Atomistic full-band simulations of silicon nanowire transistors: Effects of electron-phonon scattering

Mathieu Luisier and Gerhard Klimeck
Network for Computational Nanotechnology and Birck Nanotechnology Center, Purdue University, 465 Northwestern Avenue, West Lafayette, Indiana 47907, USA

(Received 27 July 2009; revised manuscript received 22 September 2009; published 14 October 2009)

An atomistic full-band quantum transport simulator has been developed to study three-dimensional Si nanowire field-effect transistors in the presence of electron-phonon scattering. The nonequilibrium Green's function (NEGF) formalism is solved in a nearest-neighbor $sp^3d^5s^*$ tight-binding basis. The scattering self-energies are derived in the self-consistent Born approximation to inelastically couple the full electron and phonon energy spectra. The band dispersion and the eigenmodes of the confined phonons are calculated using a dynamical matrix that includes the bond and the angle deformations of the nanowires. The optimization of the numerical algorithms and the parallelization of the NEGF scheme enable the investigation of nanowire structures with diameters up to 3 nm and lengths over 40 nm. It is found that the reduction in the device drain current, caused by electron-phonon scattering, is more important in the ON state than in the OFF state of the transistor. Ballistic transport simulations considerably overestimate the device ON currents by artificially increasing the charge injection mechanism at the source contact.

DOI: 10.1103/PhysRevB.80.155430 PACS number(s): 73.63.−b, 05.60.Gg, 72.10.−d, 63.20.kd

I. INTRODUCTION

Semiconductor nanowires (NWs) have emerged as potential active components of future integrated circuits beyond the conventional metal-oxide-semiconductor field-effect transistor (MOSFET) technology.[1] Nanowires made of different materials, crystal orientations, cross-section shapes, and dimensions have been successfully synthesized opening perspectives in both electronics and optoelectronics.[2–10] As FETs, nanowires provide an excellent electrostatic control of the channel due to their one-dimensional (1D) transport properties and can therefore operate at lower supply voltages than the traditional MOSFETs.

The fabrication of nanowire FETs is still a technology under development that requires further innovations before challenging state-of-the-art MOSFETs. Physics-based device simulations can support the experimental work to accelerate the development of NW FETs, reduce their cost, identify their strength and weakness, and demonstrate their scalability down to the sub-20 nm range. Classical and semiclassical approaches such as drift diffusion[11] and the Boltzmann transport equation[12] are very popular among the semiconductor industry for their ability to deliver simulation results in a short time. However, at the nanometer scale, these methods fail and must be replaced by computationally more intensive but physically more accurate quantum-mechanical models that include an atomistic device representation.

Ballistic simulations of NW FETs based on an effective-mass treatment of the Schrödinger equation[13–15] represent a substantial advancement as compared to drift diffusion and the Boltzmann equation but they do not correctly capture the quantization of the energy levels in structures with a cross section smaller than 5×5 nm^2.[16] To overcome the limitations of the effective-mass approximation, full-band three-dimensional (3D) transport simulations of NW FETs based on tight-bindinglike orbital basis have been recently reported.[17–20] In these atomistic studies, the diameters of the nanowires do not exceed 1.5 nm, with the notable exception of Ref. 20, where a more efficient numerical approach[21,22] has been used that enables the simulation of devices with diameters up to 10 nm.

