Improving the Accuracy of the Schrödinger-Poisson Solution in CNWs and CNTs

M. Rudan, A. Gnudi, E. Gnani, S. Reggiani, G. Baccarani

Advanced Research Center on Electronics Systems (ARCES) and Department of Electronics (DEIS)

University of Bologna, Via Toffano 2/2, 40125 Bologna, Italy

Phone: +39-051-20-93016, Fax: -93073, email: mrudan@arces.unibo.it

Abstract

The Schrödinger equation, or the coupled Schrödinger and Poisson equations, are transformed into an integral equation. Back-substituting from the original equations allows one to approximate the numerical corrections to any order without the need of calculating derivatives of the unknown function of order larger than one. Typical applications are in the numerical analysis of quantum transport in nanowires and nanotubes in the ballistic regime.

Coupled Schrödinger-Poisson Equations

In the full-quantum analysis of ballistic transport in cylindrical-nanowire (CNW) and carbon-nanotube (CNT) transistors, the problem is often solved by decoupling the Schrödinger equation along the radial (r) and logitudinal (z) coordinate [1]. After discretizing the latter from the source ($z = 0$) to the drain ($z = L$) end of the channel, the Schrödinger equation is solved over the transverse section at each grid node z_i. Being this a closed-boundary problem, it yields a set of eigenvalues that provide the ground energies of the subbands for the longitudinal problem. The latter is then tackled by solving the longitudinal Schrödinger equation with open-boundary conditions (this is also called *quantum-transmitting boundary method*, QTBM). In this part of the solution the total energy E of the electron may take any value within each subband. For the electrons that are injected from the source (drain) into the channel the weight of each E is prescribed by the Fermi statistics f and the density of states g of the source (drain) lead.

The solution of the Schrödinger equation in the silicon CNW is based upon the parabolic-band approximation. It is assumed that the longitudinal coordinate z corresponds to the [001] crystallographic direction and that the two valleys aligned with such direction have the indexes 3 and 6. As the valley of index 3 is centered at $(0, 0, k_{30})$ it is convenient to use for the wave function the replacement $\psi \leftarrow \psi \exp(-jk_{30}z)$. This centers the valley at the origin. Using the cylindrical coordinates one finds for the kinetic part of the Hamiltonian,

$$\mathcal{T}_3 = -\frac{\hbar^2}{2m_t}\left(\frac{\partial^2 \psi}{\partial r^2} + \frac{1}{r}\frac{\partial \psi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \psi}{\partial \varphi^2}\right) - \frac{\hbar^2}{2m_l}\frac{\partial^2 \psi}{\partial z^2}, \quad (1)$$

where φ is the angular coordinate and m_t, m_l the transversal and longitudinal mass, respectively. A similar expression holds for $\mathcal{T}_6$. If the potential energy U were independent of φ, then the Hamiltonian $\mathcal{H}_3 = \mathcal{T}_3 + U$ would be rotationally invariant. The situation is different for the other valleys.

For instance, assuming that the two valleys aligned with the [100] crystallographic direction have the indexes 1 and 4 one finds

$$-\frac{2}{\hbar^2}\,\mathcal{T}_{1(4)}(r, \varphi) = c_1\frac{\partial^2 \psi}{\partial r^2} + c_2\left(\frac{1}{r}\frac{\partial \psi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \psi}{\partial \varphi^2}\right) +$$

$$+ c_3\left(\frac{1}{r^2}\frac{\partial \psi}{\partial \varphi} - \frac{1}{r}\frac{\partial^2 \psi}{\partial r\partial \varphi}\right) + \frac{1}{m_t}\frac{\partial^2 \psi}{\partial z^2}, \quad (2)$$

with $c_1 = \cos^2(\varphi)/m_l + \sin^2(\varphi)/m_t$, $c_2 = \cos^2(\varphi)/m_t + \sin^2(\varphi)/m_l$, $c_3 = (1/m_l - 1/m_t)\sin(2\varphi)$. The above shows that the kinetic part $\mathcal{T}_1$ ($\mathcal{T}_4$) of the Hamiltonian $\mathcal{H}_1$ ($\mathcal{H}_4$) has a periodicity of π. The kinetic part of the Hamiltonian for the other two valleys is derived from (2) by interchanging m_l and m_t or, equivalently, by replacing φ with $\varphi + \pi/2$. Hence the periodicity of $\mathcal{T}_2$, $\mathcal{T}_5$ is the same as that of $\mathcal{T}_1$, $\mathcal{T}_4$.

