A fast approach to the simulation of silicon nanowire transistors

Paolo Marconcini

Abstract—In this article the transport properties of field-effect transistors based on thin silicon nanowires are studied limiting the computationally demanding atomistic analysis to the determination of a restricted set of parameters which are then used in a less-detailed description of the overall device. In particular, at the beginning of the simulation a tight-binding analysis of a simplified structure is performed, in order to find the main low-level properties of the considered silicon nanowire. These quantities are then passed to another part of the algorithm which, using a continuum, higher-level and thus simpler description, derives the electrical characteristics of the complete nanowire transistor, through the self-consistent solution of the electrostatic and transport equations. I show that this multi-level approach allows to obtain results in very good agreement with fully atomistic simulations, with a considerable time saving.

Keywords—multi-level simulation, nanoelectronic devices, silicon nanowire transistors

I. INTRODUCTION

In the last decades, the demand for smaller and more powerful devices has led to a progressive downscaling of electronic components. Nowadays, the characteristic dimensions of electronic devices have reached the nanometer scale, approaching the few-atom limit, and an intense research activity is focused on the development of nanoelectronic devices.

In this context, the activity of computer simulation plays a key role, both in order to fully understand the physical behavior of the devices and to improve their performances. Indeed, at these dimensions, quantum-mechanical effects have to be explicitly considered to correctly analyze and predict the electrical behavior of a device and also atomic details can have a strong impact on its characteristics. However, an atomistic, full-quantum simulation of the overall device requires a great computational effort, especially if the number of atoms to consider is quite large and the crystalline structure of the involved materials is complex. This makes such a kind of approach impractical for example when an in-depth exploration of the design parameter space is needed, and therefore the simulations have to be repeated for a large number of device geometries or operating conditions.

A good trade-off between the accuracy of the results and the efficiency of the simulation program can often be reached using a hierarchical multi-scale approach. Using this simulation strategy, the system under consideration is studied using a hierarchy of levels of treatment. For each fundamental property and interaction of the system the right level of approximation is identified: the adoption of very detailed (low-level) descriptions (such as atomistic, ab initio models) is limited to small sub-regions of the device or to a simplified structure, while the overall device is treated with less accuracy (i.e. at a higher level). The results obtained at the lower level (which are strictly related to local interactions and atomic details) are passed to the higher level in the form of a limited number of condensed parameters, which are sufficient for a less-detailed description of physical interactions. In general, some data can also be passed from the upper level to the lower level, in order to allow a reciprocal interaction between the different levels of approximation. For example, in the case I will describe in this article I will make a tight-binding, atomistic description of a simplified structure consisting of the infinite periodic repetition of the unit cell of the transistor channel, and then I will pass a reduced number of parameters (the effective masses and band edges of the dispersion relations and the transversal local density of states) to a higher-level, continuum, envelope-function (i.e. effective mass) description of the overall device. While this represents a two-level description, a hierarchical simulation approach can in general include several levels of approximation (as sketched in Fig. 1), the higher of which could be a lumped parameter model of a nanoelectronic circuit, analogous to that used in common electronic circuit simulators as Spice.

![Fig. 1 Schematic representation of a hierarchical multi-level approach to the simulation of nanodevices.](image)

The field of nanoelectronics includes both the last generations of CMOS transistors, scaled to the nanometer size through the adoption of new technological solutions, and more innovative devices, conceived to overcome the scaling problems of CMOS technology. This last category includes
devices which differ from CMOS ones e.g. for the adopted materials (for example semiconductor heterostructures [1-6], graphene [7-14], carbon nanotubes [15-17] or single molecules [18]), information carriers (for example spintronic devices [19], where the information is carried and stored by the electron spin, instead of by electron charge), operation principles (for example QCA’s [20-22] or quantum computing [23]), or geometries (as FinFets [24] or nanowire transistors [25]).

