment.

4.4. Numerical simulations of the double-gate transistor

In this section, the subband decomposition approach is applied to the double-gate transistor. The obtained numerical results are used to illustrate the numerical efficiency of this representation whose the simulation times are compared to those obtained with the full 2D approach.

We denote by L and t, respectively, the length size and the thickness of the channel, the thickness of the oxide layer under the two gates is defined by t_{ox}. We recall that n_D is the positive doping profile and that the whole calculation domain is restricted along the transport direction by x_s and x_d (see Fig. 3). In our simulations, we take $x_d - x_s = 18$ nm, $L = 10$ nm, $t = 5$ nm, $t_x = 1$ nm, $n_D = 10^{26}$ m^{-3}, $\mu = 0$ and the temperature $T = 300$ K. Denoting m_e the electron mass, then the transverse and longitudinal effective mass in silicon are, respectively, equal to $m_t^* = 0.19m_e$ $m_l^* = 0.98m_e$. The effective mass in oxide is chosen as $0.5m_e$ and the relative dielectric constant ϵ_r of the silicon and the oxide are, respectively, given by 11.9 and 3.8. The

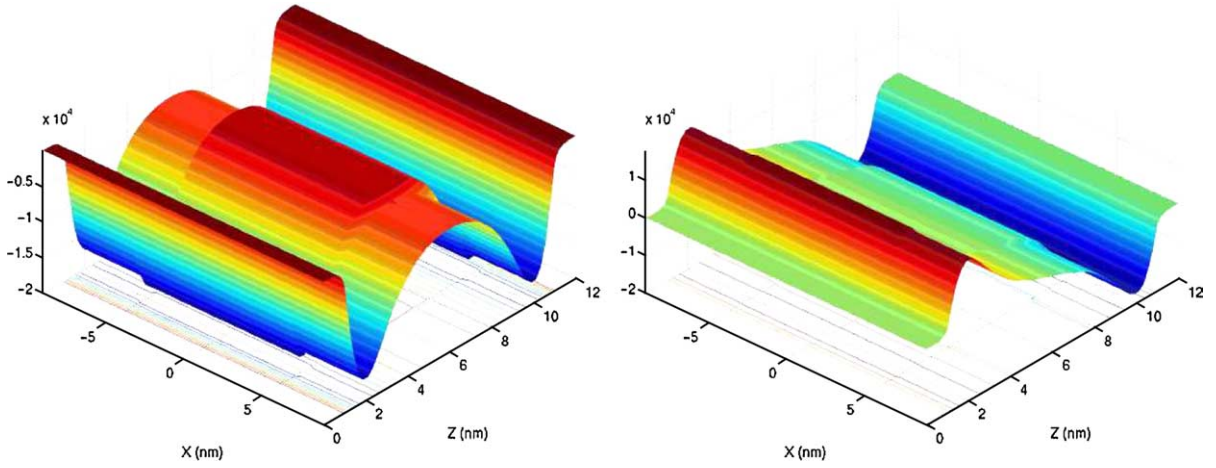

Fig. 7. Variations of the first two eigenfunctions $\chi_1(z;x)$ (on the left) and $\chi_2(z;x)$ (on the right) along the x transport direction numerically obtained using the assumption of degenerate states. We note now a regular behavior for the eigenfunctions in the device.

features of the grid mesh are $N_x = 72$ and $N_y = 75$, whereas the maximum number of modes is equal to $M = 8$.

Fig. 8 shows the obtained electron density in the device at equilibrium. In this case, the thickness of the device is large enough to yield two electron channels in the active region.

Table 2 compares the computer resources required to obtain the same numerical results at equilibrium by the full 2D approach and by the subband representation. These calculations show that the subband decomposition approach is approximately 20 times faster than the full 2D one, allowing to reduce a 5 h simulation time to a 15 min on a bi-compaq DS20, to get 10 points on the I–V curves. These times are proportional to the number of processors due to the perfect parallelism of both models. The speed improvement factor depends on: (i) the number of processors and the machine architecture, (ii) the size of the transport problem when the size of the transverse one is fixed (that would increase the size of the linear systems but not of the eigenvalue problems), (iii) the number of points in the confined direction when the number of points in the transport one is fixed (that would increase the size the linear system for the full approach while in the subband approach the size of the linear system keeps unchanged while the size of the eigenvalue problems increases). All these quantitative considerations for the simulations are not exposed in this article. Some others simulations time results obtained on a Linux cluster are given for the 3D silicon nanowire devices in [43,44].

Some results on the I–V characteristics of the DG-MOSFET are shown in Fig. 9 as a function of the gate potential V_{GS}. In the sequel, results are shown for $V_{GS} = 0.5$ V.

The variations of the first energy levels $E_i(x)$ along the x transport direction for the two configurations of the effective mass in the z confined direction are given in Fig. 10 at equilibrium for $V_{DS} = 0$ V and out of equilibrium for $V_{DS} = 0.5$ V. In the latter case, six modes are located at the drain contact below the upper limit of the energy $\mu_s + 4k_BT$ if $m_z^* = m_l^*$ (the thermal energy $4k_BT$ is next to 0.1 eV and the chemical potential is equal to zero at equilibrium $\mu_s \equiv \mu = 0$), while only three modes appear if $m_z^* = m_t^*$. These modes are the propagating ones where the electron transport takes place from the source to the drain contacts, and then the higher modes are defined as the evanescent ones.

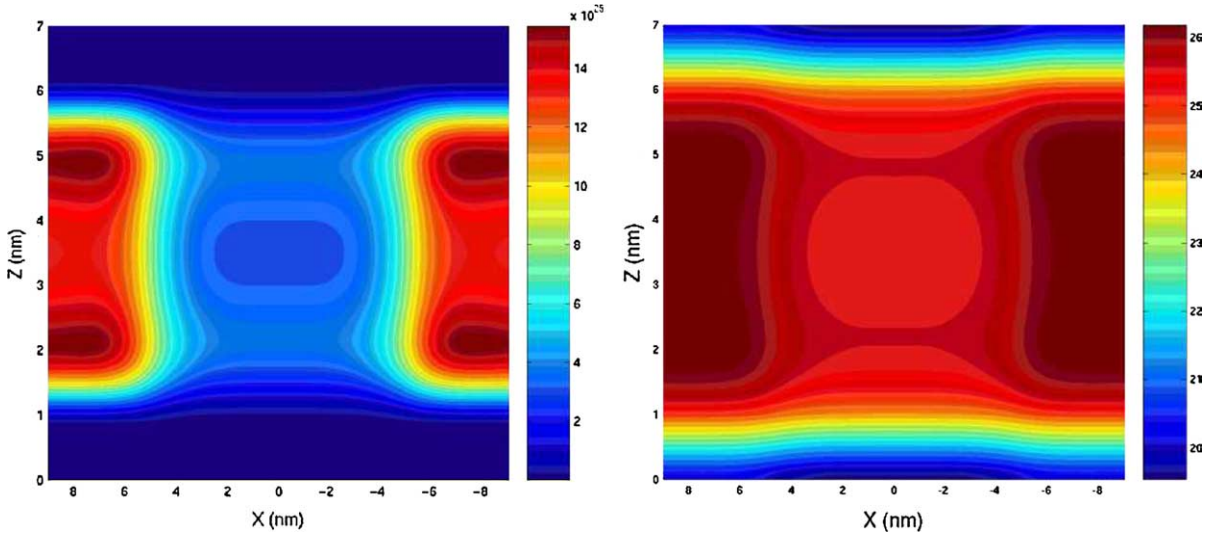

Fig. 8. Electron density at the equilibrium state. The left figure is plotted using a linear scale while we use a logarithmic one on the right figure to point out the presence of the two channels.

Table 2
List of the number and the nature of the operations require to solve the problem using the full 2D approach and the subband representation one (for only one point in the I–V curve)

	Full 2D approach	Subband decomposition with $M = 8$
Size of the "transport Schrödinger matrix"	5400	576
Number of linear "transport Schrödinger matrix"	~1500	~1500
Size of the "confined Schrödinger matrix"	None	75
Number of linear "confined Schrödinger matrix"	None	72
Number of iterations on the potential	~5–10	~5–10
Simulation time factor (on a bi-compaq DS20)	1	0.05

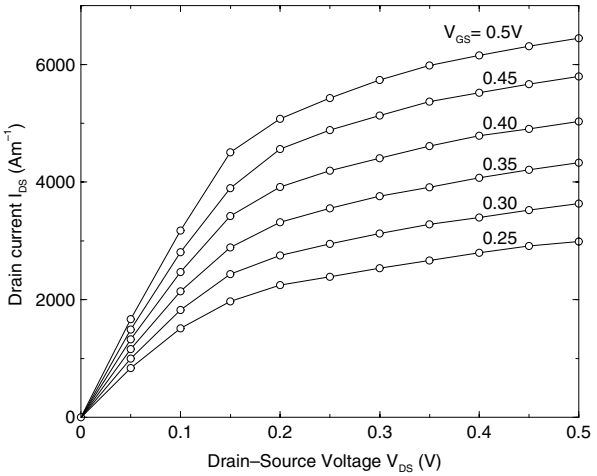

Fig. 9. I–V characteristics for the device obtained using the subband representation. The gate source voltage V_{GS} is changed with a step of 0.05 V.

