# Self-Consistent Modeling of Open Quantum Devices

Richard Akis, Lucian Shifren and David Ferry

Department of Electrical Engineering, Arizona State University
Tempe, AZ 85287-5706, USA
richard.akis@asu.edu

## ABSTRACT

In this paper, we describe a method of simulating electron transport in semiconductor devices that operate in the quantum regime. Specifically, devices formed in which the electrons are confined to two dimensions (2D) and transport is ballistic. Modeling such structures using a finite difference approach, we describe how the conductance can be calculated using a numerically stabilized variant of the transfer matrix approach derived from the 2D Schrödinger equation. Examining the example of a quantum point contact, we also describe how this method can be efficiently coupled to a Poisson solver to allow self-consistency to be achieved.

***Keywords***: quantum transport, heterostructures, numerical methods, finite difference.

## 1. INTRODUCTION

The Semiconductor Industry Association (*SIA*) projects that by the end of 2009 devices will employ 0.05 μm gate lengths and have oxide thickness of less than 1.5 nm. Groups from Toshiba and Lucent Bell Labs have already fabricated *n*-channel MOSFETs with effective gate lengths below 25 nm [1], showing that these feature sizes are feasible. In this sub-micron regime, electron transport is expected to be dominated by quantum effects. In MOSFETs and GaAs/GaAlAs heterostructures, the situation is commonly one in which the electrons are already confined along one of the dimensions, forming a subband structure[2]. If the electrons are free to move along the two remaining directions, they constitute a 2 dimensional electron gas (2DEG). Moreover, if the dimensions of the device are significantly smaller than the mean free path and the coherence length, then transport should be ballistic, ie. the primary source of scattering is from the boundaries[2]. Under such conditions, interference effects due to the wave-like nature of the electrons become paramount and a fully quantum mechanical treatment is required for an accurate simulation of such a device. In this paper, we describe a numerical method for doing such calculations based on a stabilized variation of the transfer-matrix method. We also show this 2D Schrödinger solver can be efficiently coupled to a Poisson solver to allow self-consistency to be achieved.

## 2. CALCULATING THE CONDUCTANCE

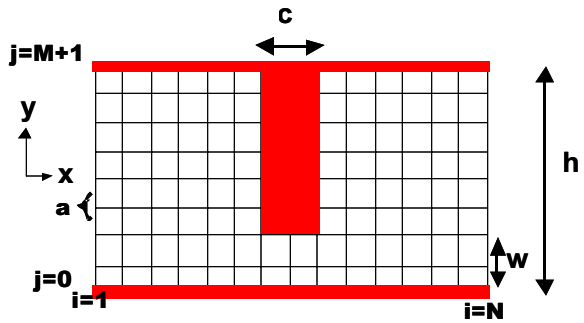



Figure 1: Schematic for the quantum device we consider in this paper, a quantum point contact. The dark regions represent places where the potential, $V_{i,j}$, is high. Importantly, the lattice spacing, $a$, is in reality much smaller than it is illustrated here.

To illustrate the method, we simulate one of the simplest quantum devices, a quantum point contact (QPC). As shown schematically in Fig. 1, our QPC is constituted by a narrowing or bottleneck occurring inside a quantum wire. We assume that this device is constructed within a GaAs/GaAlAs heterostructure, so that the effective mass is 0.067 and the electrons are confined in a 2DEG, so that the z-direction need not be considered.

We perform our simulations on a square finite-difference lattice with lattice constant, $a$. Position can thus be specified by integers: $x=ia$ and $y=ja$. Keeping only the lowest order terms in the approximations of the derivatives, the 2D Schrödinger equation becomes

$$-t(\psi_{i+1,j}+\psi_{i-1,j}+\psi_{i,j+1}+\psi_{i,j-1})+(V_{i,j}+4t)\psi_{i,j} \qquad (1)$$
$$=E\psi_{i,j}, \text{ where } t=\frac{\hbar^2}{2m^*a^2},$$

$V_{i,j}$ represents the potential at site $i,j$ and $E$ is the energy. Since we are interested in current flow, the typical situation which we consider is one in which the device is enclosed inside an ideal quantum wire, which extends outward to $\pm\infty$ along the $x$-axis. Along the top and bottom boundaries we use Dirichlet boundary conditions, so for a wire M

lattice spacings high,

$$\psi_{i,j=0} = \psi_{i,j=M+1} = 0. \qquad (2)$$

Given this, the wavefunction along a particular slice $i$ on the $x$-axis can be specified by a M-dimensional vector and equation (1) can be rewritten as a matrix equation:

$$\mathbf{H_{0i}}\vec{\psi}_i - t\,\vec{\psi}_{i+1} - t\,\vec{\psi}_{i-1} = E\,\vec{\psi}_i . \qquad (3)$$