Apart from the electron and hole bandstructure and 3D electrostatic charging effects, the interactions of the free carriers with their environment (impurities, open surfaces, or lattice vibrations) strongly affect the performances of NW FETs. For example, electron-phonon scattering is expected to drastically deteriorate the ON current of devices with diameters below 5 nm.[10] Modeling such many-body, inelastic interaction phenomena at a quantum-mechanical level is computationally very challenging. The nonequilibrium Green's function (NEGF) formalism[23–26] has established itself as a powerful approach to account for electron-phonon scattering in device simulations. Due to the complexity of the physical models, the 3D simulators that include dissipative scattering are based on the effective-mass approximation (no atomistic representation) and decouple the current transport direction from the transverse directions of confinement (mode-space approach reducing transport to a 1D problem).[27–31]

There have been some attempts to include both a multiband tight-binding model and electron-phonon scattering into a single device simulator using the NEGF formalism but they have been restricted to 1D structures only, such as resonant-tunneling diodes[32,33] and quantum well solar cells[34] or to molecules composed of very few atoms, typically less than 20.[35–37] Finally, alternative methods to NEGF have been considered to treat electron-phonon scattering in ultrascaled NW FETs such as the Pauli master equation[38] and a modified Monte Carlo approach.[39]

In this paper, a 3D, atomistic simulation approach that combines a full-band model and electron-phonon scattering in the self-consistent Born approximation is presented. The nearest-neighbor $sp^3d^5s^*$ tight-binding method and the NEGF formalism are used together to simulate quantum

transport in Si nanowire FETs. The electron-phonon interaction is calculated directly from the tight-binding Hamiltonian and from the band dispersion and the eigenmodes of the phonons confined in the nanowire structures. Despite the heavy computational burden of this atomistic approach, circular gate-all-around (GAA) NW FETs longer than 40 nm, with diameters up to 3 nm, and composed of more than 14 000 atoms are considered. The availability of large supercomputers and the efficient parallelization of the different tasks are essential to reach these dimensions that no other full-band simulator, even ballistic, has been able to treat so far.[40] The degradation of the ON current due to electron-phonon scattering in Si NW FETs with diameters of 2, 2.5, and 3 nm is investigated as an application.

The paper is organized as follows: in Sec. II, the simulation approach is introduced from the rigorous calculation of the electron-phonon self-energies and of the Green's functions to the necessary simplifications and approximations to keep the computational burden manageable. Details about the numerical implementation are given in Appendix. Applications of the formalism to different Si NW FET structures and the analysis of their ON currents are shown in Sec. III. The paper is concluded in Sec. IV and an outlook on future device simulations is given.

II. THEORY

A. General formalism description

Electron-phonon scattering is treated in the framework of the NEGF formalism under steady-state conditions.[23–26,41,42] The lesser ($G^<$), the greater ($G^>$), and the retarded (G^R) Green's functions are expressed in a nearest-neighbor tight-binding basis according to[21,32]

$$\sum_l \{[\mathbf{E}-\mathbf{E}_n-\mathbf{V}(\mathbf{R}_n)]\delta_{l,n}-\mathbf{H}_{nl}-\mathbf{\Sigma}^{RB}_{nl}(E)-\mathbf{\Sigma}^{RS}_{nl}(E)\}\mathbf{G}^R_{lm}(E) = \delta_{nm}, \tag{1}$$

$$\mathbf{G}^{\gtrless}_{nm}(E)=\sum_{l_1 l_2}\mathbf{G}^R_{nl_1}(E)[\mathbf{\Sigma}^{\gtrless B}_{l_1 l_2}(E)+\mathbf{\Sigma}^{\gtrless S}_{l_1 l_2}(E)]\mathbf{G}^A_{l_2 m}(E), \tag{2}$$

$$\mathbf{\Sigma}^R_{nm}(E)=\frac{1}{2}[\mathbf{\Sigma}^>_{nm}(E)-\mathbf{\Sigma}^<_{nm}(E)] + i\mathcal{P}\int\frac{dE'}{2\pi}\frac{\mathbf{\Sigma}^>_{nm}(E')-\mathbf{\Sigma}^<_{nm}(E')}{E-E'}, \tag{3}$$

