Quasi-three-dimensional Green's-function simulation of coupled electron waveguides

M. Macucci
Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Via Diotisalvi 2, I-56126 Pisa, Italy

A. Galick and U. Ravaioli
Beckman Institute, University of Illinois, 405 N Mathews, Urbana, Illinois 61801
(Received 21 February 1995)

A quasi-three-dimensional (3D) simulation of a quantum waveguide coupler has been performed, computing the self-consistent transverse potential along the electron waveguides and then solving the transport problem with a modified recursive Green's-function method. Results have been obtained for the tunneling conductance between the two waveguides as a function of coupling length and gate biases. A clear structure of conductance peaks is observed, strongly dependent on both the drain and the source biases. Such dependence has been investigated in greater detail for an idealized model, allowing a fast numerical simulation. A ridgelike conductance pattern has been obtained, which can be interpreted as a characteristic signature to be looked for when searching for the evidence of 1D-to-1D tunneling in experimental data.

I. INTRODUCTION

The availability of advanced nanolithographic techniques has allowed the fabrication of microstructures in which the wave nature of electrons becomes important, and new device functionalities based on quantum effects can be achieved. A particularly interesting idea is the one of the quantum field-effect directional coupler proposed by del Alamo and Eugster.[1] It is substantially the electronic equivalent of the directional couplers that have been developed and are being used in the field of integrated optics. Two electron waveguides obtained by lateral confinement with metallic gates on top of a heterostructure, which provides the vertical confinement are brought into close proximity, so that tunneling between them is made possible. The height and width of the potential barrier between the two waveguides depends on the bias applied to a central gate, which can, therefore, be used to modulate the amount of coupling between the two channels. The widespread interest in such a device is motivated both by the potential for applications as a switching element in digital circuits and by the fabrication and simulation challenges it poses.

Theoretical investigations of simple[2–4] and more elaborate[5–7] models have been performed, and several parameters, such as the transfer length and the tunneling conductance have been computed. At the same time, important experimental results have been achieved, with the clear observation of a one-dimensional–to–two-dimensional (1D-to-2D) tunneling current[8–10] and the possible detection of 1D-to-1D tunneling.[11]

In Ref. 3, 1D-to-2D tunneling has been modeled for a leaky electron waveguide, while Refs. 2 and 4 deal with 1D-to-1D tunneling between two electron waveguides, each in the single-mode regime. Coupling between the waveguides is assumed to be constant along their whole, infinite length. A structure closer to those fabricated by Eugster and del Alamo[10,11] is investigated in Refs. 6 and 7, where two symmetric multimode waveguides coupled for a finite length are considered. In Ref. 5, a simplified coupler structure was studied in the monomode regime at 4.2 K, using realistic potentials obtained from a self-consistent solution for 2D slices.

We have been interested in a detailed simulation of the earlier structure studied by Eugster and del Alamo,[10] with the purpose of understanding why 1D-to-1D tunneling could not be observed in this case and of determining the main features that should reveal it in experimental results. The gate geometry that we study has been slightly modified with respect to the one of the actual device, in order to simplify the numerical simulation. Our model is quasi-three-dimensional in the sense that we subdivide the structure into a number of slices perpendicular to the direction of current flow, then a 2D Schrödinger and Poisson problem is solved in each slice and, finally, transport properties are evaluated by means of a modified recursive Green's-function formalism. Our approach represents a good approximation, as long as the variation of the potential landscape along the longitudinal direction, perpendicular to the slices, is slow enough (quasiadiabatic). In this way, we retain most of the relevant features of the real structure, while making the problem computationally manageable.

Even in the quasiadiabatic approximation, the amount of computational work is still very large: simulations in which the bias voltage of only one electrode is swept over a range of values are feasible, but simultaneous variation of two parameters, such as drain and source bias, would require too much CPU time. For this reason, we have also studied an idealized model, which retains most of the relevant features, including some of the gate bias effects. Numerical calculations for a range of source and drain values show an interesting pattern for the tunneling conductance. We believe that a similar pattern should be

the typical signature of 1D-to-1D tunneling to be looked for in experimental results.

In Sec. II, we present a detailed description of the device models that we have used in our simulation, including a discussion of the techniques applied for the self-consistent solution of the 2D Schrödinger-Poisson problem. The idealized model is also presented in this section. The modified recursive Green's-function technique is presented in Sec. III, with an emphasis on the differences with respect to the standard procedure. Finally, numerical results are reported and discussed in Sec. IV, both for the realistic and the idealized models.

II. DEVICE MODELS

We have modeled the heterostructure used in the experiments of Ref. 8. The layer arrangement is shown in Fig. 1: a thick layer of intrinsic GaAs is grown on top of an undoped GaAs substrate, followed by a 7.5 nm spacer layer of undoped $Al_xGa_{1-x}As$. Then a 42 nm layer of n doped $Al_xGa_{1-x}As$ (the actual dopant concentration will be discussed later in this section) follows, capped by a 5 nm layer of GaAs. The dashed lines at the surface of the heterostructure indicate the presence of surface charge, whose density must be determined, while the dashed line near the intrinsic $Al_xGa_{1-x}As$-GaAs interface represents the 2D electron gas formed by modulation doping.

Three metal gates have been indicated on top of the heterostructure. The sizes and relative distances reported in the figure correspond to those in the coupling region, where tunneling between the two waveguides can take place. The central gate, which controls the height and width of the potential barrier separating the waveguides, has been labeled with G and will be referred to in the following as "central gate." One of the lateral gates has been labeled S (source) and the other one D (drain), in analogy with the terminology used for MOS devices. The geometry of the gates in the plane parallel to the heterostructure surface is shown in Fig. 2(a), where all the measures are expressed in micrometers. The shaded areas represent the actual gate geometries. The geometry of the central gate of our model is exactly correspondent to that of the real device, while for the lateral gates we have chosen a slightly different shape, indicated by the dashed lines. In this way, we have kept unaltered the central 0.45 μm long region, where most of the coupling takes place, and limited the number of modes in the outer regions, which would have become unmanageable if we had chosen to model the full transition to the 2D electron gas. A 3D view of the device is shown in Fig. 2(b), where the dark-shaded areas correspond to the cross sections of the electron waveguides.

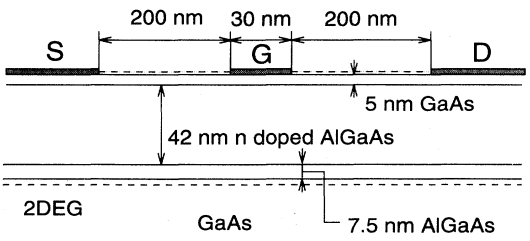

FIG. 1. Layer arrangement of the simulated heterostructure.

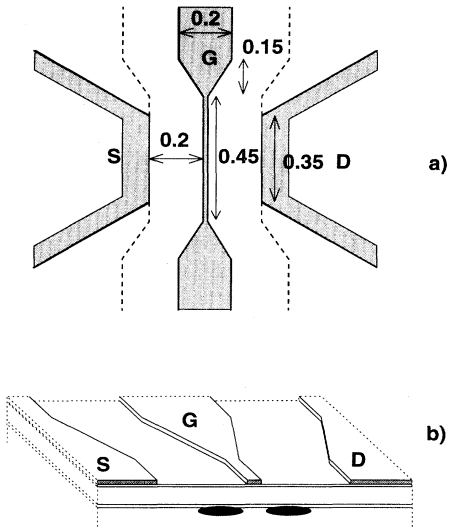

FIG. 2. (a) Gate layout for the actual device (shaded areas) and modified layout used in the simulations. (b) View of a cross section of the device: the dark-shaded regions correspond to the quantum wires.