We note that the drain potential effect involves a shift in the energy levels at the drain contact as shown in Fig. 11. This shift is expected to be exactly equal to the amplitude of the associated drain bias voltage. However, due to the Neumann boundary conditions on the potential applied at the drain and the source contact, we note a drop on the energy level in the source contact while no bias voltage is applied on it. This problem is related to the ballistic transport assumption and we refer [39] for the physical interpretation of this phenomena.

The choice of $M = 8$ modes in our simulations means that we take into account about two evanescent modes when $m_z^* = m_l^*$ and $V_{DS} = 0.5$ V. However, the $M = 8$ value is an overestimated choice when it implies more than two evanescent modes as for a smaller drain potential (or at equilibrium) or for $m_z^* = m_t^*$. Therefore, the number of transverse mode M considered in the simulations can be adapted to speed up the numerical convergence of the subband approach for each effective mass and potential configurations.

Fig. 12 shows a close agreement between numerical results on the I–V characteristics when less coupling operates between the modes. Indeed, the numerical results are almost not changed if only two modes are coupled, and they lead to very small differences with the zero and the one coupling mode models. Therefore, it is not necessary to calculate all the extra-diagonal matrix elements A^{ij} of the matrix (39) to get satisfactory results, and then the reduction of the number of non-zero elements can be significant. Let us call the m-coupling modes model the one for which A^{ij} is set to zero for $|i - j| > m$. Also, in the 0-coupling mode

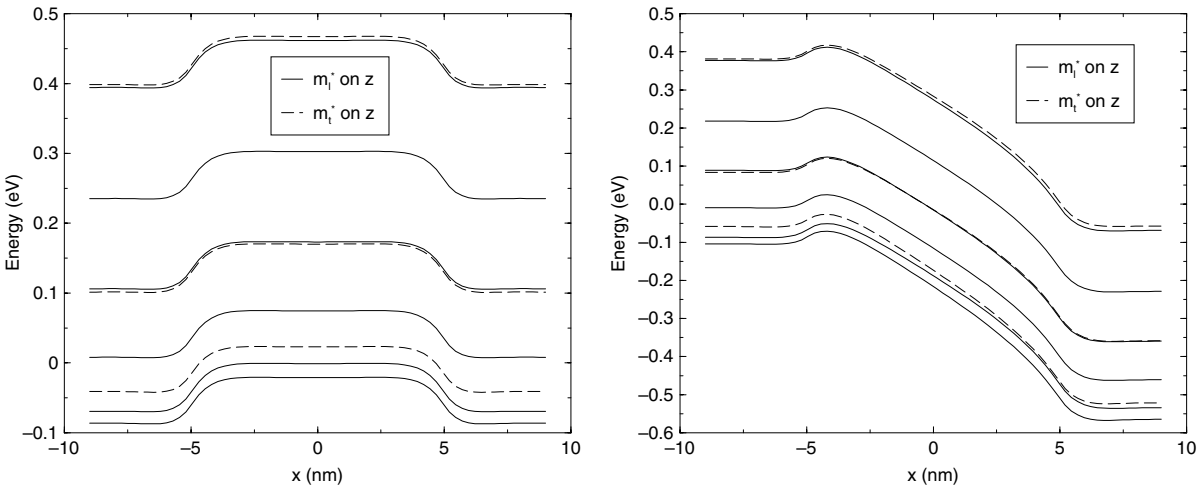

Fig. 10. Energy levels $E_i(x)$ along the x transport direction for the two effective mass configuration in the z direction, at the equilibrium state (on the left) and out of equilibrium with $V_{DS} = 0.5$ V (on the right).

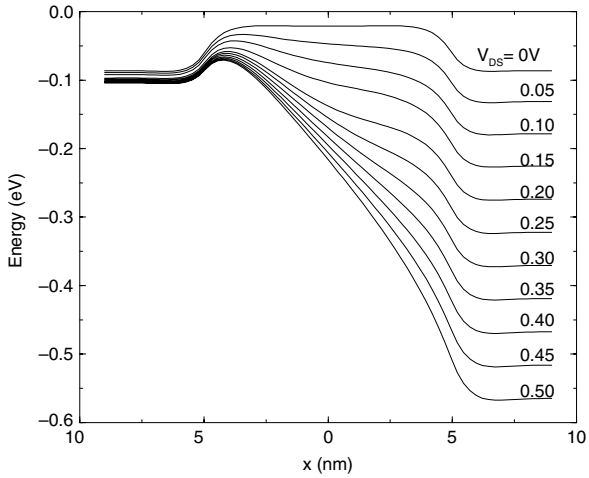

Fig. 11. Example of the behavior of the first mode $E_1(x)$ for the m_l^* configuration in the z direction, as function of the drain current voltage V_{DS}.

model, the modes are not coupled and then the non-diagonal terms in Eq. (18) are neglected. The matrix (39) is thus a tridiagonal one and its computation requires standard numerical procedure. We could also go beyond this simplified model and put $\forall i$, $\partial_x \chi_i(z;x) = 0$ in Eq. (18) as suggested in [17,29], then the diagonal terms c_{ii}^2 now disappears (we recall that $c_{ii}^1 = 0$). In our simulations, we found exactly the same results using this ''no-coupling'' model than using the 0-coupling mode model where we have shown that the obtained results are in a good agreement with the solution. This means that the potential U is expected to keep almost the same shape in the z confined direction along the x transport direction (this is true in our devices since the oxide, the silicon and the gates layers are uniform along the transport direction). Therefore, the eigenfunctions $\{\chi_i\}$ should be almost the same along the x direction (the eigenvalues can be different), and then the no-coupling model appears as a very good approximation.

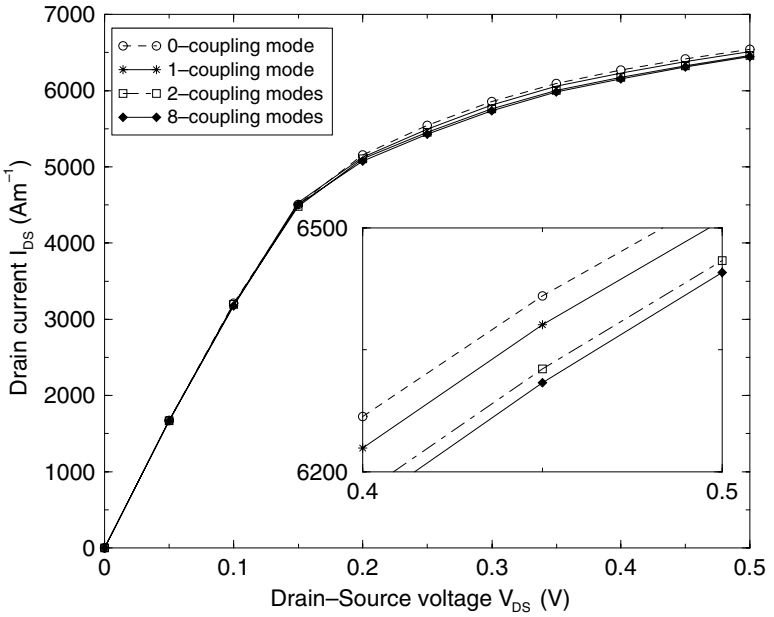

Fig. 12. *I–V* characteristics obtained using different modes coupling configurations are presented. All the results are close each others. However, the caption figure shows very small differences if the model does not involve at least two coupling modes.

The "no-coupling" model has been called by others authors [27–29] a global adiabatic approximation, while the "0-coupling" model is called a diagonal adiabatic approximation. In their simulations to calculate the conductance associated to some class of devices with smooth constriction, it appears that the "no-coupling" model is a better approximation than the "0-coupling" one. In [29], they point out that the "0-coupling" model may be inconsistent since it keeps diagonal terms while neglecting equally important non-diagonal ones. However as we will show further, the non-adiabatic mode mixing becomes important in such non-uniform devices, and off-diagonal terms have to be definitely included in the simulations to get correct results.

In order to estimate the coupling effects between the modes, we analyze the reflection–transmission coefficients. If the coupling effects are small, an electron injected from the source on the mode i will mainly stay on this mode. When the coupling is strong enough, the electron can be transmitted or reflected in the others modes as shown for example in Fig. 13. We define, respectively, R_i and T_i the reflection and the transmission coefficients in the mode i [40,41] by

$$\begin{aligned} R_i &= \frac{k^{\mathrm{s}}_{x,i}}{k^{\mathrm{s}}_{x,i_0}} |\varphi^i(x_{\mathrm{s}})|^2 \quad \text{if } i \neq i_0, \\ R_{i_0} &= |-1+\varphi^i(x_{\mathrm{s}})|^2, \\ T_i &= \frac{k^{\mathrm{d}}_{x,i}}{k^{\mathrm{d}}_{x,i_0}} |\varphi^i(x_{\mathrm{d}})|^2 \quad \forall i. \end{aligned} \tag{52}$$

We note that $R_{i_0} + \sum_i (R_i + T_i) = 1$, and for the devices where no coupling effects occur, we should have $R_{i_0} + T_{i_0} = 1$.