$$\mathbf{G}^A_{nm}(E)=\mathbf{G}^{R\dagger}_{mn}(E). \tag{4}$$

The indices l, m, and n refer to the atomic positions $\mathbf{R}_l$, $\mathbf{R}_m$, and $\mathbf{R}_n$, respectively. The matrices $\mathbf{E}$ (diagonal, injection energy), $\mathbf{E}_n$ (diagonal, on-site energy), $\mathbf{V}(\mathbf{R}_n)$ (diagonal, self-consistent electrostatic potential at position $\mathbf{R}_n$), $\mathbf{H}_{nl}$ (nearest-neighbor coupling between atom n and l), $\mathbf{\Sigma}^B_{nl}(E)$ (boundary self-energy, different from 0 only if atoms n and l are directly connected to the semi-infinite device contacts, computed as in Ref. 21), $\mathbf{\Sigma}^S_{nl}(E)$ (scattering self-energy between atoms n and l), and $\mathbf{G}_{nm}(E)$ (Green's functions between atoms n and l) are of size $t_B\times t_B$, where t_B is the number of orbitals of the tight-binding model. The formulation is general for electron and hole transport. In this work, a $sp^3d^5s^*$ basis without spin-orbit coupling is used to describe Si (Ref. 43) so that $t_B=10$. The relatively small split-off energy of Si ($\Delta_0=45$ meV) justifies the neglection of spin-orbit coupling when only electron transport is considered as here. Each atom composing the simulation domain is treated individually. In Eq. (3), $\mathcal{P}$ denotes the Cauchy principal integral value. The definition and interpretation of the on-site energy $\mathbf{E}_n$ and nearest-neighbor coupling matrix $\mathbf{H}_{nl}$ can be found in Refs. 43–45.

The nearest-neighbor connection $\mathbf{H}_{nl}$ depends on the relative position of the atoms n and l, i.e., it is a function of $\mathbf{R}_l-\mathbf{R}_n$. If lattice vibrations are neglected, the positions of the atoms are fixed and time independent so that $\mathbf{R}_n(t)=\mathbf{R}^0_n$. When the lattice starts to vibrate the atoms oscillate around their equilibrium position $\mathbf{R}^0_n$ with an amplitude $\mu_n(t)$ and

$$\mathbf{R}_n(t)=\mathbf{R}^0_n+\mu_n(t). \tag{5}$$

This implies that $\mathbf{H}_{nl}$ becomes a function of $(\mathbf{R}^0_l-\mathbf{R}^0_n)+[\mu_l(t)-\mu_n(t)]$ and that it can be expanded around its equilibrium position to lowest order in $[\mu_l(t)-\mu_n(t)]$

$$\mathbf{H}_{nl}\approx\mathbf{H}^0_{nl}+\sum_i\frac{\delta\mathbf{H}_{nl}}{\delta(R^0_{l,i}-R^0_{n,i})}[\mu^i_l(t)-\mu^i_n(t)] \approx\mathbf{H}^0_{nl}+\sum_i\nabla_i\mathbf{H}_{nl}[\mu^i_l(t)-\mu^i_n(t)]. \tag{6}$$

The index i runs over the x, y, and z directions. In Eq. (6) the lattice vibrations are approximated as harmonic oscillations. The second, time-dependent term in Eq. (6) is assumed much smaller than $\mathbf{H}^0_{nl}$ and is treated in a perturbative way, leading to the electron-phonon scattering self-energies.[35–37]

In the second quantization picture the harmonic oscillations $\mu^i_l(t)$ are given by[46,47]

$$\mu^i_l(t)=\sum_{\lambda,\mathbf{q}}\sqrt{\frac{\hbar}{2M_l\omega_\lambda(\mathbf{q})}}f^i_\lambda(\mathbf{R}_l,\mathbf{q})[a^\dagger_\lambda(-\mathbf{q},t)+a_\lambda(\mathbf{q},t)], \tag{7}$$