Coupling between the two quantum wires in the regions with a 0.2 μm central gate is negligible, thus the wires, far from the central region, become independent and are assumed to be connected with electron reservoirs characterized by a well defined chemical potential. If a very small potential difference is applied between any two reservoirs, the conductance between them can be evaluated using the Landauer-Büttiker formula[12,13] (in the hypothesis of purely ballistic transport),

$$G = \frac{2e^2}{h} \sum_{ij} T_{ij}, \tag{1}$$

where e is the electron charge, h Planck's constant, and T_{ij} the square modulus of the transmission coefficient from mode i to mode j. At finite temperatures, this result must be averaged over energy, with the derivative of the Fermi function as a weighing factor:[14]

$$G = \int_0^\infty G(E) \frac{\partial f(E)}{\partial E} dE. \tag{2}$$

In order to compute the transmission coefficients, we must solve a scattering problem across this structure. For such purpose, we need to know the 3D potential landscape. We subdivide the structure into a number of transverse slices (transverse with respect to the direction of current flow), within which the potential can be considered longitudinally constant. As already mentioned in the introduction, we make the simplifying hypothesis of

quasiadiabatic potential variation in the longitudinal direction, so that the solution of the combined Schrödinger and Poisson equations in each section reduces to that of a 2D problem, corresponding to an infinite quantum wire.[15]

As in Ref. 16, the resolution of the transverse modes in each slice is an iteration to self-consistency of the Schrödinger equation for the wave functions and the nonlinear Poisson equation for the electrostatic potential.

The 2D self-consistent problem consists, in the effective mass approximation, of the Schrödinger equation

$$-\nabla\left[\frac{\hbar^2}{2m^*}\nabla\psi_l\right]+[V-E_l]\psi_l=0 \tag{3}$$

and the Poisson equation

$$-\nabla[\epsilon\nabla\phi]=\rho, \tag{4}$$

where the unknowns are the wave functions and energy eigenvalues (ψ_l, E_l) for states in the quantum region, and the 2D electrostatic potential ϕ. Here, m^* is the electron effective mass, ϵ is the dielectric permittivity, $\hbar$ is the reduced Planck constant, and ρ is the charge density. V represents the conduction band potential energy, given by

$$V=-e\phi+V_h+V_{\mathrm{xc}}(n), \tag{5}$$

where V_h is a step function representing the effective potential energy associated with the heterojunction discontinuity, and $V_{\mathrm{xc}}(n)$ is a correction for exchange-correlation effects, which depends on the electron density n confined in the quantum well under the heterojunction and is treated here following a simple local density function approximation.[17,18] Contributions of the image potential are negligible in this material system.

The charge density ρ in Eq. (4) is expressed in terms of the electron density n, the hole density p, and the total density of ionized dopants $N_D^+ - N_A^-$ as

$$\rho=e[p-n+N_D^+-N_A^-]. \tag{6}$$

The electron density in the conduction channel is obtained from the wave functions ψ_l calculated from Eq. (3) as

$$n=\sum_{l=1}^{\infty}N_l\psi_l^2. \tag{7}$$

The occupancy N_l of the lth eigenstate for the 1D electron gas is expressed in terms of the Fermi level E_F, the Boltzmann constant k_B, and the temperature T, through the Fermi-Dirac integral of order $-1/2$, which for GaAs gives

$$N_l=\frac{1}{\pi}\left(\frac{2m^*k_BT}{\hbar^2}\right)^{1/2}F_{-1/2}\left(\frac{E_F-E_l}{k_BT}\right). \tag{8}$$

The Fermi-Dirac integrals in Eq. (8) are evaluated by an efficient rational function approximation.[19]

The partial differential equations (3) and (4) are discretized on a nonuniform rectangular grid using a five-point finite difference scheme based on the box-integration method.[20] This results in sparse symmetric pentadiagonal matrices for both equations. The resulting standard eigenvalue problem for the discretized Schrödinger equation, on the fine grids necessary to resolve the wave functions, has a matrix which is too large for dense solvers such as in EISPACK, but can be effectively solved with projection methods[21] that target only the energy eigenvalues corresponding to significant occupancy N_l. For these calculations, we employ the Chebyshev-Arnoldi method[22] that was originally developed for nonsymmetric eigenvalue problems associated with dielectric waveguides. In this approach, a Chebyshev preconditioning[22] is used to amplify the desired part of the eigenvalue spectrum for faster convergence of the Arnoldi solution method.[21] The Chebyshev preconditioning is particularly efficient because it is implemented by applying a simple three-term recursion.[22] Another effective approach for the solution of the eigenvalue problem is the Ritz iteration procedure RITZIT (Ref. 23) that was applied earlier to quantum wire calculations.[15]

The Poisson equation (4) is highly nonlinear, because the classical densities of holes and ionized impurities are nonlinear functions of the electrostatic potential ϕ. In the structures under consideration, the n-type layers are practically depleted of carriers and the nonlinearity mainly occurs in the p-type substrate. The ionized acceptor density in the substrate is modeled, under the assumption of quasiequilibrium, using the classical statistics as given in Ref. 24,

$$N_A^-=\frac{N_A}{1+4\exp\left(\dfrac{-e\phi+V_h-E_g+E_a-E_F}{k_BT}\right)}, \tag{9}$$

where E_g is the energy band gap and E_a is the acceptor ionization energy. Because of this nonlinear expression, the transition from depletion layer to neutral region deep in the GaAs substrate must be adequately resolved by a local mesh refinement, since the overall solution is strongly affected by the extent of the depletion layer. This problem is particularly acute at the low operation temperature $T = 1.6$ K considered here. Regarding the electron density n in the quantum channel, no closed form expression is available, so n is held constant in the Poisson equation, and is updated only after the Schrödinger equation has been solved.

The nonlinear Poisson equation is solved using a Newton method. At each iteration step a Jacobian linear system is obtained and is solved using an efficient variant of the conjugate gradient method with a red-black reordering as described in Ref. 15. The solution domain for the Poisson equation is extended into the vacuum above the semiconductor, in order to resolve the surface boundaries of the top semiconductor layer. A Schottky barrier of 0.9 eV is considered for the metal contacts.

The iteration procedure for the coupled Schrödinger/Poisson problem is modified depending on how close we are to the solution. Initially, we use a simple fixed-point iteration, where the equations are solved in succession in a decoupled fashion, and an adaptive

underrelaxation is used to stabilize the iteration. We apply the relaxation to the electron density n in the quantum channel, which is fixed in the Poisson equation, but can change dramatically after the Schrödinger equation is solved, and hence is the quantity with the highest sensitivity to variations of the solution. It was found that direct relaxation of the electron density n, rather than of the quantity $\ln(n)$ as done in Ref. 15, is more effective. The convergence of the underrelaxation scheme becomes inefficient when the solution is approached, and the procedure switches to a Newton iteration to accelerate the speed of convergence to self-consistency. We adopt a version of Newton's method which is Jacobian-free and based on the generalized minimum residual method.[25]

Considerable difficulties in relating simulations to measurements are due to the uncertainty in determining the actual doping of the semiconductor layers of the sample, and the amount of surface charge which accumulates at the semiconductor-vacuum interface, which strongly depends on the fabrication process. The numerical model was, therefore, calibrated using indirect measurement information. We first adjusted the uniform $Al_xGa_{1-x}As$ donor doping concentration, so that the threshold for formation of the 2D electron gas under the source contact was $V_S = -0.6$ V as noted in the experiments (V_S is the source electrode voltage). This was done with a 1D self-consistent calculation along a direction perpendicular to the heterojunction under the source. Then, a uniform surface charge density was introduced so that, with the doping selected above, the model reproduced the experimental threshold for the formation of the 1D electron gas in the quantum wires at $V_S = V_D = -2$ V, and $V_G = -0.85$ V, V_D and V_G being the drain and the gate electrode voltages, respectively. Assuming a typical background (unintentional) acceptor doping of $N_A = 5 \times 10^{14}$ cm^{-2} throughout the sample, we determined an $Al_xGa_{1-x}As$ donor doping of $N_D = 10^{18}$ cm^{-3} and a surface charge density equivalent to -3.3×10^{12} cm^{-2}.