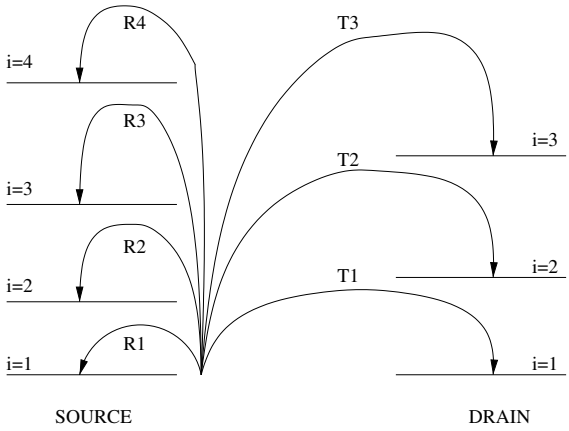

Fig. 13. Schematic representation of the reflection and the transmission waves in the source and drain related to only one incoming wave from source in the mode 1. We denote the modes i the energy levels, respectively, equal to $E_i(x_s)$ in the source and $E_i(x_d)$ in the drain.

Fig. 10 shows that the energy levels in the devices are much close to each other in the effective mass configuration with $m_z^* = m_l^*$ than with $m_z^* = m_t^*$. Therefore, the coupling effects between the modes are expected to be more important in the $m_z^* = m_l^*$ configuration which have been chosen to illustrate the following numerical results on reflection–transmission coefficients.

For one incoming wave from the source in the mode 1, the obtained reflection and transmission coefficients on this first mode as function of the energy spectra are given in Fig. 14. At equilibrium the wave keeps mainly located in the mode 1 since $R_1 + T_1 \simeq 1$ for all the energy values. While with a strong bias voltage on the drain $V_{DS} = 0.5$ V, the coupling effects are a little bit more important, since $R_1 + T_1$ is just lower than one. These results explain that a very small portion of the wave is reflect or transmit in the other modes and then it appears small differences between the coupling modes models in the I–V curves in the saturation regime (see Fig. 12).

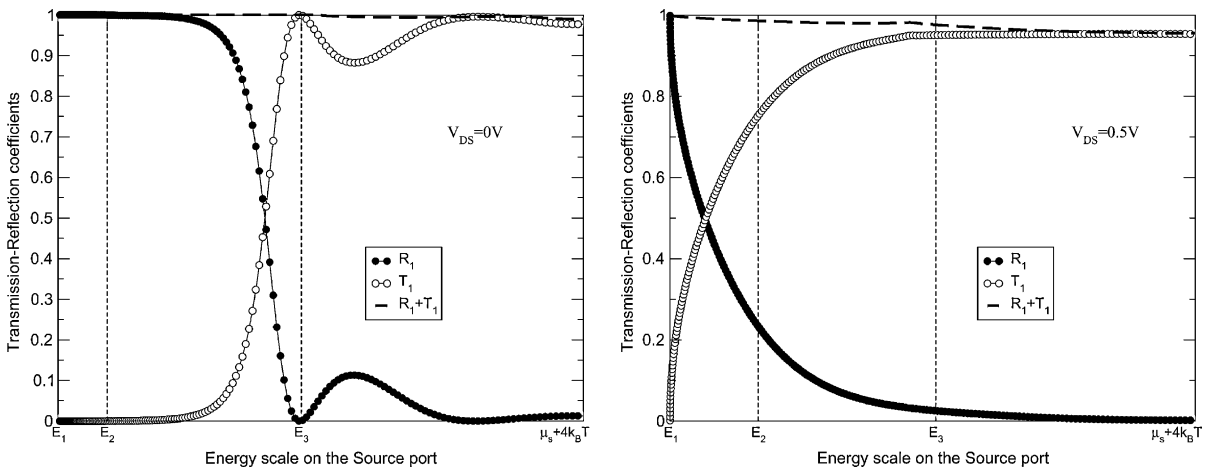

Fig. 14. Reflection and transmission coefficients in the mode 1 as function of the injection energy. These results are obtained for only one incoming wave from the source in the mode 1 and for the $m_z^* = m_l^*$ configuration. The summation of the reflection and the transmission coefficients is considered equal to one for all the energy values at the equilibrium state (on the left) and next equal to one out of equilibrium for $V_{DS} = 0.5$ V (on the right).

In order to investigate the coupling effects and the performance of the subband approach, we propose now to prospect the characteristics of a new device whose the potential variations in the x transport direction present a non-uniform behavior. In this device displayed in Fig. 15, the additional oxide layer leads to squeeze the channel region and then the electron density is now located in only one channel. We thus expect to enhance the importance of the coupling effects between the modes. In the following results, we will consider $l_s = 4$ nm, $L_s = 2$ nm and $t_s = 1.5$ nm.

Fig. 16 compares the I–V characteristics obtained within different coupling modes models. Note that we obtain a very good agreement between results only if two coupling modes are considered. While the obtained currents with both the no-coupling model, the 0-coupling mode model and the 1-coupling mode model show some net differences which increase with the drain potential in the saturation regime. These models overestimate the results since the quantum resistance due to the overall coupling modes is not taken into account.

As previously introduced, a detailed study of the coupling modes effects can be given using the results on the reflection and transmission coefficients in Fig. 17 for equilibrium and in Fig. 18 for out of equilibrium situations. These results show that $R_1 + T_1$ is not equal to one while we consider only one incoming mode from the source in the mode 1. In this case important reflections and transmissions on the others modes appear even if the device is at equilibrium. For high source–drain voltages, the 3-coupling modes model is required since the reflection and the transmission of the wave in the mode 4 is possible. However, we note that the values of the reflection and the transmission coefficients in the mode 4 are important enough only next to the upper energy limit.

To summarize, we showed that at least a 2-coupling modes model is preferred for the study of the double-gate MOSFET, and it is required for the study of "exotic" structures showing non-uniform behaviors. The number of needed modes also depends on the geometry of the structure and the uniformity of the potential in the transport direction. This number can be estimated within the study of the transmission and the reflection coefficients.

However, we would like to point out that the subband decomposition approach presented in this article cannot be applied to particular systems with a squeezed channel structure where $l_s = 0$ nm in Fig. 15.

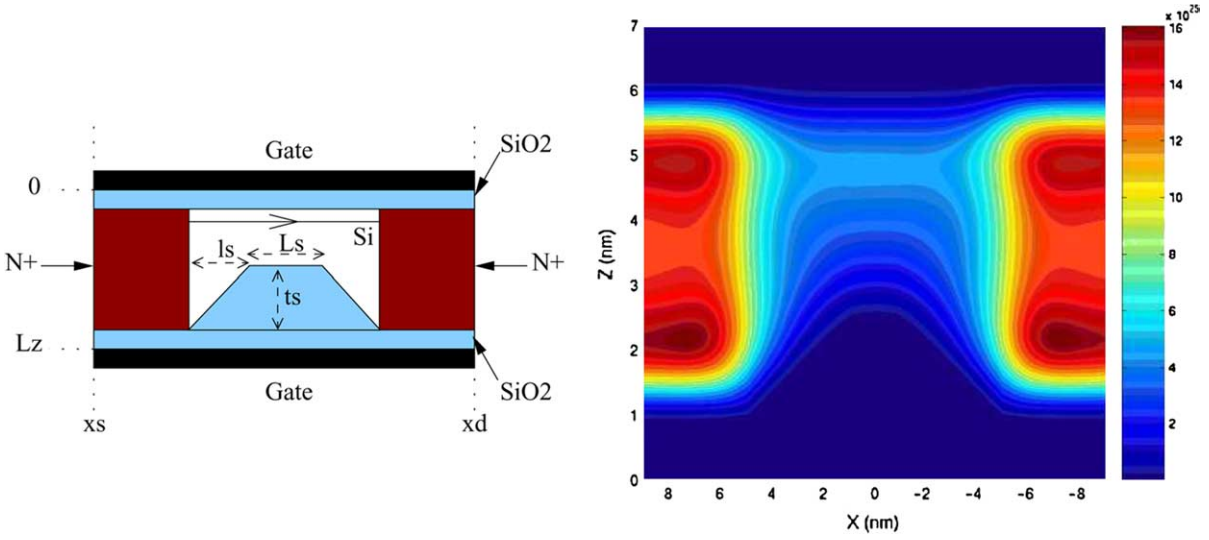

Fig. 15. Representation of a squeezed channel structure (on the left) and its electron density at equilibrium state (on the right). The electrons are led in only one channel in the structure which it is expected to show the importance of the mode coupling effects in the subband decomposition approach.