Combining this with the trivial equation, $\vec{\psi}_i = \vec{\psi}_i$ , one can derive a transfer-matrix equation that relates adjacent slices:

$$\begin{bmatrix} \vec{\psi}_i \\ \vec{\psi}_{i+1} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & \left(\dfrac{\mathbf{H_{0i}} - E}{t}\right) \end{bmatrix} \begin{bmatrix} \vec{\psi}_{i-1} \\ \vec{\psi}_i \end{bmatrix} = \mathbf{T_i} \begin{bmatrix} \vec{\psi}_{i-1} \\ \vec{\psi}_i \end{bmatrix} . \qquad (4)$$

Since the quantum wire acts as a waveguide, the actual current is carried by the propagating modes of the wire. Thus, we begin the calculation by solving the eigenvalue problem for the transfer-matrix on the first slice:

$$\mathbf{T}_1 \begin{bmatrix} \vec{\psi}_1 \\ \vec{\psi}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{12} & \mathbf{T}_{22} \end{bmatrix} \begin{bmatrix} \vec{\psi}_1 \\ \vec{\psi}_0 \end{bmatrix} = \lambda \begin{bmatrix} \vec{\psi}_1 \\ \vec{\psi}_0 \end{bmatrix}. \qquad (5)$$

The eigenvectors of (5) have the form

$$\begin{bmatrix} \vec{u}_m(\pm) \\ \lambda_m(\pm)\vec{u}_m(\pm) \end{bmatrix}, \qquad (6)$$

and given there are $q$ propagating wave modes ($|\lambda| = 1$) and $M$-$q$ evanescent modes ($|\lambda| \neq 1$) the corresponding eigenvalues can be expressed as

$$\lambda_m(\pm) = e^{\pm i k_m a},\ m = 1, \ldots, q \qquad . \qquad (7)$$

$$\lambda_m(\pm) = e^{\mp \kappa_m a},\ m = q+1 \ldots M$$

The ± symbol refers to the fact that the modes actually come in pairs, those that travel to the right (+) and those to the left (−). For the transmission problem, it is useful to collect these together in a 2M × 2M matrix

$$\mathbf{T}_0 = \begin{bmatrix} \mathbf{U}_+ & \mathbf{U}_- \\ \mathbf{U}_+ \boldsymbol{\Lambda}_+ & \mathbf{U}_- \boldsymbol{\Lambda}_- \end{bmatrix}, \text{where}$$

$$\mathbf{U}_\pm = [\vec{u}_1(\pm) \ \ldots \ \vec{u}_M(\pm)], \text{ and} \qquad (8)$$

$$\boldsymbol{\Lambda}_\pm = diag[\lambda_1(\pm) \ \ldots \ \lambda_M(\pm)]$$

To calculate the transmission through a device, the modes are injected from the left side only with unit amplitude. For a structure N slices long, one must thus solve the transfer matrix problem:

$$\begin{bmatrix} \mathbf{t} \\ \mathbf{0} \end{bmatrix} = \mathbf{T}_0^{-1}\mathbf{T}_N\mathbf{T}_{N-1}\ldots\mathbf{T}_1\mathbf{T}_0 \begin{bmatrix} \mathbf{I} \\ \mathbf{r} \end{bmatrix} \qquad , \qquad (9)$$

where $\mathbf{t}$ is a matrix of transmission amplitudes of waves exiting from the right part of the structure, and $\mathbf{r}$ is the matrix of amplitudes of waves reflected back towards the left. Given $\mathbf{t}$, one calculate the conductance, $G$, using the Landauer-Buttiker formula[2]:

$$G = \frac{2e^2}{h} \sum_{m,n} \frac{v_n}{v_m} \left| t_{n,m} \right|^2 \qquad (10)$$

where $t_{n,m}$ represents the transmission amplitude of mode $n$ to mode $m$ and the summation is only over propagating modes. The $v$'s are the mode velocities, which can be obtained by taking the expectation value of the current operator.