One of the main scaling problems of CMOS technology is the deterioration of electrostatic control of the channel: with the decreasing of the gate length, short-channel effects like DIBL (Drain-Induced Barrier Lowering, i.e. the effect of drain voltage on the threshold voltage, due to its electrostatic effect on the potential barrier between source and drain) increase and the gate control on the channel progressively reduces. One interesting solution is to use field effect transistors (FETs) in which the channel is represented by a nanowire. This allows to improve the electrostatic action of the gate, since it is possible to exploit the presence of several gates around the channel.

It has been shown that, down to cross sections of 4 nm×4 nm, numerical simulation of nanowire FETs can rely on an effective mass, \(\tilde{k} \cdot \tilde{p}\) approximation [26] based on the dispersion relation of bulk (unconfined) silicon. However, for smaller cross sections the effect of geometrical confinement on energy bands and on transport has to be considered and atomistic approaches need to be undertaken.

Here I describe the hierarchical multi-scale approach that I have used to simulate the electrical behavior of silicon-nanowire field-effect-transistors. In particular, I solve only once, at the beginning, the atomistic problem yielding the wave functions and energy eigenvalues of the nanowire. For each one-dimensional (1D) band, I extract, from the dispersion relationships, the corresponding effective mass and band edge, and from the wave function the transverse local density of states. These parameters are then used to solve the transport problem in a Green’s function formalism, considering an envelope-function description with decoupled bands. Comparing our results with those obtained by Luisier et al. [27] with a full-band atomistic simulation code for a triple-gate nanowire transistor, I show that the agreement is very good, with a significant reduction of the computational times. Finally I use the code to simulate the transport properties of a transistor based on a nanowire with circular section.

II. ATOMIC SIMULATION

In order to make a transport simulation of the device, a self-consistent solution of the three-dimensional (3D) Poisson (electrostatic) and Schrödinger (transport) equations with open boundary conditions is necessary. The 3D Schrödinger equation has been decoupled adiabatically in a two-dimensional (2D) equation in the transversal plane, and a 1D equation along the direction of transport (the longitudinal direction).

In order to obtain the local density of states on the transversal plane, as well as the parameters to use inside the 1D Schrödinger equation, I have performed once (at the beginning of the simulation process) an atomistic study, for a null external potential, of a simplified structure, consisting in an infinite silicon nanowire with the same section as that of the wire which represents the channel of our transistor.

For the atomistic analysis, I have adopted a real space \(sp^3d^5s^*\) nearest-neighbor tight-binding model, that I detail in the following. Since the considered nanowires have orientation [001], their unit cell includes 4 atomic planes. For each silicon atom belonging to the nanowire unit cell I have considered 10 orbitals or for the atom \(j\) (inside the unit cell) on which it is centered, I have considered the following Bloch function:

\[
B_{n,j}(k, \vec{r}) = \frac{1}{\sqrt{N}} \sum_\gamma e^{i \vec{k} \cdot \Delta \vec{z}} \varphi_{n,j}(\vec{r} - \vec{R}_j - l \Delta \vec{z}),
\]

where \(k\) is the longitudinal wave vector, \(\vec{r}\) the position vector, \(N\) the number of considered unit cells, \(l\) the index which identifies each unit cell, \(z_j\) and \(\vec{R}_j\), the longitudinal position and the three-dimensional position vector of the \(j\)-th atom of the unit cell identified by \(l = 0\), \(\Delta z\) the longitudinal length of the unit cell and \(\Delta \vec{z}\) the longitudinal unit vector, while the sum is done over all the \(N\) considered unit cells. These functions satisfy the Bloch theorem for the infinite lattice in the longitudinal direction \(z\). If we start from atomic orbitals orthonormalized using the Löwdin procedure, these Bloch functions represent an orthonormal basis on which the nanowire wave functions \(\Psi(k, \vec{r})\) can be expanded:

\[
\Psi(k, \vec{r}) = \sum_{n,j} C_{n,j}(k) B_{n,j}(k, \vec{r}).
\]