where the indices λ and $\mathbf{q}$ depict the oscillator mode and wave vector, respectively, M_l is the mass of the atom at position $\mathbf{R}_l$ while the operator $a^\dagger_\lambda(-\mathbf{q},t)[a_\lambda(\mathbf{q},t)]$ creates (annihilates) a phonon in mode λ and momentum $\mathbf{q}$. The normalized atomic displacement $f^i_\lambda(\mathbf{R}_l,\mathbf{q})$ and the phonon frequency $\omega_\lambda(\mathbf{q})$ are calculated from the following harmonic-oscillator eigenvalue problem[48]

$$\sum_{m,j}\frac{\Phi^{ij}_{nm}(\mathbf{q})}{\sqrt{M_nM_m}}f^j_\lambda(\mathbf{R}_m,\mathbf{q})-\omega^2_\lambda(\mathbf{q})f^i_\lambda(\mathbf{R}_n,\mathbf{q})=0. \tag{8}$$

The dynamical matrix $\Phi^{ij}_{nm}(\mathbf{q})$ depends on the geometry of the device structure and determines the nature of the phonons [bulk, 1D, two-dimensional (2D), or 3D confined phonons].

Starting from the definition of the electron-phonon Hamiltonian in Eq. (6) and using functional derivatives[49–51] the general form of the two-time-dependent electron-phonon self-energy $\mathbf{\Sigma}^{e\text{-}ph}_{nm}(tt')$ in the self-consistent Born approximation is obtained as

$$\mathbf{\Sigma}^{e\text{-}ph}_{nm}(tt') = i\hbar \sum_{l_1 l_2} \sum_{\lambda,\mathbf{q}} \mathbf{M}^{\lambda}_{nl_1}(\mathbf{q}) \cdot \mathbf{G}_{l_1 l_2}(tt') \cdot \mathbf{M}^{\lambda *}_{l_2 m}(\mathbf{q}) \cdot \mathcal{D}^{\lambda}(\mathbf{q};tt'), \tag{9}$$

$$\mathbf{M}^{\lambda}_{nm}(\mathbf{q}) = \sqrt{\frac{\hbar}{2\omega_{\lambda}(\mathbf{q})}} \sum_i \nabla_i \mathbf{H}_{nm} \left[\frac{f^i_{\lambda}(\mathbf{R}_m,\mathbf{q})}{\sqrt{M_m}} - \frac{f^i_{\lambda}(\mathbf{R}_n,\mathbf{q})}{\sqrt{M_n}} \right]. \tag{10}$$

The phonon Green's function $\mathcal{D}^{\lambda}(\mathbf{q};tt')$ is defined as in Ref. 47. To derive Eq. (9) the property that $f^i_{\lambda}(\mathbf{R}_n,-\mathbf{q}) = f^{i*}_{\lambda}(\mathbf{R}_n,\mathbf{q})$ has been applied. Equation (9) is valid in 3D structures only where electrons are confined along two directions (y and z) and transport occurs along the third one (x), like nanowires. In this configuration, the Green's functions $\mathbf{G}_{nm}(tt')$ have no momentum dependence, in contrary to the phonons.

Langreth theorem,[41] the replacement of the phonon Green's function by its unperturbed definition at equilibrium[47] and the consideration of the steady-state regime allow one to find the energy-dependent self-energy $\mathbf{\Sigma}^{\gtrless,e\text{-}ph}_{nm}(E)$

$$\mathbf{\Sigma}^{\gtrless,e\text{-}ph}_{nm}(E) = \sum_{l_1 l_2} \sum_{\lambda,\mathbf{q}} \mathbf{M}^{\lambda}_{nl_1}(\mathbf{q}) \{ N^{\lambda}(\mathbf{q}) \mathbf{G}^{\gtrless}_{l_1 l_2}[E \pm \hbar\omega_{\lambda}(\mathbf{q})] + [N^{\lambda}(\mathbf{q})+1] \mathbf{G}^{\gtrless}_{l_1 l_2}[E \mp \hbar\omega_{\lambda}(\mathbf{q})] \} \mathbf{M}^{\lambda *}_{l_2 m}(\mathbf{q}), \tag{11}$$