Unfortunately, equation (9) in its current form is made numerically unstable by the exponentially growing and decaying contributions of the evanescent modes that accumulate when the product of transfer matrices is taken. Usuki *et al.*[3] overcame this difficulty be rewriting the transfer matrix problem in terms of an iterative scheme. Rather than using the simple relationship given by equation (4), slices $i$ and $i+1$ can be related by:

$$\begin{bmatrix} \mathbf{C}_1^{i+1} & \mathbf{C}_2^{i+1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \mathbf{T}_i \begin{bmatrix} \mathbf{C}_1^{i} & \mathbf{C}_2^{i} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{P}_i$$

$$\text{with} \quad \mathbf{P}_i = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{P}_{i1} & \mathbf{P}_{i2} \end{bmatrix}, \qquad (11)$$

$$\mathbf{P}_{i1} = -\mathbf{P}_{i2}\mathbf{T}_{i21}\mathbf{C}_1^i,$$

$$\text{and} \quad \mathbf{P}_{i2} = [\mathbf{T}_{i21}\mathbf{C}_2^i + \mathbf{T}_{i22}]^{-1}.$$

The iteration is started by the condition $\mathbf{C}^0{}_1 = \mathbf{I}$ and $\mathbf{C}^0{}_2 = \mathbf{0}$, which starts the modes off with unit amplitude. At the end, $\mathbf{t}$ obeys the relationship:

$$= -\left(\mathbf{U}^+\ddot{\mathbf{e}}^+\right)^{-1}\left[\mathbf{C}_1^{N+1} - \mathbf{U}^+\left(\mathbf{U}^+\ddot{\mathbf{e}}^+\right)^{-1}\right] \qquad . \qquad (12)$$

The numerical stability of the Usuki *et al.* method in large part stems from the fact that the iteration implied by equation (11) involves products of matrices with *inverted* matrices. Taking such products tends to cancel out most of the troublesome exponential factors. It should be noted that the Usuki *et al.* method is really just a variation on the "cascading scattering matrix" method developed by Inkson and Ko [4].

## 3. ACHIEVING SELF-CONSISTENCY

To make the calculation self-consistent, we need to be able to calculate the electron density, which is obtained by reconstructing the electron wave functions. Usuki *et al*. outlined a method for doing this starting from the left and working back to the end of the structure. Unfortunately, it entails performing a calculation similar to that for obtaining the conductance, but for every single slice. As a result, while the time it takes to calculate $G$ goes as $N$, the time to reconstruct the wave function instead goes as $N!$, which makes it very time consuming. Since self-consistency requires the density to be recalculated numerous times, including it thus becomes impractical for most calculations. We however have found a simple way to make the reconstruction far more efficient. Instead of going from left to right, one starts at the end of the structure and works backword. Manipulating Usuki *et al.*'s equations, it can be shown that for the final slice :

$$\boldsymbol{\psi}_N = \mathbf{P}_{N2}. \qquad (13)$$

Note here that $\boldsymbol{\psi}_N$ is a matrix, the columns of which represent the separate contributions of the individual modes to the total wave function on slice $N$. Going towards the left, one then does the iteration:

$$\boldsymbol{\psi}_i = \mathbf{P}_{i1} + \mathbf{P}_{i2}\boldsymbol{\psi}_{i+1}. \qquad (14)$$

The **P**'s here are the same ones obtained during the $G$ calculation and so are recalled from memory rather than recalculated. The density at site $i,j$, given there are $q$ propagating modes becomes

$$n(x,y) = n(i,j) = \sum_{k=1}^{q} \left|\psi_{ijk}\right|^2. \qquad (15)$$

Obtaining $n(x,y)$ in this modified way takes about the same amount of time as the original $G$ calculation and can be *orders of magnitude faster* than Usuki *et al.*'s technique depending on the size of the structure.

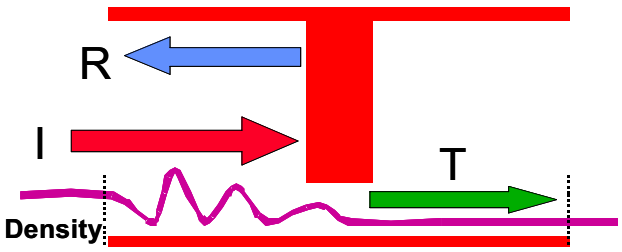



Figure 2: Illustration of the situation simulated- waves incident from the left (I) that are reflected (R) and transmitted (T). The dashed lines are where the boundary conditions of equation (16) are imposed.