We use a nonuniform grid for the solution of the Schrödinger and Poisson problem in each section, finer in the regions of higher electron density, where most of the oscillations of the wave functions take place. Thus, we have a different grid in each section, which represents a problem from the point of view of calculating the eigenfunction overlap integrals needed in the application of the recursive Green's-function method. Interpolation of the wave functions on a common grid leads to loss of orthonormality, which in turn leads to a serious degradation in the precision of the calculated transmission coefficients. In order to avoid this problem, once self-consistency of the Schrödinger-Poisson solution is reached, the resulting potential is interpolated over a finer grid, and the Schrödinger equation is solved again on this new grid, which is common to all slices. The values of the overlap integrals at each interface, together with the eigenvalues in each section, are passed to the code performing the calculation of the transmission coefficients.

For the idealized model, we have considered a simplified 2D potential landscape formed by an input section, a coupling section and an output section. The input and output sections are mirror symmetric and contain two independent waveguides with hard-wall lateral confinement and tilted straight bottom [Fig. 3(a)]. The coupling section is also made up of two waveguides, but the potential wall separating them is of finite height [Fig. 3(b)].

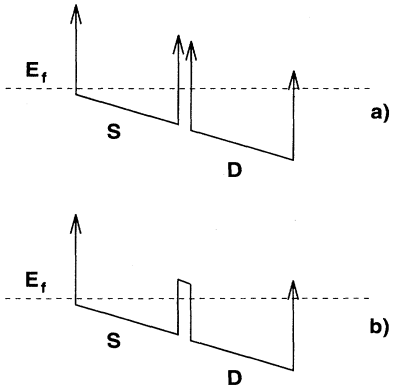

FIG. 3. Transverse potential profile for the ideal model outside the coupling region (a) and in the coupling region (b).

We assume the Fermi level as the potential reference and set the potential values at the left edge of the source waveguide and at the right edge of the drain waveguide. The potential at intermediate points is obtained through linear interpolation, except for the barrier region, where the barrier height is added to the result of the interpolation. Transverse eigenmodes and eigenvalues in all sections are readily obtained by discretizing the 1D Schrödinger equation (this is a globally 2D model, therefore, we have 1D transverse eigenmodes), with a finite differences scheme and computing the eigenvectors and eigenvalues of the resulting tridiagonal matrix with an optimized algorithm.[26] Then the eigenmode overlap integrals and the eigenvalues are passed to the Green's-function code for the calculation of the transmission coefficients, in the same way as for the realistic model.

III. MODIFIED RECURSIVE GREEN'S-FUNCTION METHOD

Our modified recursive Green's-function method is based on the standard tight-binding recursive Green's-function technique often used in the literature for the study of mesoscopic structures.[27–29] In the standard method,[29] the structure to be investigated is filled with a tight-binding square (or cubic, in 3D) lattice with coupling between nearest neighbors. The structure being investigated is subdivided into sections with longitudinally constant transverse potential profile, and the Green's functions for each section are evaluated independently, assuming it closed with hard walls at its ends. Then, coupling between each pair of neighboring sections can be introduced as a perturbation $\hat{V}$ of the Hamiltonian $\hat{H}_0$ of the uncoupled sections. The total Hamiltonian will be

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{10}$$

Once the Green's function $\hat{G}^0$ of the unperturbed Hamiltonian is known, the one $\hat{G}$ of the perturbed Hamiltonian can be obtained via the Dyson equation,

$$\hat{G} = \hat{G}^0 + \hat{G}^0\hat{V}\hat{G}. \tag{11}$$

Let us briefly summarize the procedure used to solve this implicit equation.[29] We choose a mixed representation of the operators $\hat{G}$ and $\hat{V}$: on the space of lattice sites along the longitudinal direction and on the space of transverse eigenmodes for the transverse directions. This choice will result particularly convenient, as we shall see in the following.

In our representation, a generic operator $\hat{A}$ has matrix elements $\langle j,s|\hat{A}|t,k\rangle$, where j and k are transverse modes, and s and t are longitudinal lattice locations. For our purposes, we need the Green's functions and the perturbing coupling potential between pairs of lattice locations along the longitudinal direction. They can be represented with $M \times M$ matrices, M being the total number of transverse modes considered. We shall use the notation A_{st} to indicate such matrices.

Let us consider two isolated sections, one going from lattice location a to lattice location b, and the other from c to d. From Dyson's equation, we can write

$$\begin{aligned}\left\langle a|\hat{G}|d\right\rangle &= \left\langle a|\hat{G}^0|d\right\rangle + \left\langle a|\hat{G}^0\hat{V}\hat{G}|d\right\rangle \\ &= \left\langle a|\hat{G}^0|d\right\rangle \\ &\quad + \sum_{m,n}\left\langle a|\hat{G}^0|m\right\rangle\left\langle m|\hat{V}|n\right\rangle\left\langle n|\hat{G}|d\right\rangle. \end{aligned} \tag{12}$$

Since the perturbation $\hat{V}$ acts exclusively between the sites b and c, there are only two nonzero terms in the sum over m,n:

$$\begin{aligned}\left\langle a|\hat{G}^0\hat{V}\hat{G}|d\right\rangle &= \left\langle a|\hat{G}^0|b\right\rangle\left\langle b|\hat{V}|c\right\rangle\left\langle c|\hat{G}|d\right\rangle \\ &\quad + \left\langle a|\hat{G}^0|c\right\rangle\left\langle c|\hat{V}|b\right\rangle\left\langle b|\hat{G}|d\right\rangle, \end{aligned} \tag{13}$$

and, observing that $\left\langle a|\hat{G}^0|c\right\rangle = 0$, Eq. (12) can be written

$$\left\langle a|\hat{G}|d\right\rangle = \left\langle a|\hat{G}^0|d\right\rangle + \left\langle a|\hat{G}^0|b\right\rangle\left\langle b|\hat{V}|c\right\rangle\left\langle c|\hat{G}|d\right\rangle. \tag{14}$$

By using the previously defined matrix notation and extending the same line of reasoning to the lattice location pairs c,d and b,d, we obtain

$$G_{ad} = G^0_{ad} + G^0_{ab}V_{bc}G_{cd}\ , \tag{15}$$

$$G_{cd} = G^0_{cd} + G^0_{cc}V_{cb}G_{bd}\ , \tag{16}$$

$$G_{bd} = G^0_{bd} + G^0_{bb}V_{bc}G_{cd}. \tag{17}$$

Substituting Eq. (17) into Eq. (16) and Eq. (16) into Eq. (15), and noticing that $G^0_{ad} = G^0_{bd} = 0$, we get

$$G_{cd} = G^0_{cd} + G^0_{cc}V_{cb}G^0_{bb}V_{bc}G_{cd}, \tag{18}$$

from which we finally obtain the explicit expression

$$G_{cd} = (1 - G^0_{cc}V_{cb}G^0_{bb}V_{bc})^{-1}G^0_{cd}. \tag{19}$$

The matrix G_{aa} can be obtained with a completely analogous procedure.[29] Thus, starting from one end of the device, we recursively add one section at a time, considering its addition as a further perturbation, until we reach the other end and obtain the Green's functions across the whole structure.