If, starting from the quantum mechanical equation of the system

\[
\hat{H} \Psi(k, \vec{r}) = E(k) \Psi(k, \vec{r}),
\]

where \(\hat{H}\) is the Hamiltonian operator, we substitute the expansion (2) and then we project on the Bloch function \(B_{n,j}\), we obtain, exploiting the completeness relation for the Bloch functions and using the Dirac notation that

\[
\sum_{n,j} \left( \langle B_{n,j}(k, \vec{r}) \rangle \right) \hat{H} \left( B_{n,j}(k, \vec{r}) \right) C_{n,j}(k) = E(k) C_{n,j}(k),
\]

which can be rewritten in the simpler matrix form:

\[
H(k) \tilde{C}(k) = E(k) \tilde{C}(k),
\]

where the Hamiltonian matrix \(H(k)\) describes the interaction between the Bloch functions. Since the Bloch functions are equal to (1), each element of \(H(k)\) can be written as a combination, with coefficients given by exponential terms depending on the geometry of the nanowire, of the Hamiltonian terms between single atomic orbitals. In our calculations only the interactions between nearest neighbor
atoms have been considered. The interaction between atomic orbitals, in its turn, has been derived using the relations given (writing the potential energy as sum of spherical potentials located on all the atoms and using a two-center approximation) by Slater and Koster [28] and, for generical atomic orbitals, by Podolskiy and Vogl [29], which express it in terms of the geometry of the nanowire and of the Slater-Koster parameters. For these last quantities, the same values proposed for bulk silicon by Boykin et al. [30] have been considered.

The matrix equation (5) is clearly an eigenproblem. Solving this equation for different values of the longitudinal wave number \( k \) I find as eigenvalues the values of the energy \( E \) as a function of \( k \), i.e. the dispersion relations of the system (its energy bands) and as eigenvectors the coefficients of the electron wave function expansion on the basis set of Bloch functions, from which the wave function can be derived.

In order to remove the effect of surface states, I have tried two alternative approaches. A method [31] requires to explicitly include in the calculation the presence on the surface of hydrogen passivation atoms, which saturate the dangling bonds of the silicon atoms located at the borders of the nanowire. In particular, for each hydrogen passivation atom I have to consider one (or two, if atomic orbitals with different spin are distinguished, as I have done when considering the spin-orbit effect) \( s \) atomic orbital (and therefore the corresponding Bloch function). The corresponding elements in the Hamiltonian matrix \( H(k) \) can be written starting from our knowledge of the nanowire geometry and using the Slater-Koster parameters fitted by Zheng et al. in [31]. An alternative, more approximate technique [32] consists in increasing the on-site energy of the outward oriented \( sp^3 \) hybridized orbitals in correspondence of the atoms at the border of the nanowire. In particular, since \( sp^3 \) hybridized orbitals are linear combinations of the original \( s \) and \( p \) atomic orbitals, the effect of this increase of the on-site energy (i.e. of the proper diagonal elements in the Hamiltonian matrix between \( sp^3 \) hybridized orbitals) on the Hamiltonian matrix relating the \( s \) and \( p \) orbitals can be found with a simple change of base. Since I have noticed that both the approaches capture the main features of the dispersion relations, I have adopted the second one, in order to include a smaller number of atoms in the calculation, in this way reducing the dimension of the Hamiltonian matrix and the times required to find its eigenvalues and eigenvectors.

In Fig. 2 I show the atomic structure of an infinite silicon nanowire with a 2.1 nm×2.1 nm square cross-section with the edges oriented in the \([1 0 0]\) and \([0 1 0]\) directions and the longitudinal axis in the \([0 0 1]\) direction, and in Fig. 3 I show its energy bands.

In Fig. 4 I show the bands obtained for the same nanowire with and without the effect of spin-orbit interaction. Since only small changes in the dispersion relations can be observed, the spin-orbit effect has been neglected and I have considered only 10 atomic orbitals for each silicon atom.