Before an actual self-consistent calculation, one must put appropriate boundary conditions on $n(x,y)$. For this, we assume Neuman boundary conditions on the ends:

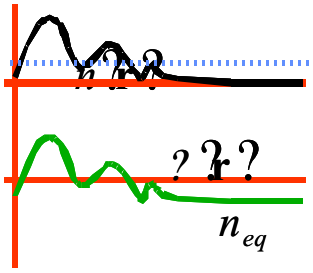
$$\left[\frac{dn}{dx}\right]_{i=0} = \left[\frac{dn}{dx}\right]_{i=N} = 0. \qquad (16)$$

This assumption holds if the region of interest, ie. the device, is far enough from the ends that the reflected waves do not interfere with the boundaries. While only strictly true when the transmitted and reflected waves are subjected to damping, this assumption is sufficiently accurate in general to allow a convergent solution to Poisson's equation, and has been used previously in calculations performed in similar quantum wire [5]. To stabilize the calculation during the iteration to solve Poisson's equation, we follow previous work [5-7] and fix one dimensional density at the entrance and exit slices, and allow the Fermi energy to adjust itself in order to keep the number of propagating modes constant. To determine the new adjusted Fermi level, we solve [5-7] :

$$N_{1D} = \frac{2}{\pi}\sum_{E_l<E_F}\left[\frac{2m^*(E_F - E_l)}{\hbar^2}\right]^{\frac{1}{2}}, \qquad (17)$$

where $N_{1D}$ is the 1D density at the quantum wire entrance to the device and $E_l$ is the energy level of the $l^{\text{th}}$ mode as calculated from the Schrödinger equation.

Figure 3: A cross-section along the x-direction of the electron density n(***r***), the ionized impurity density $n_{eq}$, and the net density $\rho(\mathbf{r})$.




The *total* self-consistent potential is a combination of the Hartree, exchange and the correlation potentials. The formulation for the exchange and the correlation potential can be found in Ref. [5]. The Hartree potential is calculated via an iterative Poisson solver (for programming simplicity, we used simultaneous over-relaxation, though more sophisticated solving methods could have been employed). The Poisson solver solves the usual equation

$$\nabla^2\phi(\mathbf{r}) = -\frac{e}{\varepsilon}\rho(\mathbf{r}) \qquad (18)$$

where $\rho$, the density, is given by

$$\rho(\mathbf{r}) = n_{eq} - n(x,y). \qquad (19)$$

Here, $n_{eq}$ is the density of positively ionized donors in the structure, which is taken to be equal in magnitude to the average electron density of the entire structure so that the system is charge neutral. As shown in Fig. 3, what ρ ends up representing are the fluctuations in the density. The Poisson equation is solved after every wave function density calculation until we reach two levels of convergence, namely, the Poisson and the Schrödinger solutions both need to converge for the system to have full self-consistency.

## 4. RESULTS

For our simulation, we use $h=100$ nm, $c=20$ nm and $w=20$ nm for the constriction (refer to Fig. 1 for the meaning of these parameters). The Fermi energy assumed was 18.84 meV, which gives rise to 5 propagating modes for the given dimensions in the wire itself, while the QPC allows only two propagating modes. We begin by solving Schrödinger's equation assuming the hard-wall boundaries shown in Fig. 1 and neglecting the self-consistent solution for the potential. The resulting electron density for this case is shown in Fig. 4(a). Note that two collimated beams exit the QPC at distinct angles. These correspond to the two propagating modes which are *diffracted* by the QPC. This diffraction, or electron beam collimation, a

byproduct of the electron momentum in the *y*-direction being quantized in the constriction which thus defining the exit angles for the electron *waves*, constitutes the primary effect that quantum mechanics produces in this system.

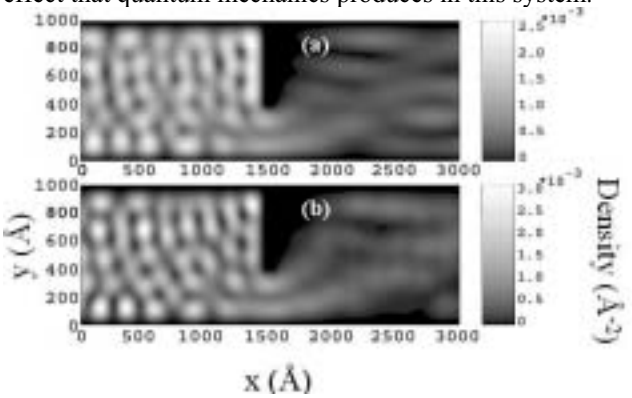



Figure 4: Density plots for (a) a bare hardwall potential and (b) a calculation with full self-consistency.