The transmission coefficients can then be obtained from the Green's functions through straightforward relations.[29,30] If the transverse potential profile in each section is constant, there is no mode admixture within its length, and the related Green's-function matrix is diagonal. Each diagonal matrix element represents the Green's function for a single 1D tight-binding chain at the energy available for longitudinal motion in the relative subband, and can be derived with some simple algebra[29] both for the finite-length intermediate sections and the semi-infinite sections corresponding to the leads.

In other words, our initial 3D (or 2D, in the case of the idealized model) transport problem has been reduced down to a collection of 1D problems, with interaction between different modes only at the interfaces between different sections.

The elements of V_{st} are just the overlap sums between the eigenmodes ψ at location s and those at location t, multiplied by the coupling potential $v = -\hbar^2/(2m^*\delta^2)$,[29] where δ is the tight-binding lattice constant:

$$\langle j|V_{st}|k\rangle = v\sum_i \psi^{*s}_j(i)\psi^t_k(i), \qquad i = 1,Q\ , \tag{20}$$

Q being the total number of lattice sites on a transverse section.

Up to this point we have described the application to a 3D geometry of the standard recursive Green's-function method. We shall now examine the drawbacks of this method and the modifications we have introduced in order to overcome such problems. The tight binding energy–wave vector dispersion relation is cosinusoidal,

$$E = 2v[\cos(k\delta) - 1], \tag{21}$$

thus it represents a good approximation of the continuum parabolic relation only for values of the energy E that are very small compared to v. This introduces a constraint on the value of δ, the lattice spacing, which must be chosen small enough as to make the hopping potential v much larger than the maximum energy that we expect to consider in our simulation. A small lattice spacing leads, however, to a very large number of lattice sites in the transverse sections. In the recursive Green's-function technique of Ref. 29, the number of transverse modes considered corresponds to the number of lattice sites in the transverse sections, therefore, the size of the matrices to be inverted for the evaluation of expressions such as that in Eq. (19) becomes very large, whenever a wide energy range must be investigated.

In the case of the realistic model, this approach would become unfeasible, because the number of sites in a 2D cross section is extremely large even for coarse values of the lattice spacing. We notice that only a relatively

small number of transverse modes actually contributes to the determination of the transport parameters: those which exhibit a significant occupancy plus a few evanescent modes.

The observation that in the procedure summarized above the tight-binding formalism becomes important only in the numerical implementation of the connection between two neighboring sections has led us to the development of an approach in which the tight-binding lattice is preserved only in the longitudinal direction, while transverse sections are dealt with in an independent manner. In other words, we preserve the scheme of multiple 1D tight-binding chains (each of which corresponds to a transverse mode) interacting only at the interfaces between different transverse sections, but evaluate the transverse eigenvalues and the overlap integrals using a completely independent discretization.

Thus, Eqs. (10)–(19) remain formally unchanged, while Eq. (20) becomes

$$\langle j|V_{st}|k\rangle = v\int_S \psi_j^{*s}(\vec{r})\psi_k^t(\vec{r})d\vec{r}, \tag{22}$$

where S represents the surface of the cross section between the longitudinal sections s and t. The integral will actually be performed with an appropriate discretization, which is not related with the longitudinal tight-binding lattice. The longitudinal lattice spacing can now be independently chosen, small enough as to allow good precision over the required energy range, without influencing the overall computational complexity.

Formally, our modified method is rather similar to the scattering matrix approach, which has often been applied to the investigation of mesoscopic structures.[31] Implementation in a computer program is straightforward and computationally efficient.

For our realistic model, we have divided the structure into a number of transverse slices (Fig. 4) in each of which the potential is computed assuming a constant gate geometry corresponding to the one in the middle of the slice. As discussed earlier, the hypothesis of quasiadiabatic variation of the potential in the longitudinal direction is made, so that for each slice, we solve a 2D problem equivalent to that of infinite quantum wires defined by the transverse gate geometry of the slice.

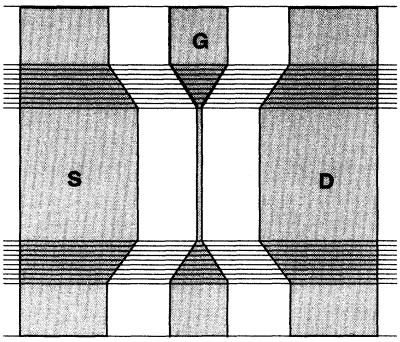

FIG. 4. Subdivision of the device into a number of transverse slices.

Far from the coupling region, the two waveguides are separate and can be represented with two independent semi-infinite leads. In the region where they approach each other, a certain number of slices is needed, capable of resolving the gate geometry, and, finally, the coupling region, with constant gate geometry, is represented with a single slice. We have experimented with the number of slices and observed that no significant improvement was reached with more than a total of 21 sections, including the semi-infinite input and output leads and the coupling region. Due to the longitudinal symmetry of the device, we need to solve for the self-consistent potential only in 11 slices: one containing the semi-infinite leads, nine for the transition region, and one for the coupling region. We have increased the number of transverse modes being included in our calculation until no significant difference could be observed: a total of 18 modes have been used for the two waveguides, with a number of modes with significant occupancy varying between 4 and 12.

IV. NUMERICAL RESULTS

We have run simulations for the realistic model only for a few important cases, since each data point requires significant CPU time on an HP735 workstation. The calculation is very time consuming, because of the time needed to reach convergence of the self-consistent Schrödinger-Poisson solver for each of the 11 slices at low temperatures.

In the first two simulations, we have held the gate and the drain voltages constant, sweeping the source voltage between -1.99 V and -1.4 V, at a temperature of 1.6 K. The gate voltage has been set at -0.6 V, corresponding to the 2D threshold. We have made this choice in order to obtain some significant coupling between the two waveguides, which would otherwise be vanishing for lower gate voltages.

In our model (Fig. 2), the source and drain gates are within a short distance from the middle gate everywhere along the longitudinal axis, therefore, they contribute to the depletion under such gate, preventing the formation of a 2D electron gas (2DEG) even under its widest sections. In the real device[8] the source and drain gates are very far apart outside the central region, and there may not be enough depletion underneath the middle gate, if biased right at the 2D threshold, to keep the two 2DEG's insulated from each other. This could be the reason why del Alamo and Eugster were not able to observe significant 1D-to-1D tunneling in their experiments[32] on this heterostructure: with a large enough negative bias voltage on the middle gate (to keep the reservoirs insulated from each other) coupling between the two waveguides becomes too small to be observed, as we shall see later in this section.