The potential energy U that appears in the Schrödinger equation is initially taken from a classical solution, and is then updated by feeding the charge density back into the Poisson equation $\nabla^2 U = q\varrho/\epsilon$. Specifically, after solving the Schrödinger equation, the charge density ϱ at each section is calculated from the electron concentration $n_e = n_S(r, z) + n_D(r, z)$, where the suffixes refer to the source and drain leads, respectively. Only the contribution of the electrons is considered because the material is intrinsic and the hole concentration is negligible in the typical operating regimes. The procedure is iterated until self-consistency is reached.

The electron concentration is calculated from the wave functions derived from the Hamiltonians discussed earlier. As the Hamiltonians are periodical in φ, the potential energy retains such periodicity. This implies that the angular coordinate φ enters the self-consistent solution of the Schrödinger-Poisson system. To avoid the complicacy the approximation used in [1] is adopted here, namely, that of averaging the coefficients c_1, c_2, c_3 over φ. This yields $c_1, c_2 \leftarrow 1/m_a \doteq (m_l + m_t)/(2m_lm_t)$, $c_3 \leftarrow 0$. As a consequence the wave function retains a dependence on φ only through the factor $\exp(j\mu\varphi)$, with $\mu = 0, \pm1, \pm2 \ldots$, and the term containing the φ-derivative in (1,2) is transformed as

$$\frac{1}{r^2}\frac{\partial^2 \psi}{\partial \varphi^2} \leftarrow -\frac{\mu^2}{r^2}\,\psi. \quad (3)$$

The variables are then scaled, namely, $r \leftarrow \sqrt{m_t/m}\,r$, $z \leftarrow \sqrt{m_l/m}\,z$ in valleys 3 and 6, while in the remaining valleys $r \leftarrow \sqrt{m_a/m}\,r$, $z \leftarrow \sqrt{m_t/m}\,z$, with m the mass

of the free electron. In this way the kinetic part of the Hamiltonian becomes isotropic and identical for all valleys. In turn, due to the variable scaling, the potential energy changes from a valley to another, $U \leftarrow U_s$. The problem is thus reduced to solving the same differential equation using different data. The solution domain is a rectangle of sides L_s, R_s, that are the scaled version of the original source-drain distance L and radius R of the device. The radius includes the conductive channel and gate insulator.

Factorization of the Eigenfunctions

The solution of the Poisson equation is tackled as a fully two-dimensional problem, in which the electric potential $-U/q$ is prescribed in the gate contact. Homogeneous Neumann conditions are used along the remaining boundaries. From the numerical standpoint the solution is achieved by discretizing the equation over a tensor-product grid superimposed to the solution domain. The potential energy found by solving the Poisson equation is not separable in general. Due to this, the solution of the Schrödinger equation is tackled by recasting the latter into a system of two coupled equations. The procedure is the same as in the Born-Oppenheimer theory and is based upon the factorization $\psi \exp(-j\mu\varphi) = \chi(r,z)\,w(z)$, where χ, w are solutions of

$$\mathcal{H}_{\mu s}(z)\,\chi_{n\mu s} = \eta_{n\mu s}(z)\chi_{n\mu s}\,, \qquad (4)$$

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}z^2} + V_{n\mu s}(z)\right] w_{kn\mu s} = E_{kn\mu s}w_{kn\mu s}\,, \qquad (5)$$

respectively, with $V_{n\mu s}(z) \doteq \eta_{n\mu s}(z) + \omega_{n\mu s}(z)$ and

$$\mathcal{H}_{\mu s} \doteq -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{\mu^2}{r^2}\right) + U_s(r,z)\,, \qquad (6)$$

$$\omega_{n\mu s}(z) \doteq -\frac{\hbar^2}{2m}\int_0^{R_s} \chi_{n\mu s}\frac{\partial^2 \chi_{n\mu s}}{\partial z^2}\,\mathrm{d}r\,. \qquad (7)$$

Eq. (4) is solved at each section z of the device, in practice at each discretization point z_i in the longitudinal direction. Index n derives from the quantization in the radial direction whereas index s counts the valleys. After $\eta_{n\mu s}$, $\omega_{n\mu s}$ are found as functions of z for each triplet n, μ, s, they are used to build-up the potential energy in (5). Besides the inherited indices n, μ, s, the eigenvalues of (5) depend on the index k, which is continuous because the equation refers to an open-boundary problem. It is useful to note that, due to the application of the scaling factors to the coordinates, the indices n and k turn out to be scaled by the inverse factor of r and z, respectively.