We now consider results with self-consistency. Since the net self-consistent potential is negative (mainly due to the inclusion of the correlation energy as part of the calculation), the tendency is for the density in the structure to increase, giving rise to more propagating modes. However, since we constrain the number of modes and the 1D density to be constant, we must adjust the local Fermi energy after each iteration of the Schrödinger equation. Once the iteration has converged, the Fermi energy in the device is now 5.46 meV. The new Fermi energy is lower than the original reference level since the self-consistency has resulted in a lower potential inside the wire. A similar result was found by Wang and Berggren[7]. The electron density with the self-consistency included is shown in Fig. 4(b). It differs from Fig. 4(a) only in minor details. This result is not particularly surprising, since non-self-consistent models have been able to well describe such structures[2].

Fig. 5(a) shows the self-consistent potential after a single iteration while Fig. 5(b) shows it for the case where convergence has been achieved. The two are quite similar. Superimposed are Bohm trajectories for the electrons, which are determined by a velocity field generated by the wavefunction solutions of the Schrödinger equation[8]. Adopting Bohm's point of view, one retains the idea that electrons act as particles rather than as waves, but that the quantum mechanics effects their motion

Writing $\psi = \mathrm{Re}^{\frac{iS}{\hbar}}$, the corresponding velocity field is [8]:

$$\mathbf{v}(\mathbf{x},t) = \frac{\nabla S(\mathbf{x},t)}{m^*} = \frac{\hbar}{m^*}\frac{\mathrm{Im}(\psi^*\nabla\psi)}{|\psi|^2}. \qquad (20)$$

As shown in Fig. 5, the Bohm trajectories reflect the dominant quantum effect, which is the diffraction.

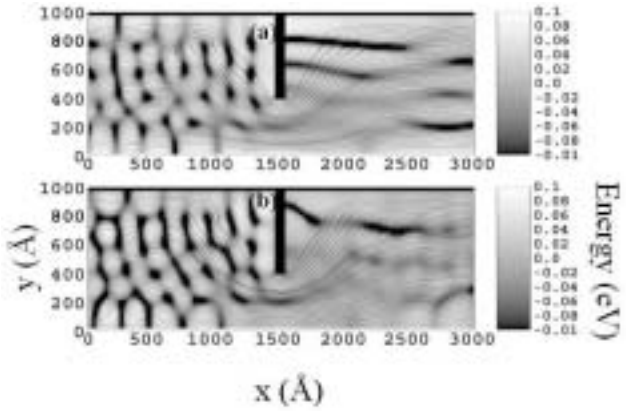



Figure 5: Potentials for (a) the first iteration of the self-consistent loop and (b) the final iteration. The lines represent Bohm trajectories.

To conclude, we have described a method of simulating ballistic nanostructures that includes self-consistency. In open structures such as the QPC shown here, this self-consistency is relatively easy to achieve. More difficult are situations where quasi-bound states are involved, such as in quantum dots, since the density changes quite rapidly. Generally, one must heavily damp the iteration process to get convergence in that situation.

## REFERENCES


[1] M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro and H. Iwai, IEEE Trans. Electron Devices 42, 1822, (1995).

[2]D.K. Ferry and S.M. Goodnick, in *Transport in Nanostructures*, Ed. by K. Board and D. R. J. Owen, Swansea UK (Cambridge Press, 1997).

[3] T. Usuki, M. Saito, M. Takatsu, R.A. Kiehl, and N. Yokoyama, Phys. Rev. B, 52, 8244 (1995).

[4] D.Y.K. Ko and J.C. Inkson, Phys. Rev. B, 38, 9945 (1988).

[5] Y. Wang, J. Wang, H. Guo and E. Zaremba, Phys. Rev. B, 52, 2738 (1995).

[6] Y. Sun and G. Kirczenow, Phys. Rev. B, 47, 4413 (1993).

[7]C. Wang and K.F. Berggren, Phys. Rev. B, 54, 14257 (1996).

[8] D. Bohm and B.J. Hiley, *The Undivided Universe*, (Routledge, 1993).