In Fig. 5 we report, for a drain bias of -1.72 V, the conductance of the drain waveguide (dashed line), of the source waveguide (dotted line) and the tunneling conductance between the two waveguides (solid line) for a coupling length of 450 nm (a) and of 400 nm (b). All conductances are expressed in units of $2e^2/h$. The drain waveguide is defined by the drain gate and the middle

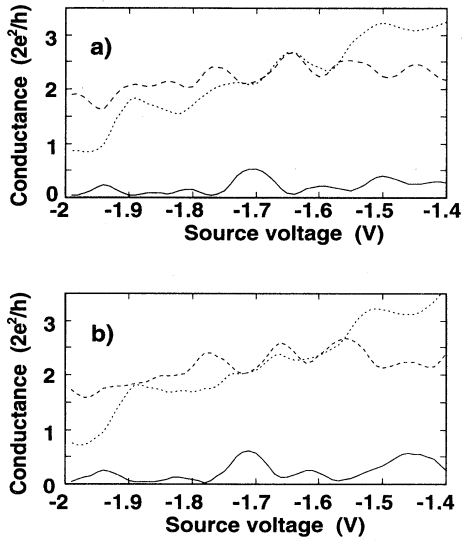

FIG. 5. Tunneling conductance between the two waveguides (solid line), drain waveguide conductance (dashed line), and source waveguide conductance (dotted line), as a function of source voltage for a coupling length of 450 nm (a) and of 400 nm (b). The drain voltage is held constant at −1.72 V and the temperature is 1.6 K.

gate, which are held at constant voltages, therefore its conductance is somewhat constant, oscillating between 1.5 and 2.5 units. The conductance of the source waveguide, instead, increases with increasing source voltage, varying between 1 and 3.5 units. Two conductance steps are visible, partially smeared by the effects of coupling to the drain waveguide and of the bend geometry. The tunneling conductance is characterized by clear peaks that correspond to the alignment of the highest subband levels in the two waveguides. The central peak occurs, as expected, at a source voltage of −1.72 V, corresponding to perfect symmetry with the drain waveguide. In this case, all levels of both waveguides are aligned, but the main contribution still comes from tunneling between the topmost occupied subbands. Structure between the peaks

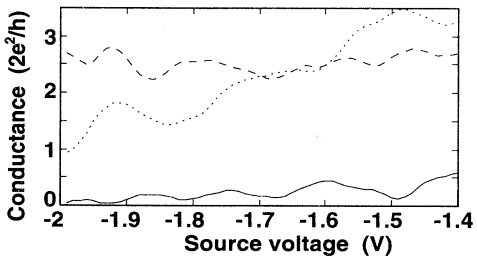

FIG. 6. Tunneling conductance between the two waveguides (solid line), drain waveguide conductance (dashed line), and source waveguide conductance (dotted line), as a function of source voltage for a coupling length of 450 nm. The drain voltage is held constant at −1.6 V and the temperature is 1.6 K.

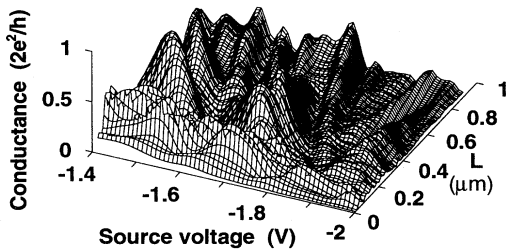

FIG. 7. Tunneling conductance between the drain and source waveguides vs source voltage and coupling length. The drain voltage is held constant at −1.72 V and the temperature is 1.6 K.

is related to tunneling between lower subbands. Notice also that dips in the drain and source waveguide conductances appear in correspondence with the peaks in the tunneling current, a direct consequence of total current conservation.

This same calculation has been repeated for a different drain bias, −1.6 V, instead of −1.72 V, for a coupling length of 450 nm. The results are shown in Fig. 6, and exhibit the expected shift in the central peak, which moves up to approximately $V_S = -1.6$ V, the bias value leading to a symmetric transverse potential landscape. The other peaks are correspondingly shifted towards higher V_S values. The meaning of the various curves is the same as in Fig. 5.

In Fig. 7, we show some more results for a drain bias of −1.72 V, varying both the source bias from −1.99 V to −1.4 V and the coupling length from 0 to 900 nm. A rich structure can be observed, originating from subband alignments, periodic oscillation of the electrons between pairs of quasidegenerate states and multiple reflections associated with the particular geometry of the wires. The oscillation of electrons between quasidegenerate states leads to fluctuations of the tunneling current as a function of the coupling length and can be qualitatively understood in the framework of coupled modes theory.[2]

In Fig. 8, we report the computed tunneling current versus coupling length for three values of the central gate bias: −0.6 V (solid line), −0.62 V (dashed line), and −0.64 V (dotted line). We notice that the period of the

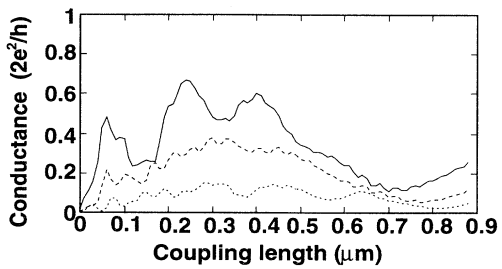

FIG. 8. Tunneling conductance vs coupling length for $V_G = -0.6$ V (solid line), $V_G = -0.62$ V (dashed line), and $V_G = -0.64$ V (dotted line). The other parameters are $V_S = V_D = -1.72$ V and $T = 1.6$ K.

main, lowest frequency, fluctuation tends to increase with decreasing gate voltage.

This can be easily understood from coupled mode theory. When the potential barrier between the two waveguides is very high (very low gate voltage), so that they can be considered independent, each waveguide has its own set of localized transverse eigenstates. By properly biasing the source and drain electrodes, it is possible to align the highest occupied eigenvalues of the two wires: if we raise the gate voltage, the barrier between the two waveguides will be lowered, the transverse states will not be localized in a single waveguide any more, and their energy eigenvalues will split by an amount increasing with increasing gate voltage. Let us define these energy eigenvalues E_1 and E_2. The wave vectors for longitudinal propagation along the two modes will be given by

$$k_i = \sqrt{\frac{2m^*}{\hbar^2}(E - E_i)}, \qquad i = 1, 2 \ , \tag{23}$$

where E is the total energy of the electron. The probability density in each waveguide will depend on the interference between the longitudinal waves associated with these k_i's, and maxima of the tunneling current will be measured when the length of the coupling section will correspond to $\pi/(k_1 - k_2)$.[2] Thus, decreased coupling leads to a reduced difference between k_1 and k_2, and, therefore, to a longer oscillation period in Fig. 8. Coupling between the two waveguides becomes rapidly negligible when the gate voltage is lowered below -0.65 V, as shown in Fig. 9, where the tunneling conductance is plotted versus coupling length and gate voltage.

We have also performed a calculation for higher temperature, 10 K. Results are reported in Fig. 10 for $V_S = -1.72$ V and coupling length $L = 450$ nm (a) and $L = 400$ nm (b). The overall behavior has not undergone significant changes with respect to the results at 1.6 K, because thermal broadening at 10 K is still much smaller than the width of the peaks in the tunneling current. The main effect of higher temperature is the one of smoothing out the results as a consequence of the broader $\partial f/\partial E$ weighing function. This is clearly shown by the data reported in Fig. 11, which represent the equivalent, at $T = 10$ K, of those of Fig. 7.

These calculations at 10 K are only for exemplification purposes, because at such temperature phase-breaking dissipative effects, which are not included in our simulation, would play a dominant role, destroying most of the interference features and washing out conductance quantization.