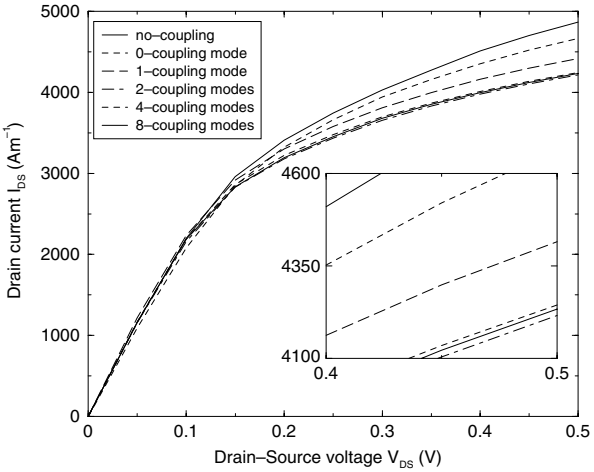

Fig. 16. *I–V* characteristics for the device with squeezed channel where different modes coupling configurations are presented. We note that the numerical results are not suitable if we do not consider at lest two modes in the subband approach. The obtained results with the 2-coupling modes, the 4-coupling modes and the 8-couplings modes model are approximately the same.

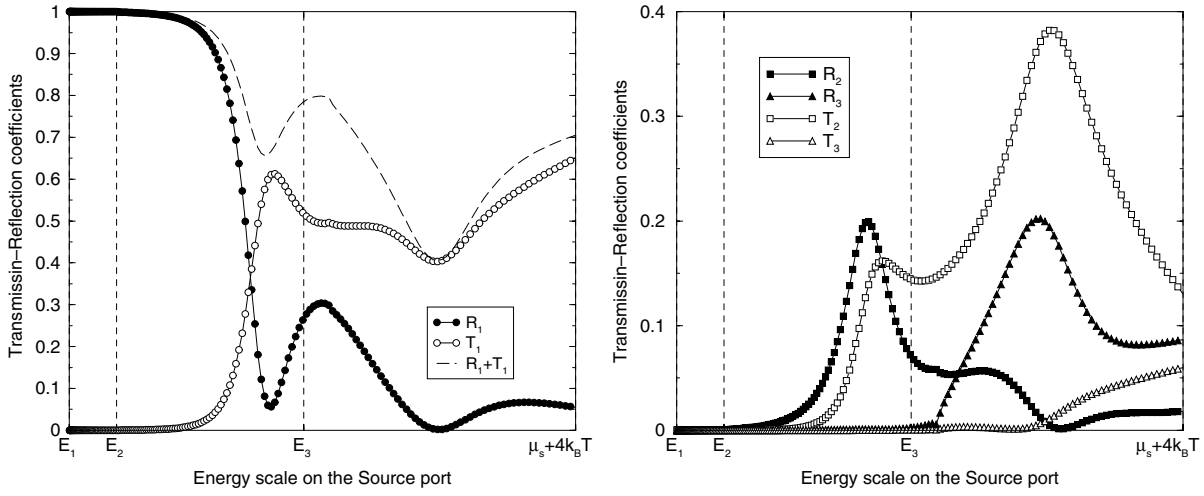

Fig. 17. Reflection and transmission coefficients in the mode 1 (on the left) and in others modes (on the right) as function of the energy spectra in the drain contact. These results are obtained for only one incoming mode from the source in the mode 1 and for the $m_z^* = m_\parallel^*$ configuration at equilibrium ($V_{DS} = 0$ V). The summation of the reflection and transmission coefficients which is considered in the left figure are very different to one for the energy values. This means that it appears in the right figure a transmission and a reflection in the others modes than 1.

Indeed, in this case the device presents a discontinuity in the oxide layer (hard constrictions), and it is not possible to define in this region the derivative of the eigenfunctions $\chi_i(z;x)$ in the x direction, to obtain the required coefficient c_{ij}^0, c_{ij}^1 and c_{ij}^2 for the subband decomposition. Another approach would be to consider smooth constrictions with a very small width l_s ($l_s \ll 1$), that is usually the case in real devices. The numerical treatment of these regions has to be done using a very high numbers of mesh points (large number of slices) to capture the variation of the constriction.

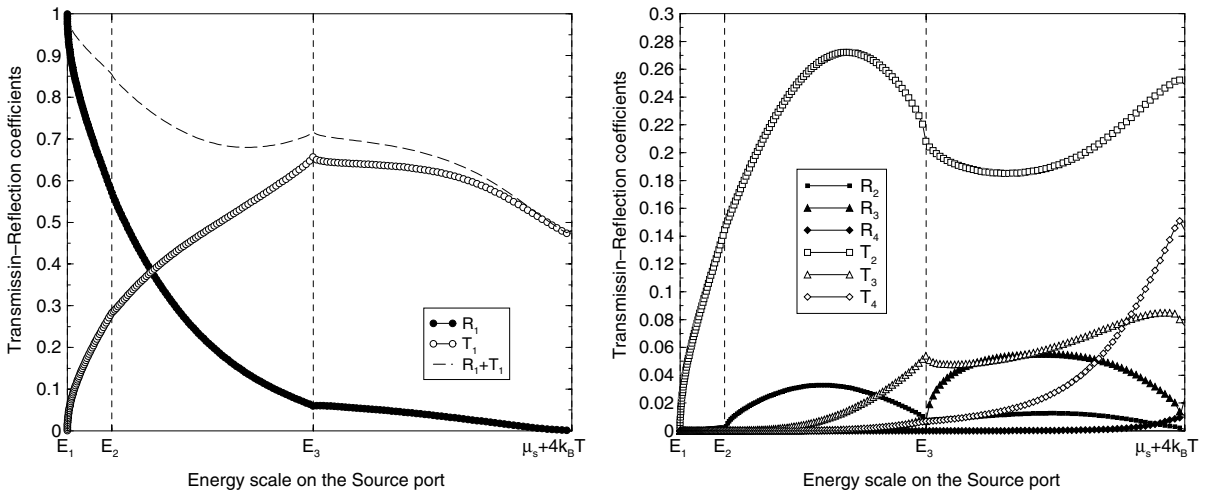

Fig. 18. The same results on the reflection and transmission coefficients than Fig. 17 but now obtained out of equilibrium for $V_{DS} = 0.5$ V.

5. An asymptotic approach of the subband decomposition

The main goal of this section is to present a quasi-full dimensional model which consists in an asymptotic approach of the subband decomposition approach presented before. In this section the potential energy $U(e_1,e_2)$ is given in the region $\Omega_1 \times \Omega_2$ where we recall that the direction e_1 denotes the confined direction(s) of the electron gas, and the direction e_2 denotes the transport direction(s) where the motion of the electrons takes place. The asymptotic approach will be used for a particular structure of potential $U(e_1,e_2)$ which implies strong confinement for the electron gas in the e_1 direction. In the following, the potential energy $U(e_1,e_2)$ will be separated arbitrarily into a potential U_1 depending on the e_1 confined direction, and a potential U_2 depending on e_1,e_2, such that

$$U(e_1,e_2) = U_1(e_1) + U_2(e_1,e_2), \tag{53}$$

where U_2 is assumed slowly varying in e_1.

In the case of the potential U_2 is equal to zero, denoting $\widehat{\chi}_i(e_1)$ and $\widehat{E}_i$ the eigenvectors and the eigenvalues solutions of Eq. (4) in the confined direction which does not depend on the e_2 transport direction

$$\begin{cases} -\frac{\hbar^2}{2m^*}\Delta_1\widehat{\chi}_i(e_1) + U_1(e_1)\widehat{\chi}_i(e_1) = \widehat{E}_i\widehat{\chi}_i(e_1), \\ \int_{\Omega_1}\widehat{\chi}_i(e_1)\overline{\widehat{\chi}}_j(e_1)\,\mathrm{d}e_1 = \delta_{i,j}, \\ \widehat{\chi}_i(e_1) = 0 \quad \text{if } e_1 \in \partial\Omega_1. \end{cases} \tag{54}$$

Therefore, the coefficients $\vec{c}^{\,1}_{ij}$ and c^2_{ij} are equal to zero and the obtained problem in the e_2 transport direction is the decoupled Schrödinger equation (7). The energy $E_i(e_2)$ is also equal to $\widehat{E}_i$ and then the wave functions $\varphi_i(e_2)$ are plane waves in the transport direction.

In the case of the potential U_2 does not depend on the e_1 confined direction (denoting $\widehat{U}_2(e_2)$ this potential), the previous eigenvectors and eigenvalues are denoted by $\widehat{\chi}_i(e_1)$ and $\widehat{E}_i$. The wave functions $\varphi_i(e_2)$ are solutions of the decoupled Schrödinger equation (7) in which the terms associated to the coefficients $\vec{c}^{\,1}_{ij}$ and c^2_{ij} are equal to zero, and where $E_i(e_2)$ is replaced by $\widehat{E}_i + \widehat{U}_2(e_2)$.

In more general cases, quantum devices are described by a more realistic potential structure since the potential U_2 depends on the e_1 confined direction too. Now, assuming that the potential $U_2(e_1,e_2)$ slowly

varies in the e_1 confined direction in the region Ω_1 where the electron gas is mainly located (see for example [42]), then we can write

$$U_2(e_1,e_2) = \widehat{U}_2(e_2) + u(e_1,e_2), \tag{55}$$

where u is a very small potential ($\|u\| \ll 1$) in the vicinity of the electron gas

$$\int_{\Omega_1} u(e_1,e_2)|\chi_i(e_1;e_2)|^2 \,\mathrm{d}e_1 \ll 1. \tag{56}$$

This condition is an appropriate one for real devices with a strong confinement.