The density of states $g_{n\mu s}(E)$ is found from the dispersion relation $E_{kn\mu s} = E_{n\mu s}(k)$. It follows that the dependence on k of the eigenfunctions w is transformed into a dependence on E. In the source and drain leads the potential energy U_s is independent of z, whence $\eta_{n\mu s} = $ const, $\omega_{n\mu s} = 0$, $w_{kn\mu s} \propto \exp(jkz)$, $E_{n\mu s}(k) = \eta_{n\mu s} + \hbar^2 k^2/2m$. As a consequence $g_{n\mu s} \propto (E - \eta_{n\mu s})^{-1/2}$. The

calculation of the density of states in the channel is more involved and must be carried out numerically.

The electron concentration n_e is finally calculated by adding up the contributions from the different eigenfunctions. The concentration due to the electrons injected from the source is

$$n_S(r,z) = \sum_{n\mu s} |\chi_{n\mu s}|^2 \int_{\eta_{n\mu s}^S}^{\infty} |w_{n\mu s}^S|^2 g_{n\mu s}^S f\,\mathrm{d}E\,, \qquad (8)$$

with $f = f(E)$ the Fermi statistics. The apex S in (8) reminds that for matching the boundary conditions of w in (5) one assumes that the electron is injected from the source (more details are given below). An expression similar to (8) holds for n_D. As mentioned earlier, the charge density $\varrho = -qn_e$ is used to solve the Poisson equation in two dimensions, and the electric potential thus found is fed back into the Schrödinger equation. The procedure is iterated until convergence is reached.

As far as the numerical load is concerned, the calculation of the charge density ϱ depicted above is by far the most demanding part of the procedure. It is therefore useful to investigate methods able to solve the whole problem without the need of calculating ϱ in an exceedingly high number of sections. For this reason we concentrate on the longitudinal part (5) of the Schrödinger equation. Dropping the indices for the sake of simplicity, letting $b(z) = 2m\,(E - V)/\hbar^2$, and using primes to indicate the derivatives, the equation becomes $w'' + bw = 0$.

As mentioned above, the latter is tackled by QTBM. This may lead to instability and possible lack of convergence. It will be shown here that the accuracy of the solution method is improved by converting (5) into an integral equation. Besides exploiting the smoothing property of the integral, the method increases the order of the solution. This is achieved without the need of involving extra nodes of the discretization grid but those belonging to the element under consideration. As a consequence, the method can be implemented into standard simulation codes leaving their structure unchanged.

Before entering into the details it is convenient to discuss the boundary conditions of the longitudinal equation (5). In fact, its solution w in the interval $[0, L_s]$ must be matched with the solutions in the leads. Considering, e.g., the electrons injected from the source at some energy $E = \eta + \hbar^2 k_S^2/(2m)$, the solutions in the leads have the form $a_1 \exp(jk_S z) + a_2 \exp(-jk_S z)$ for $z \leq 0$, and $a_5 \exp(jk_D z)$ for $z \geq L_s$. In general it is $k_D \neq k_S$ because of the bias applied between the source and drain leads. The constancy of the density of probability flux yields $R+T = 1$, where $R = |a_2/a_1|^2$ is the reflection coefficient and $T = (k_D/k_S)\,|a_5/a_1|^2$ the transmission coefficient.