The probability to find an electron in the volume \( \Omega_0 \) around
an atom is given by the integral on $\Omega_0$ of the square modulus of the electron wave function. According to (2), the electron wave function is given by a linear combination of Bloch functions, which, in their turn, are given by the combination (1) of the atomic orbitals. Since in $\Omega_0$ only the atomic orbitals of the atom inside the volume have a non negligible value, and therefore the values of the indexes $l$ and $j$ are fixed, the value of that integral is equal to

$$\frac{1}{N} \sum_{ij} C^*_{ij} C_{ij} \int_{\Omega_0} \phi_{ij}(\vec{r} - \vec{R}_j - l \Delta z) \phi_{ij}(\vec{r} - \vec{R}_j - l \Delta z) d\vec{r}$$

and therefore, exploiting the orthonormality of the atomic orbitals, is proportional to the sum of the square moduli of the coefficients $C$ corresponding to all the atomic orbitals centered on the considered atom. This quantity, projected on the transversal section of the wire, represents in a point charge approximation the transverse local density of states.

The condensed parameters that I need to extract from the atomistic simulation and that are passed to the transport simulator are the band edges (the value of the dispersion relations in the extrema points, i.e. in our case for $k=0$), the longitudinal effective masses and the transversal density of states computed for $k=0$.

The longitudinal effective mass of each energy band is equal to the ratio between the square of the reduced Planck’s constant and the second derivative of $E(k)$, which has been obtained, using a 5-points discretization formula, from the value of the dispersion relations for $k=0$ and for two small positive values of $k$ (exploiting the symmetry of the dispersion relations around $k=0$). Therefore the solution of the atomistic problem (5) for only three values of $k$ is sufficient to obtain all the required parameters.

In Fig. 5 I show, for four different sizes (differing for multiples of the lattice constant $a=5.4 \, \text{Å}$) of a silicon nanowire with square section and the same orientation considered in Fig. 2, the values of the energy gap and of the modulus of the effective masses corresponding to the lowest conduction band and to the highest valence band. As expected, the energy gap increases decreasing the wire size.

In Fig. 6 the transverse local density of states of the four lowest conduction bands of the nanowire represented in Fig. 2, computed for $k=0$ ((a) corresponds to the lowest band, (d) to the highest).
III. SELF-CONSISTENT POISSON-SCHröDINGER SOLVER

This set of parameters, obtained from the atomistic analysis, is then passed to the self-consistent Poisson-Schrödinger solver (based on the NanoTCAD ViDES code [33]), where I have adopted a less detailed, continuum description, in which the effect of atomic potentials on the electrons is taken into account through the use of the effective masses deriving from the tight-binding calculation.

In particular, the values of the band edges and of the effective masses are used to formulate the 1D envelope function (effective mass) transport equation (Schrödinger equation) in the longitudinal direction. Since I have verified that, for the small nanowires I have considered, the amount of mode mixing occurring as a result of the longitudinal variation of the potential along the channel is negligible, I have considered uncoupled 1D transport equations for the different energy bands. The value of the corresponding band edges has been translated, inside the equations, by the mean value of the potential applied to each transversal slice. Solving these 1D transport equations, with open boundary conditions, with the Non-Equilibrium Green's Functions (NEGF) formalism I have found the transmission coefficient and the longitudinal linear electron density for each band.

The product between the transversal local density of states computed with the atomistic procedure and the longitudinal linear electron density, summed up on the considered bands, gives the total 3D electron density that is needed in the electrostatic equation (Poisson equation) to find the potential inside the device.

Since in the devices I have simulated I have observed a very weak dependence of the electron wave functions on the actual transverse potential, I have used the transversal local density of states computed for a null external potential also in the presence of a nonzero potential. A point charge approximation has been assumed, where all the charge associated with an atom is spread over a volume cell.

Since the transport equation depends on the value of the potential, which is obtained from the solution of the electrostatic equation, but this last equation depends on the electron density, which derives from the transport equation, the two relations have to be solved self-consistently [34,35]. In particular, starting from an initial potential, I continue to solve the electrostatic and transport equations iteratively until the variation of the potential from an iteration to the next one drops under a fixed threshold (see the flowchart represented in Fig. 7). Once the convergence is reached, the value of the current flowing through the device is derived from the transmission coefficient using the Landauer formula.