where $N^{\lambda}(\mathbf{q})$ is the Bose distribution of the phonons with frequency $\omega_{\lambda}(\mathbf{q})$. In 3D simulation domains, Eqs. (1)–(4) and (11) cannot be directly solved and require further approximations. In a small nanowire structure with $N_A = 10\,000$ atoms, the total size of the Green's function and of the self-energy matrices is $N_A \times t_B = 100\,000$ in the $sp^3d^5s^*$ tight-binding model without spin-orbit coupling. Inverting, factorizing, and multiplying such full matrices for multiple energies goes beyond the capabilities of currently available supercomputers.

B. Simplifications and approximations

To investigate realistic nanowire structures with electron-phonon scattering some simplifications and/or approximations have to be considered in the calculation of the Green's functions, the self-energies, and the dispersion relation of phonons.

The dynamical matrix $\Phi^{ij}_{nm}(\mathbf{q})$ in Eq. (8) is based on a modified Keating model including four terms (bond-stretching, bond-bending, angle-angle, and bond-bond interactions).[52] This model uses four material constants that are optimized to accurately reproduce the bulk phonon dispersion of different semiconductors such as Si or Ge. The original bulk formalism has been first verified and then extended to calculate the confined phonon modes of nanowires, similarly to Refs. 53 and 54, automatically accounting for the acoustic and the optical phonon branches. For that purpose, the nanowire structure is assumed to be infinite and composed of identical unit cells. The 3D wave vector $\mathbf{q}$ reduces to its single x-component q aligned with the NW transport direction. As a second approximation, the environment of the nanowires (dielectric layers, substrate, and metallic gate) is omitted in the computation of the phonon bandstructure and eigenmodes. The surface atoms are therefore free to move. Discussions about the influence of the boundary conditions on the calculation of the phonon spectra can be found in Refs. 53–55.

To minimize the required memory to store the Green's functions and the self-energies as well as to reduce the computational burden, the $\mathbf{\Sigma}^{e\text{-}ph}_{nm}(E)$ are limited to on-site interactions only, i.e., $n=m$ but remain $t_B \times t_B$ blocks. There is no physical justification to this approximation but it appears to be the only way to numerically treat 3D structures such as nanowires composed of more than a dozen of atoms. Each additional connection to more distant neighbors (nearest, second-nearest, third-nearest, ...) induces an exponential increase in the time to compute the self-energy in Eq. (11). Furthermore, a large number of the off-diagonal blocks $\mathbf{G}_{nm}(E)$ must be computed if off-site self-energy interactions are taken into account. Efficient computational approaches such as the recursive Green's function (RGF) algorithm[56] fail if far off-diagonal blocks are required. The physical implications of full matrix and diagonal self-energy interactions have recently been studied in small 2D systems with respect to the influence on phase and momentum relaxation.[57] To compensate the missing self-energy terms, the magnitude of the matrix elements $\mathbf{M}^{\lambda}_{nm}(\mathbf{q})$ can be artificially increased.[58] This has not been tested in this study and it can only be speculated that scattering is underestimated.

To avoid the calculation of several off-diagonal blocks $\mathbf{G}_{nm}$ and to ensure current conservation in the diagonal self-energy approximation, Eq. (11) is simplified to

$$\mathbf{\Sigma}^{\gtrless,e\text{-}ph}_{nn}(E) = \sum_{l} \sum_{\lambda,q} \mathbf{M}^{\lambda}_{nl}(q) \{ N^{\lambda}(q) \mathbf{G}^{\gtrless}_{ll}[E \pm \hbar\omega_{\lambda}(q)] + [N^{\lambda}(q)+1] \mathbf{G}^{\gtrless}_{ll}[E \mp \hbar\omega_{\lambda}(q)] \} \mathbf{M}^{\lambda *}_{ln}(q). \tag{12}$$