The matching of the boundary conditions requires that w and w' are known at 0 and L_s, which is not true in general. The problem is overcome by considering, instead of w, two fundamental solutions u, v of $w'' + bw = 0$. Letting

$$w = a_3 u + a_4 v\,, \qquad (9)$$

with a_3, a_4 undetermined coefficients, and using the boundary conditions $u_0 = u(0)$, $u'_0 = 0$, $v_0 = 0$, $v'_0 = 1/u_0$, the solution of the equations $u'' + bu = 0$ and $v'' + bv = 0$ is fully determined. From this, one calculates $u_s = u(L_s)$, $u'_s = u'(L_s)$, $v_s = v(L_s)$, $v'_s = v'(L_s)$. It is worth noting that in order to determine the above values it is not necessary to actually solve both equations for u and v, because v can be reconstructed from u. Finally, the elimination of a_3, a_4 by the matching of the boundary conditions yields the transmission coefficient

$$T = \left[\frac{1}{2} + \frac{(u'_s)^2 + (k_D u_s)^2}{4 u_0^2 k_S k_D} + \frac{(v'_s)^2 + (k_D v_s)^2}{4 k_D/(u_0^2 k_S)} \right]^{-1} .$$ (10)

It is known from the theory that the fundamental solutions u, v are proportional to u_0 and $1/u_0$, respectively. As a consequence u_0 can be chosen arbitrarily, and the expression (10) is fully general. The coefficient a_1 of the injected wave can be chosen arbitrarily as well, so that the matching of the boundary conditions provides also the expressions for a_3, a_4 in terms of a_1. One finds the relations

$$\frac{u_0 a_3}{a_1} = \frac{k_S k_D u_0 v_s + j k_S u_0 v'_s}{D/2} ,$$ (11)

$$\frac{a_4}{u_0 a_1} = \frac{j k_S u'_s/u_0 - k_S k_D u_s/u_0}{D/2} ,$$ (12)

with $D \doteq k_D k_D u_0 v_s - u'_s/u_0 + j(k_S u_0 v'_s + k_D u_s/u_0)$. Thanks to this, the solution in the interval $[0, L_s]$ is eventually determined through (9) as

$$\frac{w}{a_1} = \frac{u_0 a_3}{a_1} \frac{u}{u_0} + \frac{a_4}{u_0 a_1} u_0 v .$$ (13)

Solution by Integral Equations

As mentioned above it is sufficient to tackle only the first fundamental solution u. Taking any node z_i as origin, and considering the grid element $h_{i+1} = z_{i+1} - z_i$, the Schrödinger equation to be solved has the form $u'' + bu = 0$. The integral-form solution is

$$u_{i+1} = u_i + u'_i h_{i+1} + \int_{z_i}^{z_{i+1}} (z - z_{i+1}) \, bu \, \mathrm{d}z .$$ (14)

The integral in (14) is the correction to be approximated numerically. It is calculated by expanding bu into a Taylor series around z_i. Back-substituting from the original equation, the expansion can be carried out to any order without introducing derivatives of u of order higher than the first. For instance, the coefficient of the third-order term of the expansion reads $[(b'''_i - 4b'_i b_i)u_i + (3b''_i - b_i^2)u'_i]/6$. It follows that the practical issue here is the accuracy with which the derivatives of b are calculated. In any case, if the expansion of bu is truncated to order λ, it is seen by inspection that the integral in (14) is accurate to the $(\lambda + 2)$th power of h_{i+1}. The two-order gain is intrinsic to the double integration that brings the original differential equation to the form (14). No matter where the truncation occurs, to calculate the integral in (14) it is necessary to know only u_i and u'_i,

so the values of the unknown at other grid nodes are not involved. The derivative of u that is necessary to proceed from the $(i+1)$th to the $(i+2)$th element in (14) is given by $u'_{i+1} = u'_i - \int bu \, \mathrm{d}z$, which is accurate to the $(\lambda + 1)$th power of h_{i+1}.

From the practical standpoint one notes that at the time of solving the Schrödinger equation the function b is known. In fact it is available from the solution of the transversal equation (4) carried out at each grid node z_i at the beginning of each iteration step. Truncating, e.g., to the second order the expansion of bu yields a fourth-order accuracy, representing a fair improvement over the standard finite-difference or box-integration methods, which are second-order accurate. This is reached at no additional cost and without the necessity of separating the Poisson equation, which in fact is solved in two dimensions.