Using the first order stationary perturbation theory, we obtain

$$E_i(e_2) = \widehat{E}_i + \int_{\Omega_1} (U-U_1)|\widehat{\chi}_i|^2 \,\mathrm{d}e_1 + \mathrm{O}(\|u\|^2), \tag{57}$$

and

$$\chi_i(e_1;e_2) = \widehat{\chi}_i(e_1) + \sum_{j\neq i}\left(\int_{\Omega_1}(U-U_1)\widehat{\chi}_i\widehat{\chi}_j \,\mathrm{d}e_1\right)\frac{\widehat{\chi}_j}{\widehat{E}_i-\widehat{E}_j} + \mathrm{O}(\|u\|^2). \tag{58}$$

The model that we call the quasi-full dimensional model consists in solving the Schrödinger equations (54) in the e_1 confined direction as well as a decoupled Schrödinger equation in the e_2 transport direction, where $E_i(e_2)$ is replaced by its approximation given in (57). The latter equation is given in Ω_2 by

$$-\frac{\hbar^2}{2m^*}\Delta_2\varphi^i_{\varepsilon,Q}(e_2) + \left(\int_{\Omega_1}(U-U_1)|\widehat{\chi}_i|^2 \,\mathrm{d}e_1\right)\varphi^i_{\varepsilon,Q}(e_2) = (\varepsilon-\widehat{E}_i)\varphi^i_{\varepsilon,Q}(e_2), \tag{59}$$

where the terms associated to the coefficients $\vec{c}^{\,1}_{ij}$ and c^2_{ij} are taken equal to zero. This latter approximation will be justified in the next section. The expressions of the open boundary conditions for this Eq. (59) are described in Appendix B for the 3D electron waveguide devices presented in Table 1.

In the full dimensional problem, we had to solve only one Schrödinger equation (4) in the domain Ω while the quasi-full dimensional model involves for each mode i to solve one Schrödinger equation (59) in the reduced domain Ω_2 (the single eigenvalue problem (54) is computed only once).

5.1. Error estimates of the quasi-full dimensional model

In this section, we propose to give formally an error estimates for the quasi-full dimensional model versus the full dimensional one. In order to simplify our purpose, this estimation is limited to the non-self-consistent problem (the potential is given) and the boundary conditions are assumed the same for the two models (for example we set $u = 0$ in this boundary regions). A more rigorous mathematical analysis of the asymptotic model can be found in [45,46].

Proposition. *Denoting $\varphi^{i_0}_{\varepsilon}$ and $\varphi^{i_0}_{\varepsilon,Q}$ the wave functions solutions of Eq. (7), respectively, with and without the terms associated to the $\vec{c}^{\,1}_{ij}$ and c^2_{ij} coefficients, and where i_0 means the only incoming mode in the left or the right contact, then we can formally write*:

$$\varphi^i_{\varepsilon} = \mathrm{O}(\|u\|) \quad \textit{for } i \neq i_0, \tag{60}$$

$$\varphi^i_{\varepsilon,Q} = 0 \quad \textit{for } i \neq i_0, \tag{61}$$

$$\varphi^{i_0}_{\varepsilon,Q} = \varphi^{i_0}_{\varepsilon} + \mathrm{O}(\|u\|^2). \tag{62}$$

Proof. For a very small potential u ($u \ll 1$), the functions φ^i_ε and $\varphi^i_{\varepsilon,\mathrm{Q}}$ are equal each others at least at the first approximation order of u

$$\varphi^i_{\varepsilon,\mathrm{Q}} = \varphi^i_\varepsilon + \mathrm{O}(\|u\|). \tag{63}$$

The functions $\varphi^i_{\varepsilon,\mathrm{Q}}$ are solutions of Eq. (7) where the terms associated to the coefficients $\vec{c}^{\,1}_{i0i}$ and $b_{i_{0i}}$ are equal to zero, then it means to have $\varphi^i_{\varepsilon,\mathrm{Q}} = 0$ for $i \neq i_0$ (the wave function is physically carried on the only incoming mode i_0). The result (60) is then obtained using (63).

Using the stationary perturbation theory (58), one can show that the coefficients $\vec{c}^{\,1}_{i_0 i}$ defined in Eq. (8) can be rewritten as

$$\vec{c}^{\,1}_{i_0 i} = \frac{1}{\widehat{E}_i - \widehat{E}_{i_0}} \int_{\Omega_1} \widehat{\chi}_i \widehat{\chi}_{i_0} \nabla_2 U(e_1, e_2)\, \mathrm{d}e_1 + \mathrm{O}(\|u\|^2). \tag{64}$$

This coefficient $\vec{c}^{\,1}_{i_0 i}$ is also of the first order of u (since the the derivative of the potential U is of the first order of u). Therefore, the first term of the right member in the equation of the $c^2_{i_0 i}$ coefficients (10) is of the first order of u if $i \neq i_0$ (equal to zero else since $\vec{c}^{\,1}_{i_0 i_0} = 0$). While the second term is of the second order of u using (58). To summarize the approximation orders of the coefficients $\vec{c}^{\,1}_{i_0 i}$ and $c^2_{i_0 i}$ are given by

$$\vec{c}^{\,1}_{i_0 i_0} = 0, \quad \vec{c}^{\,1}_{i_0 i} = \mathrm{O}(\|u\|), \tag{65}$$

$$c^2_{i_0 i_0} = \mathrm{O}(\|u\|^2), \quad c^2_{i_0 i} = \mathrm{O}(\|u\|). \tag{66}$$

Therefore, all the non-diagonals terms in Eq. (7) associated to the coefficients $\vec{c}^{\,1}_{i_0 i}$ and $c^2_{i_0 i}$ are of the second approximation order of u since with the result of Eq. (60), the non-diagonal function φ^i_ε is of the first approximation order of u. Since all the neglected terms in the quasi-full dimensional model are of the second approximation order of u (the diagonal $c^2_{i_0,i_0}$ term is also of the second approximation order of u), then the functions $\varphi^{i_0}_\varepsilon$ and $\varphi^{i_0}_{\varepsilon,\mathrm{Q}}$ are equal each others at the second approximation order of u. □

Proposition. *Denoting $n(e_1,e_2)$ and $n_Q(e_1,e_2)$ the electron densities associated, respectively, to the full model and the quasi-full dimensional model, then we can formally write*:

$$n_Q(e_1, e_2) = n(e_1, e_2) + \mathrm{O}(\|u\|). \tag{67}$$

Denoting $n_s(e_1,e_2)$ and $n_{s,Q}(e_1,e_2)$ the surface densities which are defined by the integration of the densities over the confined direction, then

$$n_{s,Q}(e_1, e_2) = n_s(e_1, e_2) + \mathrm{O}(\|u\|^2). \tag{68}$$

Proof. The calculation of the electron density is related to a summation of the amplitude of the wave functions over all the energy configurations of the incoming waves in the devices. If we use the symbol '$\sim$' to denote this sum, then, using the result (6), the expressions of the electron densities are equivalent to

$$\begin{aligned} n(e_1, e_2) \sim\ & |\varphi^{i_0}_\varepsilon(e_2)|^2 |\chi_{i_0}(e_1; e_2)|^2 + \sum_{i \neq i_0} |\varphi^i_\varepsilon(e_2)|^2 |\chi_i(e_1; e_2)|^2 \\ & + 2\Re\left(\sum_j \sum_{i>j} \varphi^j_\varepsilon(e_2) \varphi^i_\varepsilon(e_2) \chi_j(e_1; e_2) \chi_i(e_1; e_2) \right), \end{aligned} \tag{69}$$

$$n_Q(e_1, e_2) \sim |\varphi^{i_0}_{\varepsilon,Q}(e_2)|^2 |\chi_{i_0}(e_1; e_2)|^2, \tag{70}$$

where $\Re(y)$ is the real part of y and i_0 denotes the only one incoming mode. The last term of Eq. (69) is defined as an interference term. A term in the first order of u appears in this interference term for $j = i_0$ or $i = i_0$ when the wave functions in Eq. (69) are replaced by the results (60) and (62). The result (67) is then established.