The position index l runs over all the nearest neighbors of the atom located at $\mathbf{R}_n$, generally four atoms, except at the nanowire surface. A combination like $\mathbf{M}_{nl_1} \cdot \mathbf{G}_{l_1 l_2}(E) \cdot \mathbf{M}_{l_2 n}$, where l_1 and l_2 are two different nearest neighbors of n is technically possible but it would break current conservation if a self-energy of the form $\mathbf{\Sigma}_{l_1 l_2}(E)$ (off-site connection) does not exist. However, for computational reasons such self-energies are neglected as mentioned above.

While the electron-phonon interaction has been made local in space in Eq. (12), it remains nonlocal (inelastic) in energy, one energy E being coupled to many other energies $E' = E \pm \hbar\omega_{\lambda}(q)$. In a typical nanowire structure, the number of phonon modes λ exceeds 100 and the phonon 1D Brillouin zone (BZ) needs to be described by about 50 q

points. Hence, one energy point is connected to at least $2\times100\times50=10\,000$ other points [the factor 2 comes from the plus and minus sign in $\pm\hbar\omega_\lambda(q)$]. This results in several numerical problems. First, this number is larger than the total number of points in the energy grid ($N_E=1500-2000$). Second, since the energy points are distributed over many CPUs (see Appendix), too much interprocessor communication would be required to exchange massive Green's function data, which is not recommended in parallel computing. An approximate solution to this problem has been developed and implemented. If the sum over the phonon momentum vector q is replaced by an integration over the phonon 1D BZ of length L_{BZ}, the schematic form of the electron-phonon self-energy is given by

$$\begin{aligned}\boldsymbol{\Sigma}^{e\text{-}ph}_{nn}(E) &\propto \sum_{l,\lambda,q}\sum_{ij}\nabla_i\mathbf{H}_{nl}\cdot\mathbf{G}_{ll}[E\pm\hbar\omega_\lambda(q)]\cdot\nabla_j\mathbf{H}_{ln}\cdot\underbrace{\frac{\hbar}{2\omega_\lambda(q)}\cdot\left[\frac{f^i_\lambda(\mathbf{R}_l,q)}{\sqrt{M_l}}-\frac{f^i_\lambda(\mathbf{R}_n,q)}{\sqrt{M_n}}\right]\cdot\left[\frac{f^{j*}_\lambda(\mathbf{R}_n,q)}{\sqrt{M_n}}-\frac{f^{j*}_\lambda(\mathbf{R}_l,q)}{\sqrt{M_l}}\right]}_{\mathbf{V}^{ij}_{nlln}[\omega_\lambda(q)]}\\ &\propto \sum_{l,\lambda}\sum_{ij}\int_{\mathrm{BZ}}\frac{dq}{L_{\mathrm{BZ}}}\mathbf{V}^{ij}_{nlln}[\omega_\lambda(q)]\cdot\nabla_i\mathbf{H}_{nl}\cdot\mathbf{G}_{ll}[E\pm\hbar\omega_\lambda(q)]\cdot\nabla_j\mathbf{H}_{ln}.\end{aligned}\tag{13}$$

The coupling strength factor $\mathbf{V}^{ij}_{nlln}[\omega_\lambda(q)]$ has been introduced to clarify the notation but the multiple energy connections remain. To reduce them, it is assumed that the lesser and greater Green's functions $\mathbf{G}^{\gtrless}_{ll}(E)$ slowly vary over a small energy range $E-\Delta_E\leq E\leq E+\Delta_E$ in the presence of electron-phonon scattering. This is the case for most of the energy points. Then the sum over λ and the integration over q are replaced by a sum over the phonon energy ω_{PH}