The method depicted here may further be extended when the coefficient $b(z)$ of the Schrödinger equation is derivable from a one-dimensional Poisson equation $b'' + \sigma = 0$. The integral form of the latter is

$$b(z) = b_i + b'_i(z - z_i) + \int_{z_i}^{z} (\zeta - z) \, \sigma \, \mathrm{d}\zeta .$$ (15)

Typically the boundary conditions for the Poisson equation are prescribed on b at the channel ends. Because of this it is necessary to preliminarily determine b'_0 from $b(L_s) = b_0 + b'_0 L_s + \int_0^{L_s} (\zeta - L_s) \sigma \, \mathrm{d}\zeta$. As noted earlier, when (14) and (15) are tackled the charge density is known at each node from the previous iteration. Thus, a polynomial interpolation can be used for calculating $\sigma(\zeta)$ in (15). The simplest one involves only the nodes of element h_{i+1}, viz., $\sigma = \sigma_i + \sigma'_i(\zeta - z_i)$ which makes $b(z)$ a third-degree polynomial. Using more interpolating nodes would provide a higher degree. Remembering the discussion above, thanks to the analytical form of b the expansion for calculating the integral in (14) can be carried out to any order using u_i and u_{i+1} only. In the examples shown below a third-order expansion is used for both b and u.

Results and Conclusion

As a first model problem we choose a solution exhibiting a near divergence at the drain end of the channel, namely, $w = 1 - \log[(1-z)/(1+z)]$. For simplicity a scaled variable is used here, such that the source and drain ends are at $z = 0$ and $z = 1 - \epsilon < 1$, respectively. The separability of the Poisson equation is also assumed, so that the analytical form of σ is derived from $\sigma = \mathrm{d}^2(w''/w)/\mathrm{d}z^2$. After calculating the boundary conditions for w and b at the source, the open integration from source to drain is carried out several times using coarser and coarser grids with uniform node spacing. The accuracy is monitored at the drain end where, as expected, the error is the largest. The relative error of the method proposed here is compared with that of the standard Box-Integration Method (BIM) in Table I, where the relative errors at $z = 1 - \epsilon$ are compared for different grid sizes.

TABLE I

$\epsilon = 0.1$, $w(1-\epsilon) = 3.94$	Relative error	
Number of nodes N	BIM	This work
500	1.0×10^{-3}	2.8×10^{-6}
200	2.4×10^{-3}	1.6×10^{-5}
100	4.5×10^{-3}	5.4×10^{-5}
50	8.1×10^{-3}	1.6×10^{-4}
20	1.5×10^{-2}	3.6×10^{-4}

TABLE II

$w(1) = 1$		Relative error	
N	ν	BIM	This work
500	1	7.8×10^{-5}	1.4×10^{-11}
500	4	1.1×10^{-3}	5.5×10^{-8}
500	16	1.6×10^{-2}	2.2×10^{-4}
50	1	6.9×10^{-3}	1.2×10^{-6}
50	4	3.2×10^{-2}	4.5×10^{-3}

As a second model problem we choose an oscillating solution $w = \cos(2\pi\nu z)$ whose oscillations are made sharper by increasing the spatial frequency ν. The source and drain ends are at $z = 0$ and $z = 1$, respectively. No separability of the Poisson equation is assumed here, so the coefficient of the Schrodinger equation is prescribed directly as $b = (2\pi\nu)^2$. The procedure is similar to that of the first model problem. The open integration from source to drain is carried out several times changing both the grid coarseness and spatial frequency, and the accuracy is monitored at the drain end. The relative error of the method is compared with that of the Box-Integration Method (BIM) in Table II, where the relative errors at $z = 1$ are compared for different grid sizes.

In conclusion, the method illustrated in this paper improves the accuracy of the solution of the Schrödinger equation, or of the coupled solution of the Schrödinger and Poisson equations, when a one-dimensional approximation is applicable. The improvement is achieved without the need of extra calculations of the charge density which, as illustrated above, is typically the heaviest part of the computation. On the other hand, the better accuracy on the longitudinal wave function improves also the calculation of the charge density, thus positively influencing the number of global iterations.

References

[1] E. Gnani, A. Marchi, S. Reggiani, M. Rudan, G. Baccarani, *Comparison of device performance and scaling properties of cylindrical-nanowire (CNW) and carbon-nanotube (CNT) transistors*, Proc. SISPAD-2006, Monterey, CA (2006).