IV. TRANSPORT RESULTS

In order to test the above described approach, I have simulated with our code the triple-gate silicon-nanowire field-effect-transistor studied by Luisier et al. [27] with an atomistic full-band self-consistent computation. Its transversal cross-section is represented in Fig. 8, while a three-dimensional sketch of the device is reported in Fig. 9. The silicon nanowire has the 2.1 nm×2.1 nm square cross-section already represented in Fig. 2, the oxide is 1 nm thick, the channel is 13 nm long and the source and drain reservoirs are 10 nm long, doped with a donor concentration equal to $10^{20}$ cm$^{-3}$.

Fig. 7 Flowchart of the simulation code ($V_{l,i}$ and $V_i$ are the potentials at the beginning and at the end of each iteration, $\Delta$ is the threshold that I have chosen for the variation of the potential).

Fig. 8 Transversal cross-section of the considered 2.1 nm×2.1 nm triple-gate silicon-nanowire transistor.
In the Section II, I have shown the results obtained from the atomistic analysis of the corresponding infinite nanowire for a null external potential. Only the lowest 4 conduction bands have been included in the transport calculations, since hole concentration is negligible at the considered bias points. In general, for high drain-to-source voltages and low gate voltages, confined states in the valence band can occur, degrading the behavior of the device in the off-state. This, however, is not the case for the device being considered, because of the high energy gap (≈ 2 eV), which makes such an effect negligible. However, it could be easily included in our approach.

In Fig. 10(a) I compare the output characteristics computed by means of our method (solid circles) with those obtained with the atomistic full-band approach (solid lines). No fitting with the results by Luisier et al. has been performed, except for a tuning of the gate work function. As can be seen, our results are in very good agreement with full atomistic simulations, at a significantly smaller cost in terms of required computational resources; in particular, a substantial speed up from the computational point of view has been obtained, by a factor of about 30.

The same quantitative agreement is achieved also in the subthreshold regime, as demonstrated by the transfer characteristics for $V_{DS}=0.4$ V in Fig. 10(b).

An analogous simulation has been performed for a transistor in which the nanowire with square cross-section is substituted by a wire with an approximately circular section and a diameter of 2.1 nm (shown in Fig.11). In Figs. 12 and 13 I show the dispersion relations and the transverse local densities of states in $k=0$ for the lowest conduction band obtained from an atomistic analysis of the corresponding infinite nanowire. In Fig.14 instead I report the transfer characteristic obtained for $V_{DS}=0.4$ V using the self-consistent Poisson-Schrödinger solver for the new transistor.
Fig. 12 Dispersion relations of an infinite silicon nanowire with the cross-section shown in Fig. 11.

Fig. 13 Transverse local densities of states in \( k=0 \) for the lowest conduction band of the infinite silicon nanowire with the cross-section shown in Fig. 11.

Fig. 14 Transfer characteristic for \( V_{DS}=0.4 \) V of the field-effect transistor based on the silicon nanowire represented in Fig. 11.

Elements of non-ideality such as surface roughness, vacancies or ionized impurities can indeed induce mode mixing, which, however, can still be included in our hierarchical approach at the expense of computing the transversal density of states for each slice and of using a multi-band effective mass equation (instead of considering uncoupled bands).

V. CONCLUSION

Here I have studied the transport behavior of silicon nanowire transistors, using a simulation approach in which the results of an atomistic analysis, limited to a simplified structure (an infinite wire with the same cross-section as the transistor channel), are passed in the form of a small number of parameters (the transversal density of states, the effective masses and the band edges) to a less-detailed (envelope function) description of the overall device.

This two-level simulation has shown a very good quantitative agreement with full atomistic calculations, with a notable reduction of the required execution times.

ACKNOWLEDGMENT

I acknowledge the important contribution of Dr. Gianluca Fiori and Prof. Giuseppe Iannaccone to the development of the simulation code and to the achievement of part of the reported results.

REFERENCES