The surface densities are obtained by integrating the densities over the e_1 confined direction

$$n_s(e_2) \sim |\varphi_\varepsilon^{i_0}(e_2)|^2 + \sum_{i \neq i_0} |\varphi_\varepsilon^{i}(e_2)|^2, \tag{71}$$

$$n_{s,\underline{Q}}(e_2) \sim |\varphi_{\varepsilon,\underline{Q}}^{i_0}(e_2)|^2, \tag{72}$$

where the interference term in Eq. (71) disappears. The result (68) is obtained using (60) and (62). □

Proposition. *Denoting I_p and $I_{p,\underline{Q}}$ the current densities at the contact p associated, respectively, to the full model and the quasi-full dimensional model, then we can formally write*:

$$I_{p,\underline{Q}} = I_p + \mathrm{O}(\|u\|^2). \tag{73}$$

Proof. As shown for the expression the electron densities in (69) and (70), one can show that the current densities can be given by the following expressions:

$$I_p \sim \int_{\gamma_p} \Im\left\{ \varphi_\varepsilon^{i_0}(e_2) \frac{\partial \varphi_\varepsilon^{i_0}(e_2)}{\partial e_{2\|}^p}|_{\gamma_p} \right\} \mathrm{d}\gamma_p + \sum_{i \neq i_0} \int_{\gamma_p} \Im\left\{ \varphi_\varepsilon^{i}(e_2) \frac{\partial \varphi_\varepsilon^{i}(e_2)}{\partial e_{2\|}^p}|_{\gamma_p} \right\} \mathrm{d}\gamma_p, \tag{74}$$

$$I_{p,\mathrm{Q}} \sim \int_{\gamma_p} \Im\left\{ \varphi_{\varepsilon,\mathrm{Q}}^{i_0}(e_2) \frac{\partial \varphi_{\varepsilon,\mathrm{Q}}^{i_0}(e_2)}{\partial e_{2\|}^p}|_{\gamma_p} \right\} \mathrm{d}\gamma_p, \tag{75}$$

where $\Im(y)$ is the imaginary part of y, i_0 denotes the only one incoming mode, γ_p is the trace of the contact p in the domain Ω_2, and $e_{2\|}^p$ is the exterior normal derivative on γ_p. The result (73) is obtained using (60) and (62). □

To summarize, the error estimate results show that the obtained electron densities are in the first order between the quasi-full dimensional model and the full model while the surface densities are in the second order between the two models. In our case, the latter quantity is more significant since the use of the quasi-full dimensional model is related to devices with a strong confinement for the electron gas (with slowly variation of the potential U_2 in the e_1 direction). Moreover, the error estimate on the current densities is also in the second order. Therefore, the quasi-full dimensional model is a good approximation to the full one.

5.2. Remarks

We conclude this section with some remarks about the quasi-full dimensional model:

- In practice, the quasi-full dimensional model only involves to solve one eigenvalue problem on e_1 and the size of the matrix used to solve the wave function $\varphi_{\varepsilon,\mathrm{Q}}(e_2)$ is divided by the number of coupling modes M compared to the solution $\varphi_\varepsilon(e_2)$ of the coupled Schrödinger problem in the subband representation.
- In order to solve the eigenvalue problem (54), the quasi-full dimensional model requires to know the potential U_1 which mainly carries the dependence on the potential U in the e_1 confined direction. In the other words, the vertical potential U_1 is a potential which "must be seen" by the electrons. In practice, we define the potential U_1 as a weighted average in the e_2 direction of the potential U

$$U_1(e_1) = \frac{\int_{\Omega_2} U(e_1, e_2) n_s(e_2)\, de_2}{\int_{\Omega_2} n_s(e_2)\, de_2}. \tag{76}$$

In order to ignore the contributions of deserted zones in the e_2 direction, the chosen weight is the surface density n_s which depends on the self-consistent process.

- In the whole domain, the wave function of the quasi-full dimensional model $\Psi_{E,Q}(e_1,e_2)$ is given using (6), (58) and (61) by

$$\Psi_{\varepsilon,Q} = \varphi^i_{\varepsilon,Q}(e_2)\left(\widehat{\chi}_i(e_1) + \sum_{j\neq i}\left(\int_{\Omega_1}(U-U_1)\widehat{\chi}_i(e_1)\widehat{\chi}_j(e_1)\, de_1\right)\frac{\widehat{\chi}_j(e_1)}{\widehat{E}_i - \widehat{E}_j}\right), \tag{77}$$

where i is chosen as the only one incoming mode. In this expression, we can show that the second approximation order on the surface density and the current density is still preserved, even if the second term of the latter expression is neglected. The model is presented in [23].

5.3. The quasi-full dimensional model applied to the electron waveguide devices

The electron waveguide devices are often based on the local constriction (by means of split gates on the top) of the 2D electron gas of a very high mobility heterostructure. In this device, the electron transport can be considered to be ballistic at low temperatures. The electrical properties depend on quantum interferences and they become controllable by the bias voltage applied to the gates and/or by the potential difference between the chemical potentials of the input and output terminal of the devices. Based on these principles, different structures, for example T-stub, directional couplers (Fig. 19), Y-branch, rings, and crosses, etc. have been proposed in order to reproduce the passive properties of a classical microwave electromagnetic waveguide in the field of quantum mechanics with the aim of fabricating quantum transistors, electronic switches, or multiplexers.

A 3D solver for the open Schrödinger–Poisson system was implemented. However, the subband decomposition of the problem is expected to drastically decrease the simulation time. Assuming that the electron gas is both confined in the y and z directions as for the T-stub case, and the quantum coupler case (see Fig. 20), then the wave functions can be expanded in a multi-mode basis which depends on the x transport direction (see the electron waveguide device type II in Table 1). Therefore, the subband decomposition approach requires to solve a large number of 2D eigenvalue problems in the y,z directions along the x direction. By

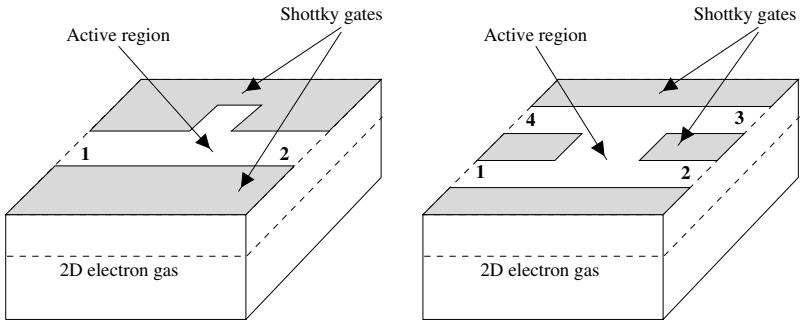

Fig. 19. Schematics of a T-stub (on the left) and a quantum directional coupler (on the right). These devices are composed by four semiconductor layers with a Shottky gates on the top (Gate, GaAs, n-AlGaAs, AlGaAs, GaAs-substrate). The T-stub device has two electron waveguides while the quantum directional has four. The electron transport takes place in a 2D plane at the AlGaAs/GaAs interface.

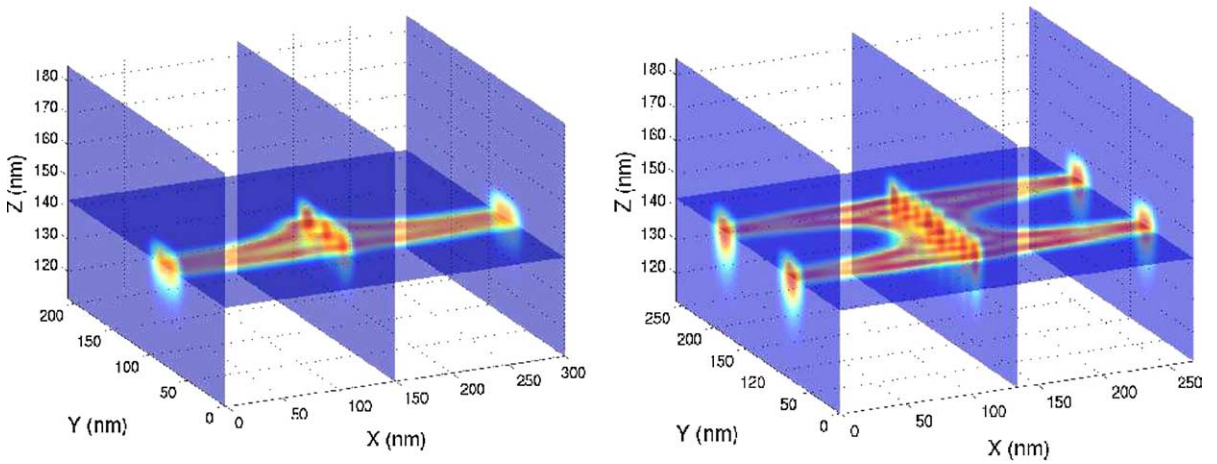

Fig. 20. The density profile after the quantum model convergence for the T-stub (left) and the directional coupler (right). The electron gas is located a few nanometer below the AlGaAs/GaAs interface (at $z = 150$ nm).

this procedure, the eigenpairs can also be computed in a reduced domain in y,z where the electron gas is a priori assumed to be localized. This general method was recently applied to study 3D silicon nanowire transistors [43,44], and a challenging trace minimization algorithm for the computation of generalized eigenvalue problems are proposed in [20].

However, due to the strong confinement of the electron gas in the z direction, we propose to use a quasi-3D approximation to describe the quantum transport. The wave functions can be now expanded in a multimode basis which depends on the transport x,y directions (see the electron waveguide device type I in Table 1). In this way, more complex configurations for the devices than those presented in Fig. 19 could be considered (for instance with both waveguides in the x and y directions). The method would only require to solve a large number of 1D eigenvalue problems in the z direction. However, the electron gas is not located in the regions under the gates since the electron energy is then smaller than the potential energy. These classically forbidden regions involve vanishing waves in the quantum regime and the probability to find an electron is then next to zero. To overcome these difficulties and since all the electrons are mainly concentrated in a quantum well next to the AlGaAs/GaAs interface (with the assumption of a slowly variation of the potential U_2 in the z direction), it appears suitable to use the quasi-3D approximation of the model which is described in the previous section.