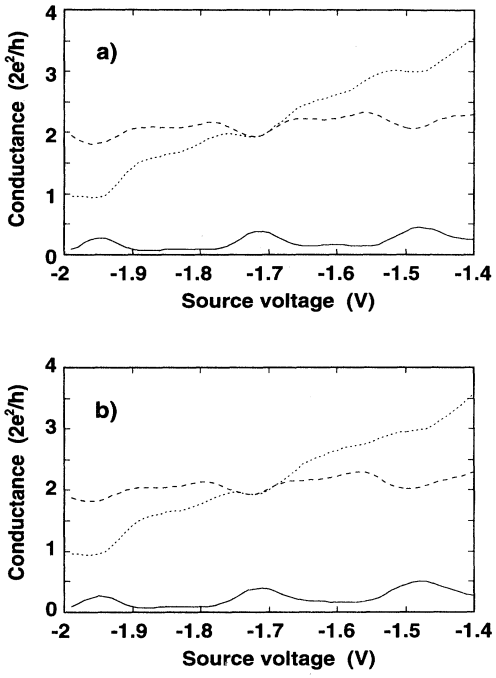

FIG. 10. Tunneling conductance between the two waveguides (solid line), drain waveguide conductance (dashed line), and source waveguide conductance (dotted line), as a function of source voltage for a coupling length of 450 nm (a) and of 400 nm (b). The drain voltage is held constant at -1.72 V and the temperature is 10 K.

The main conclusion we can draw from the analysis of the results obtained from our device model is that with this heterostructure 1D-to-1D tunneling is difficult to achieve, mostly because of the very small coupling between the two waveguides, except for gate bias values too close to the pinch-off voltage for the 2DEG. Another relevant point is represented by the drain-bias dependent shift in the source bias values corresponding to the con-

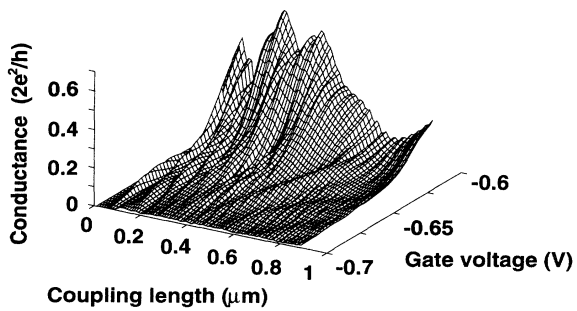

FIG. 9. Tunneling conductance vs coupling length and gate voltage. The other parameters are $V_S = V_D = 1.72$ V and $T = 1.6$ K.

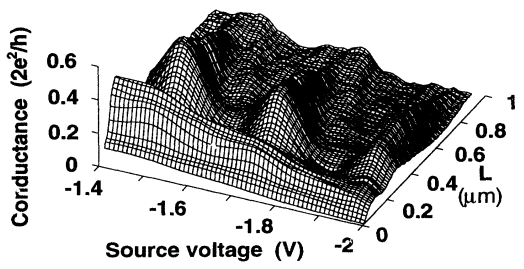

FIG. 11. Tunneling conductance between the drain and source waveguides vs source voltage and coupling length. The drain voltage is held constant at -1.72 V and the temperature is 10 K.

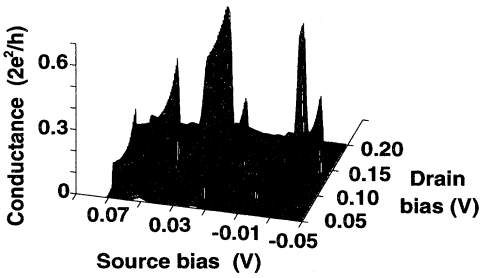

FIG. 12. Tunneling conductance for the idealized model vs source and drain voltage at 1.6 K. The drain and source waveguides are 30 nm wide and are separated by a 0.4 eV barrier in the 450 nm long coupling region

ductance peaks, which does not appear clearly in the experimental data from Ref. 11. Even though this experiment has been performed with a somewhat different heterostructure, the behavior of the conductance bumps for 1D-to-1D tunneling should be in qualitative agreement with our simulation. In order to gain a better understanding of these features, in particular, of the conductance patterns expected when both the source and drain voltage are varied (as in the experiment of Ref. 11), we have performed some calculations on the idealized model described in Sec. II.

The results for the simulation of two quantum wires 30 nm wide coupled for a length of 450 nm through a barrier with a height of 0.4 eV and a width of 5 nm are shown in Fig. 12. The potential of the left edge of the source channel is swept between -0.05 and 0.07 V with respect to the Fermi level, while the potential at the right end of the drain (see Sec. II) is swept between 0.05 and 0.2 V. The tunneling conductance exhibits relatively narrow ridges approximately parallel to the $V_S = V_D$ direction. The reason for this alignment can be quickly understood if we think of the symmetric bias condition. As previously discussed with reference to the realistic model, symmetric bias implies the alignment condition for all of the pairs of transverse eigenvalues in the two waveguides, and, in particular, for the eigenvalues closest to the Fermi level. Such alignment is maintained if we vary both V_S and V_D at the same time, giving rise to the leftmost ridge visible in Fig. 12. The other ridges, corresponding to different alignments, are exactly parallel to the same direction, because in our idealized model varying the bias values while keeping $V_S - V_D$ constant is equivalent to sliding the Fermi level through a constant potential landscape: when the Fermi level goes through the pair of aligned eigenvalues, they start contributing to the tunneling current; then as the Fermi level gets further up, their contribution slowly decreases until it vanishes.

The main features of these results are better visible in a gray-scale plot [Fig. 13(a)], where black represents conductance 0 and white conductance 1. The gray scale is logarithmic, therefore very small details become visible, such as the "rays" radiating from the end of each ridge. They are the result of multiple reflections between the ends of the coupling region, and, thus, their spacing depends on such a length. This can be verified by comparing Fig. 13(a) with Fig. 13(b), where the results for a similar calculation are shown. All the parameters are the same as in the case of Fig 13(a), except for the coupling length, which is now 650 nm instead of 450 nm. The spacing between the "rays" becomes clearly smaller, as a consequence of the reduced wave vector associated with multiple reflections within the new, longer, coupling section.

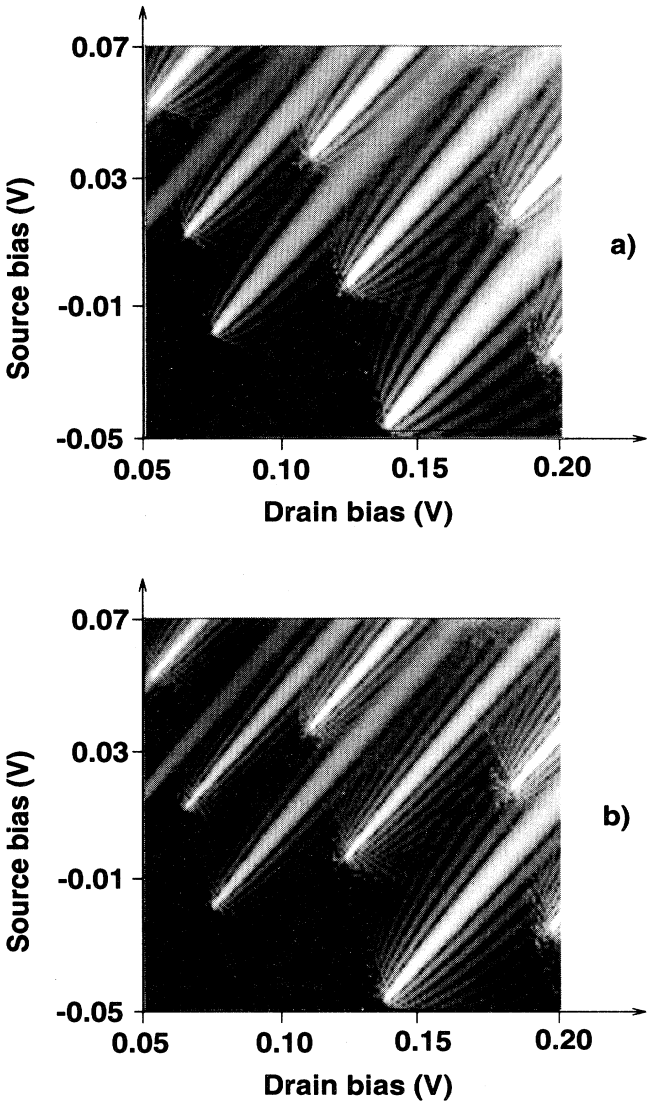

FIG. 13. Gray-scale plot of the tunneling conductance for the idealized model vs source and drain voltage at 1.6 K. The drain and source waveguides are 30 nm wide and are separated by a 0.4 eV barrier in the 450 nm (a) or 650 nm (b) coupling region.