Finally, the derivation of the quasi-3D model simultaneously takes into account the confinement of electrons in the 2DEG and the fact that the electrostatic potential is completely 3D (using the 3D Poisson's equation). A detailed description about this model was achieved in [23], where a quantitative satisfactory agreement between the numerical results was obtained by the two models (3D and quasi-3D) through the simulation of the T-stub device and the quantum directional coupler at equilibrium (see Fig. 20). The convergence speed of the quasi-3D quantum model is approximately 200 times faster than the 3D model. The results obtained with the quasi-full dimensional model approach on a 3D silicon nanowire transistors in [43,44] (which is called a fast uncoupled mode approach) also show a drastic improvement in the simulation times.

6. Conclusions

The subband decomposition approach described in this article allows to transform a 2D or 3D transport problem in the nanostructures into a less dimensional one. In this approach, the confined directions for the

electron gas are not explicitly considered in the transport problem, and the obtained transport Schrödinger equation exhibits coupling modes effects along the transport direction through extra-diagonal terms. A detailed description of the numerical implementation in the nanoscale MOSFETs was presented and the obtained results showed that much less numerical efforts is required to get the I–V characteristics with the subband approach than with the full 2D one. Moreover, the performance of the method depends on the number of coupling modes which are taken into account. This number has to be adapted to each particular device structure. The derivation of an asymptotic approach of the subband decomposition was also presented and applied to 3D electron waveguide devices. The reduction of the simulation time of the quasi-3D model compared to the full 3D one is quite significant.

Further improvements to enhance the numerical efficiency of the self-consistent procedure can be done at (i) the discretization level using for example new FEM basis functions as presented in [47], where a WKB basis function allows to drastically reduce the number of slices, (ii) the numerical algorithm level using new parallel techniques to solve the linear systems and the generalized eigenvalues problems [20].

Finally, let us mention that a general treatment for the simulation of the quantum transport problems was implemented in our multi-dimensional parallel finite element code NESSIE. The simulator is flexible enough to study a wide range of characteristics of complex geometries using various self-consistent model as: semi-classical (Thomas–Fermi), hybrid (Schrödinger and Thomas–Fermi), full quantum model (Schrödinger or NEGF), quantum model based on the subband decomposition with or without asymptotic approaches (space-mode approach and quasi-full dimensional model).

Acknowledgments

We gratefully acknowledge helpful discussions with Prof. O. Vanbésien, Dr. Mireille MOUIS and Dr. Florian Méhats. This work benefited from the financial support of the European IHP network Ref. HPRN-CT-2002-00282 entitled "Hyperbolic and Kinetic Equations: Asymptotics, Numerics, Analysis", the CNRS project "Transport dans les nanostructures" (Action Spécifique STIC), the National Science Foundation through the Network for Computational Nanotechnology Grant EEEC-0228390.

Appendix A. Open boundary conditions for the full model

In the following presentation, we shall assume the existence of a e_2 transport direction in the domain $\partial\Omega_{1,2}$. Open boundary conditions are defined at the contacts with the reservoirs, while we assume that the wave function is vanishing ($\psi_\varepsilon(e_1,e_2) = 0$) elsewhere at the boundaries. Denoting Γ_p the frontier associated to the contact p where the transport direction e_2 can be locally separated into $(e^p_{2\|}, e^p_{2\perp})$. The potential is assumed transverse inside the contact p and along the $e^p_{2\|}$ directions, such as $U(e_1, e^p_{2\perp})$ in these regions. Therefore, if we consider that $\psi_\varepsilon = 0$ on $\partial\Gamma_p$, the $e_{2\perp}$ direction becomes a confined direction too (in addition to the e_1 direction) and then we can define a set of eigenfunctions $\{\widetilde{\chi}^p_i(e_1, e^p_{2\perp})\}$ and eigenvalues $\{E^p_i\}$, associated to each contact p. The wave function in the contact p can be expanded in the latter multi-mode basis as

$$\psi_\varepsilon(e_1, e^p_{2\|}, e^p_{2\perp}) = \sum_{i=1}^{\infty} \varphi^{i,p}_\varepsilon(e^p_{2\|})\widetilde{\chi}^p_i(e_1, e^p_{2\perp}), \tag{A.1}$$

where $\varphi^{i,p}_\varepsilon(e^p_{2\|})$ are the scattering states which are described by plane waves. The propagating or evanescent nature of the waves depends on the following carrier mode conditions:

$$\begin{cases} \varphi^{i,p}_\varepsilon(e^p_{2\|}) = a^{i,p}_\varepsilon \exp(-\mathrm{i}k^p_i(\varepsilon)e^p_{2\|}) + b^{i,p}_\varepsilon \exp(\mathrm{i}k^p_i(\varepsilon)e^p_{2\|}) & \text{if } i \leqslant I_p(\varepsilon), \\ \varphi^{i,p}_\varepsilon(e^p_{2\|}) = b^{i,p}_\varepsilon \exp(-k^p_i(\varepsilon)e^p_{2\|}) & \text{if } i > I_p(\varepsilon), \end{cases} \tag{A.2}$$

where $a_\varepsilon^{i,p}$ is the amplitude of the incoming wave in the contact p and the mode i, $b_\varepsilon^{i,p}$ is the amplitude of the outgoing wave or the coefficient of the evanescent one, I_p is the number of propagating mode of the contact p given by

$$I_p = \sup_i\{\varepsilon > E_i^p\}, \tag{A.3}$$

and k_i^p is the wave vector associated to the mode i of the contact p

$$k_i^p(\varepsilon) = \left(\frac{2m^*}{\hbar^2}|\varepsilon - E_i^p|\right)^{1/2}. \tag{A.4}$$

At the contact Γ_p, the wave function and its derivative along the longitudinal $e_{2\|}^p$ direction are, respectively, given by (we note $e_{2\|}^p = 0$)

$$\psi_{\varepsilon|_{\Gamma_p}} = \sum_{i=1}^{I_p}(a_\varepsilon^{i,p} + b_\varepsilon^{i,p})\widetilde{\chi}_i^p(e_1, e_{2\perp}^p) + \sum_{i=I_p+1}^{\infty} b_\varepsilon^{i,p}\widetilde{\chi}_i^p(e_1, e_{2\perp}^p) \tag{A.5}$$

and

$$\partial_{e_{2\|}^p}\psi_{\varepsilon|_{\Gamma_p}} = \sum_{i=1}^{I_p} ik_i^p(\varepsilon)(-a_\varepsilon^{i,p} + b_\varepsilon^{i,p})\widetilde{\chi}_i^p(e_1, e_{2\perp}^p) - \sum_{i=I_p+1}^{\infty} k_i^p(\varepsilon)b_\varepsilon^{i,p}\widetilde{\chi}_i^p(e_1, e_{2\perp}^p). \tag{A.6}$$

Denoting the operator $\Pi_i^p(\psi_\varepsilon)$ as the projection of the wave function ψ_ε on the ith mode (projector), then we obtain

$$\Pi_i^p(\psi_\varepsilon) = \langle\psi_\varepsilon|\widetilde{\chi}_i^p\rangle\widetilde{\chi}_i^p = \begin{cases} a_\varepsilon^{i,p} + b_\varepsilon^{i,p} & \text{if } i \leqslant I_p(\varepsilon), \\ b_\varepsilon^{i,p} & \text{if } i > I_p(\varepsilon), \end{cases} \tag{A.7}$$

where $\langle A|B\rangle$ denoting the scalar product between the function A and B on $L^2(\Gamma_p)$ in the ket-bra notation.

Finally, the open boundary conditions of ψ_ε on Γ_p can be rewritten, independently of the unknown coefficient $b_\varepsilon^{i,p}$, as

$$\mathscr{D}_{\varepsilon,p}(\psi_\varepsilon) + \mathscr{N}_p(\psi_\varepsilon) = A_{\varepsilon,p}, \tag{A.8}$$

where $A_{\varepsilon,p}$ and the operators $\mathscr{D}_{\varepsilon,p}, \mathscr{N}_p$ are defined as follows:

$$\mathscr{D}_{\varepsilon,p}(\psi_\varepsilon) = -\sum_{i=1}^{I_p} \mathrm{i}k_i^p(\varepsilon)\Pi_i^p(\psi_\varepsilon) + \sum_{i=I_p+1}^{\infty} k_i^p(\varepsilon)\Pi_i^p(\psi_\varepsilon), \tag{A.9}$$

$$\mathscr{N}_p(\psi_\varepsilon) = \partial_{e_{2\|}^p}\psi_{\varepsilon|_{\Gamma_p}}, \tag{A.10}$$

$$A_{\varepsilon,p} = -2\sum_{i=1}^{I_p} \mathrm{i}k_i^p(\varepsilon)a_\varepsilon^{i,p}\widetilde{\chi}_i^p(e_1, e_{2\perp}^p). \tag{A.11}$$

Appendix B. Open boundary conditions for the models using the subband decomposition approach

Assuming the existence of e_1 confined directions in the whole domain $\Omega_{1,2}$, then the particular devices whose these boundary conditions could be account for are described in Table 1 as the 3D electron waveguide devices type I and II, as well as the 2D nanoscale MOSFETs. We also assume that the eigenfunctions

$\chi_i(e_1;e_2)$ used in the subband decomposition in (6) are exactly the same functions at the contacts p than the functions $\widetilde{\chi}_i^p(e_1, e_{2\perp}^p)$ defined in Appendix A. However, for the 3D electron waveguides device type II and for the 2D nanoscale MOSFETs, the local confined $e_{2\perp}^p$ direction of $\widetilde{\chi}_i^p$ is already included in the e_1 directions of χ_i. We, respectively, obtain for these devices in the real space $\widetilde{\chi}_i^p(y,z) \equiv \chi_i(y,z;x_p)$ and $\widetilde{\chi}_i^p(z) \equiv \chi_i(z;x_p)$.