The results for a calculation of the tunneling conductance on a wider, symmetric range of drain and source bias values are reported in Fig. 14, where V_S and V_D vary in the range $0 - 0.4$ V and the coupling section is

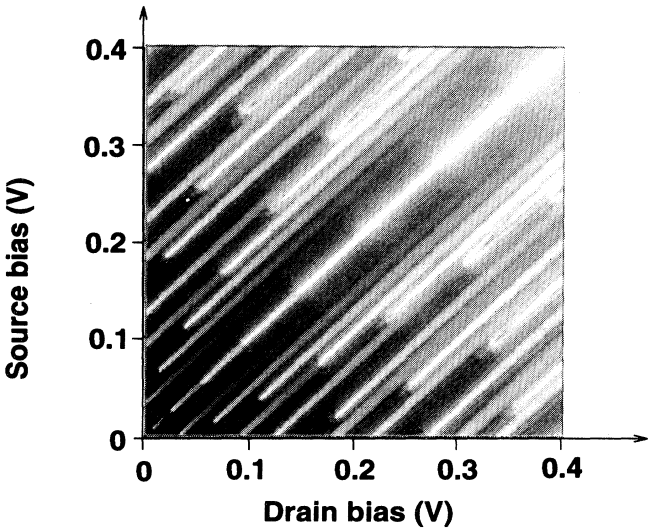

FIG. 14. Gray-scale plot of the tunneling conductance for the idealized model vs source and drain voltage at 1.6 K. The drain and source waveguides are 30 nm wide and are separated by a 0.4 eV barrier in the 650 nm long coupling region.

650 nm long. We notice that, moving towards higher bias values along $V_S = V_D$, the conductance goes suddenly up when the Fermi level goes through the first pair of aligned eigenvalues, then falls down and rises again whenever a new pair of eigenvalues is crossed. For all the other, nonsymmetric bias conditions, ridges have only one maximum, followed by a long decay, because only one pair of eigenvalues satisfies the alignment condition.

From these results, we deduce that a sharp difference in the orientation of the ridges is expected between the 1D-to-2D and the 1D-to-1D tunneling cases. In the 1D-to-2D case, maxima occur for values of the voltage applied to the gate defining the 1D channel, which correspond to the addition of each new subband. These values are independent of the bias applied to the other gate, as long as this is not low enough as to create 1D confinement. On the $V_S - V_D$ plane, conductance ridges will be parallel to the V_D or V_S axis, depending on whether conduction in the source or drain channel, respectively, is quantized. For 1D-to-1D tunneling, instead, the ridges will be approximately aligned along the $V_S = V_D$ direction, as previously discussed.

A similar alignment of ridges is expected to occur in the case of the realistic model, for which we have not performed a complete bidimensional scan, because of the excessive CPU time required. The data shown in Figs. 5,6 are consistent with such ridge pattern. The only qualitative difference we expect with respect to the ideal model is represented by some possible curvature of the ridges, due to the more complex dependence of the eigenvalues on the bias voltages. The overall pattern of ridges running along a diagonal direction on the $V_S - V_D$ plane should be the distinctive signature of 1D-to-1D tunneling. Therefore, it is possible that the conductance pattern observed in the plots of Ref. 11 is due to remnants of 1D-to-2D tunneling, rather than actual 1D-to-1D tunneling.

V. CONCLUSIONS

We have developed a method for the quasi-three-dimensional simulation of quantum waveguide couplers, based on the self-consistent calculation of the transverse confining potential and on a modified recursive Green's-function algorithm for the evaluation of the transport properties and, in particular, of the conductances.

Tunneling between two electron waveguides has been investigated, focusing on the device fabricated by Eugster and del Alamo.[8] Our results show that coupling between the two waveguides in this structure is rather small, and that a bias very close to the 2D threshold is needed for the central gate in order to observe significant tunneling. This explains why 1D-to-1D tunneling has not been experimentally observed in the device based on this particular heterostructure: the central gate cannot be biased too close to the 2D threshold, otherwise far from the coupling region insulation between the 2DEG reservoirs would deteriorate. From the data obtained both for the quasi-three-dimensional model and for the idealized 2D model, for which a more extensive exploration of the parameter space has been possible, we deduce the characteristic signature that should be revealing of the presence of 1D-to-1D tunneling and does not appear in currently available experimental results:[11] a pattern of ridges approximately parallel to the $V_D = V_S$ direction in the plot of the tunneling conductance versus V_D and V_S. Our simulation method can be a useful tool for the design of future implementations of this promising quantum effect device.

ACKNOWLEDGMENTS

We thank Jesus del Alamo for valuable discussion on the experimental implementation of quantum waveguide couplers. We also thank Thomas Kerkhoven for helpful discussions on efficient numerical techniques. This work has been supported by the NSF through Grant No. ECS 91-20641, by the National Research Council (CNR) of Italy and by the Italian Ministry of University and Scientific and Technological Research.

[1] J. A. del Alamo and C. C. Eugster, Appl. Phys. Lett. **56**, 78 (1989).
[2] R. Q. Yang and J. M. Xu, Phys. Rev. B **43**, 1699 (1991).
[3] L. Jiang, G. Xu, P. Jiang, D. Lu, and X. Xie, J. Phys. Condens. Matter **6**, 5957 (1994).
[4] G. Xu and P. Jiang, Semicond. Sci. Technol. **7**, 612 (1992).
[5] M. Macucci, U. Ravaioli, and T. Kerkhoven, Superlatt. Microstruct. **12**, 509 (1992).