For these two cases, the boundary conditions on $\partial\Omega_2$ of the wave function $\varphi_\varepsilon^i(e_1,e_2)$ solution of Eq. (7) can be derived from the treatment of the boundary conditions for full problem in Appendix A. We define $\Gamma_{2,p} = \partial\Omega_2 \cap \Gamma_p$ the frontiers where the open boundary conditions are now locally, then the projector $\Pi_p^i(\psi_\varepsilon)$ in Eq. (A.7) can be rewritten using Eq. (6), as

$$\Pi_i^p(\psi_\varepsilon) = \varphi_\varepsilon^{i,p}(e_{2\|}^p)\chi_i^p(e_1), \tag{B.1}$$

where $\varphi_\varepsilon^{i,p}(e_{2\|}^p)$ is locally the wave function $\varphi_\varepsilon^i(e_2)$ at the contacts p on $\Gamma_{2,p}$. The open boundary conditions for the wave function φ_ε^i are now defined for all the mode i as

$$\mathscr{D}_{\varepsilon,p}^i(\varphi_\varepsilon^i) + \mathscr{N}_p^i(\varphi_\varepsilon^i) = A_{\varepsilon,p}^i, \tag{B.2}$$

where

$$\mathscr{D}_{\varepsilon,p}^i(\varphi_\varepsilon^i) = \langle \chi_i^p | \mathscr{D}_{\varepsilon,p}(\Psi_\varepsilon)\rangle = -\mathrm{i}k_i^p(\varepsilon)\varphi_\varepsilon^{i,p}\delta_{i\leqslant I_p(\varepsilon)} + k_i^p(\varepsilon)\varphi_\varepsilon^{i,p}\delta_{i>I_p(\varepsilon)}, \tag{B.3}$$

$$\mathscr{N}_p^i(\varphi_\varepsilon^i) = \langle \chi_i^p | \mathscr{N}_p(\Psi_\varepsilon)\rangle = \partial_{e_{2\|}^p}\varphi_\varepsilon^{i,p}{}_{|_{\Gamma_{2,p}}}, \tag{B.4}$$

$$A_{\varepsilon,p}^i = \langle \chi_i^p | A_{\varepsilon,p}\rangle = -2\mathrm{i}k_i^p(\varepsilon)a_\varepsilon^{i,p}\delta_{i\leqslant I_p(\varepsilon)}. \tag{B.5}$$

Finally, we will note that the $\chi_i(z;x,y)$ eigenfunctions on z obtained with the subband decomposition approach in (6) for the 3D electron waveguide device type I (even through the asymptotic described in Section 5) are not equal at the contacts p to the eigenfunctions $\widetilde{\chi}_i^p(y,z)$ on y,z. In this case, the expression of the projector (A.7) cannot be simplified as previously.

References

[1] F.G. Pikus, K. Likharev, Appl. Phys. Lett. 71 (1997) 3661.
[2] H.S.P. Wong, D.J. Frank, P.M. Solomon, C.H.J. Wann, J.J. Welser, Proc. IEEE 87 (1999) 537.
[3] F. Assad, Z. Ren, D. Vasileska, S. Datta, M. Lundstrom, IEEE Trans. Electron. Devices 47 (2000) 232.
[4] Y. Naveh, K. Likharev, IEEE Electron. Device Lett. 21 (2000) 242.
[5] F. Sols, M. Macucci, U. Ravaioli, K. Hess, J. Appl. Phys. 66 (1989) 3892.
[6] E. Tekman, S. Ciraci, Phys. Rev. B 39 (1989) 8772.
[7] J.A. Del Alamo, C.C. Eugster, Appl. Phys. Lett. 51 (1990) 78.
[8] N. Tsukada, A.D. Wieck, K. Ploog, Appl. Phys. Lett. 56 (1990) 2527.
[9] J. Wang, H. Guo, R. Harris, Appl. Phys. Lett. 59 (1991) 3075.
[10] A. Weisshaar, J. Lary, S.M. Goodnick, V.K. Tripathi, J. Appl. Phys. 70 (1991) 355.
[11] G. Xu, Y. Yang, P. Jiang, J. Appl. Phys. 74 (1993) 6747.
[12] O. Vanbésien, D. Lippens, Appl. Phys. Lett. 65 (1994) 2439.
[13] J.J. Harris, C.T. Foxon, K.W.J. Barnham, D.E. Lacklison, J. Hewett, C. White, J. Appl. Phys. 61 (1986) 1219.
[14] L. Pfeiffer, K.W. West, H.L. Stormer, K.W. Baldwin, Appl. Phys. Lett. 55 (1989) 1888.
[15] K. Natori, J. Appl. Phys. 76 (1994) 4879.
[16] A. Svizhenko, M.P. Anantram, T.R. Govindan, B. Biegel, R. Venugopal, J. Appl. Phys. 91 (2002) 2343.
[17] R. Venugopal, Z. Ren, S. Datta, M.S. Lundstrom, D. Jovanovic, J. Appl. Phys. 92 (2002) 3730.
[18] E. Polizzi, S. Datta, IEEE Conference on Nanotechnology (IEEE-NANO), 2003.
[19] P. Havu, V. Havu, M.J. Puska, R.M. Nieminen, Phys. Rev. B 69 (2004) 115325.
[20] E. Polizzi, H. Sun, A. Sameh, Nanotech2004 Conference, 2004.

[21] S. Datta, Cambridge University Press, 1995 (Paperback Edition, 1997).
[22] M.V. Fischetti, Phys. Rev. B 59 (1999) 4901.
[23] E. Polizzi, N. Ben Abdallah, Phys. Rev. B 66 (2002) 245301.
[24] G.W. Bryant, Phys. Rev. B 44 (1991) 12837.
[25] M.L. Laughton, J.R. Barker, J.A. Nixon, J.H. Davies, Phys. Rev. B 44 (1991) 1150.
[26] J.A. Nixon, J.H. Davies, H.U. Baranger, Superlattice Microstruct. 9 (1991) 187.
[27] E. Cataño, G. Kirczenow, Phys. Rev. B 45 (1992) 1514.
[28] A. Brataas, K.G. Chao, Mod. Phys. Lett. B 7 (1993) 1021.
[29] F.A. Maao, I.V. Zozulenko, E.H. Hauge, Phys. Rev. B 50 (1994) 17320.
[30] D.S. Fisher, P.A. Lee, Phys. Rev. B 23 (1981) 6851.
[31] S. Rotter, J.Z. Tang, L. Wirtz, J. Trost, J. Burgdöfer, Phys. Rev. B 62 (2000) 1950.
[32] C.S. Lent, D.J. Kirkner, J. Appl. Phys. 67 (1990) 6353.
[33] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, LAPACK User's Guide, 3rd ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.
[34] R.W. Freund, SIAM J. Sci. Stat. Comput. 12 (1992) 425.
[35] H.K. Gummel, IEEE Trans. Electron. Devices 11 (1964) 455.
[36] A.S. Spinelli, A. Benvenutti, A. Pacelli, IEEE Trans. Electron. Devices 45 (1998) 1342.
[37] M. Macucci, A. Galik, U. Ravaioli, Phys. Rev. B 52 (1995) 5210.
[38] F. Rana, S. Tiwari, D.A. Buchanan, Appl. Phys. Lett. 69 (1996) 1104.
[39] S. Datta, Superlattice Microstruct. 28 (2000) 253.
[40] M. Büttiker, Y. Imry, R. Landauer, S. Pinhas, Phys. Rev. B 31 (1985) 6207.
[41] S. Datta, M.J. McLennan, Rep. Prog. Phys. 53 (1990) 1003.
[42] J.H. Davies, Semicond. Sci. Technol. 3 (1988) 995.
[43] J. Wang, E. Polizzi, M. Lundstrom, IEDM Conference, 2003.
[44] J. Wang, E. Polizzi, M. Lundstrom, J. Appl. Phys. (submitted).
[45] O. Pinaud, J. Math. Phys. 45 (2004) 2029.
[46] N. Ben Abdallah, F. Méhtas, O. Pinaud, SIAM J. Math. Anal. (to appear).
[47] N. Ben Abdallah, O. Pinaud, J. Comput. Phys. (submitted).