[6] G. Xu, M. Yang, and P. Jiang, J. Appl. Phys. **74**, 6747 (1993).
[7] J. Wang, Y. Jiang, and H. Guo, Phys. Rev. B **46**, 2420 (1992).
[8] C. C. Eugster and J. A. del Alamo, Phys. Rev. Lett. **67**, 3586 (1991).
[9] C. C. Eugster, J. A. del Alamo, M. J. Rooks, and M. R. Melloch, Appl. Phys. Lett. **60**, 642 (1992).
[10] C. C. Eugster, J. A. del Alamo, M. R. Melloch, and M. J. Rooks, Phys. Rev. B **48**, 15 057 (1993).
[11] C. C. Eugster, J. A. del Alamo, M. J. Rooks, and M. R. Melloch, Appl. Phys. Lett. **64**, 3157 (1994).
[12] R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).
[13] M. Büttiker, IBM J. Res. Dev. **32**, 306 (1988).
[14] B. J. van Wees, L. P. Kouwenhoven, E. M. M. Willems, C. J. P. M. Harmans, J. E. Mooji, H. van Houten, C. W. J. Beenakker, J. G. Williamson, and C. T. Foxon, Phys. Rev. B **43**, 12 431 (1991).
[15] T. Kerkhoven, A. Galick, U. Ravaioli, J. H. Arends, and Y. Saad, J. Appl. Phys. **68**, 3461 (1990).
[16] S. E. Laux and F. Stern, Appl. Phys. Lett. **49**, 91 (1986).
[17] L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2046 (1971).
[18] F. Stern and S. Das Sarma, Phys. Rev. B **30**, 840 (1984).
[19] H. M. Antia, Astrophys. J. Suppl. Ser. **84**, 101 (1993).
[20] R. S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962), p. 184.
[21] Y. Saad, in *Matrix Pencils, Proceedings*, edited by P. Havsbad, B. Kågström, and A. Ruhe, Springer Lecture Notes in Mathematics Vol. 973 (Springer-Verlag, Berlin, 1983), p. 121.
[22] A. T. Galick, T. Kerkhoven, and U. Ravaioli, IEEE Trans. Microwave Theory Tech. **40**, 699 (1992).
[23] H. Rutishauser, Numer. Math. **16**, 205 (1970).
[24] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981).
[25] Y. Saad and M. H. Schultz, SIAM J. Sci. Stat. Comput. **7**, 856 (1986).
[26] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1989).
[27] D. J. Thouless and S. Kirkpatrick, J. Phys. C **14**, 235 (1981).
[28] F. Guinea and J. A. Vergés, Phys. Rev. B **35**, 979 (1987).
[29] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, J. Appl. Phys. **66**, 3892 (1989).
[30] A. D. Stone and A. Szafer, IBM J. Res. Dev. **32**, 384 (1988).
[31] R. L. Schult, H. W. Wyld, and D. G. Ravenhall, Phys. Rev. B **41**, 12 760 (1990).
[32] J. A. del Alamo and C. C. Eugster (private communication).

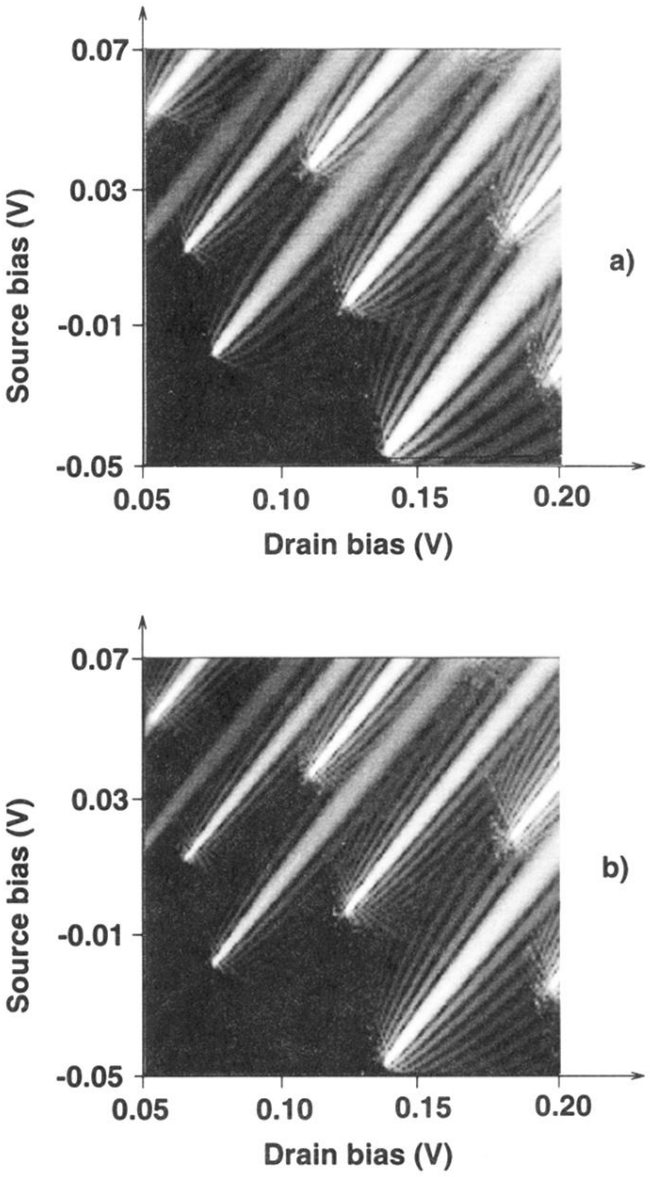

FIG. 13. Gray-scale plot of the tunneling conductance for the idealized model vs source and drain voltage at 1.6 K. The drain and source waveguides are 30 nm wide and are separated by a 0.4 eV barrier in the 450 nm (a) or 650 nm (b) coupling region.

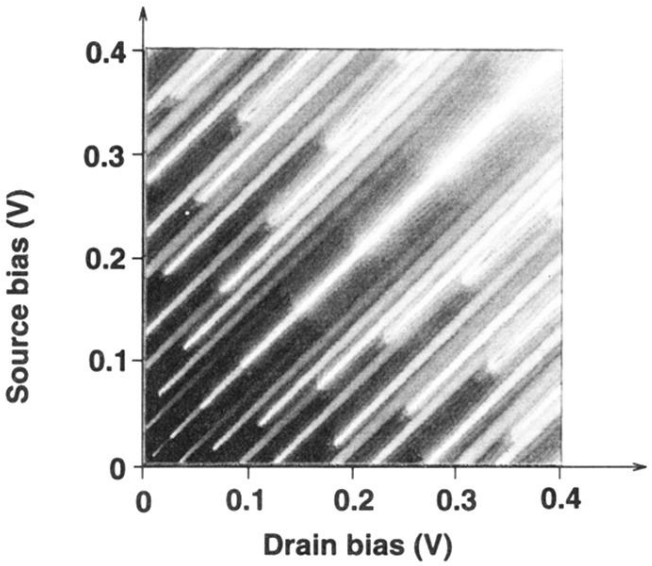

FIG. 14. Gray-scale plot of the tunneling conductance for the idealized model vs source and drain voltage at 1.6 K. The drain and source waveguides are 30 nm wide and are separated by a 0.4 eV barrier in the 650 nm long coupling region.

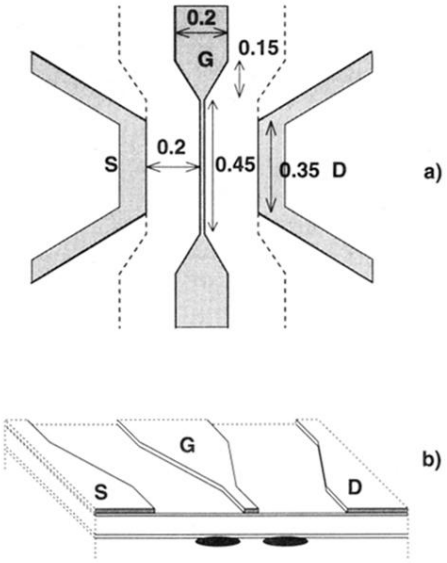

FIG. 2. (a) Gate layout for the actual device (shaded areas) and modified layout used in the simulations. (b) View of a cross section of the device: the dark-shaded regions correspond to the quantum wires.

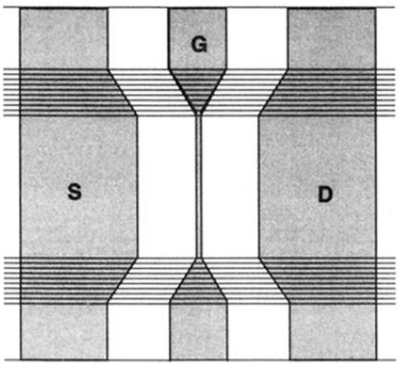

FIG. 4. Subdivision of the device into a number of transverse slices.