$$\begin{aligned}\boldsymbol{\Sigma}^{\gtrless e\text{-}ph}_{nn}(E) = \sum_{l,i,j}\sum_{\omega_{PH}}&\mathcal{V}^{ij}_{nlln}(\omega_{PH})\cdot\nabla_i\mathbf{H}_{nl}\\ &\cdot\{N^{PH}(\omega_{PH})\cdot\mathbf{G}^{\gtrless}_{ll}(E\pm\hbar\omega_{PH})+[N^{PH}(\omega_{PH})\\ &+1]\cdot\mathbf{G}^{\gtrless}_{ll}(E\mp\hbar\omega_{PH})\}\cdot\nabla_j\mathbf{H}_{ln},\end{aligned}\tag{14}$$

$$\begin{aligned}&\mathcal{V}^{ij}_{nlln}(\omega_{PH})\\ &\quad=\sum_\lambda\int_{\mathrm{BZ}}\frac{dq}{L_{\mathrm{BZ}}}\mathbf{V}^{ij}_{nlln}[\omega_\lambda(q)]\big|_{\omega_{PH}-\Delta_E/\hbar\leq\omega_\lambda(q)\leq\omega_{PH}+\Delta_E/\hbar}.\end{aligned}\tag{15}$$

The sum over λ and q does not completely disappear but it is moved into the coupling factor $\mathcal{V}^{ij}_{nlln}(\omega_{PH})$ that is precomputed at the beginning of any simulation. It has been numerically verified that a value $\Delta_E=1$ meV is small enough to ensure accurate results and that increasing the number of ω_{PH} points, i.e., decreasing the value of Δ_E below 1 meV, does not significantly change the results any more but slows down the computation time. Values of Δ_E comprised between 0.25 and 2 meV have been tested resulting in a variation in the device current by less than 5%. In Si the largest phonon energy amounts to approximately 60 meV so that the sum over ω_{PH} can be restricted to 2×30 points instead of 10 000 if $\Delta_E=1$ meV is assumed.

In Eq. (3), the retarded self-energy $\boldsymbol{\Sigma}^R_{nn}(E)$ is composed of two parts, the second one requiring the evaluation of a Cauchy principal integral coupling all the energies together. For the same computational reasons as above, this integral term cannot be calculated and it is therefore neglected in this paper. Previous studies have showed that this simplification does not introduce significant errors in the calculation of the device current.[28,59] As a future improvement, the retarded self-energy $\boldsymbol{\Sigma}^R_{nn}(E)$ will be derived directly from Eq. (9) following the same procedure as for the lesser and greater $\boldsymbol{\Sigma}^{\gtrless}_{nn}(E)$ terms.

Finally, the nearest-neighbor matrix elements of the Slater-Koster table[44] are derived to obtain the $\nabla_i\mathbf{H}_{nm}$ in Eq. (14). Only the relative angle dependence of each atom with respect to its neighbors (the l, m, and n directional cosines) is considered to calculate $\nabla_i\mathbf{H}_{nm}$, not the distance (bond-length) dependence of the matrix elements. The latter requires that the matrix elements $\mathbf{H}_{nm}$ are scaled by a factor $(d_0/d)^\eta$, where $d_0(d)$ is the ideal (altered) bond length.[45] If the bond-length dependence is included, the $\nabla_i\mathbf{H}_{nm}$ terms increase, as the electron-phonon interaction, leading to a decrease in the device current. Despite the overestimation of the device current the bond-length dependence of $\mathbf{H}_{nm}$ is neglected since the η values for Si do not appear to be unequivocally defined in the literature.[45,60,61]

C. Convergence criteria

Equations (1)–(4) and (14) are self-consistently solved till convergence is reached. A parallel RGF algorithm (up to 4 CPUs) has been developed to accomplish this task as explained in Appendix. Two criteria are used to check the convergence of the self-consistent Born treatment of the Green's functions and self-energies. The successive iterations of Eqs. (1)–(4) and (14) are stopped when the carrier and the current densities, calculated as in Refs. 62 and 63 vary within an error of 1‰ and 1%, respectively.

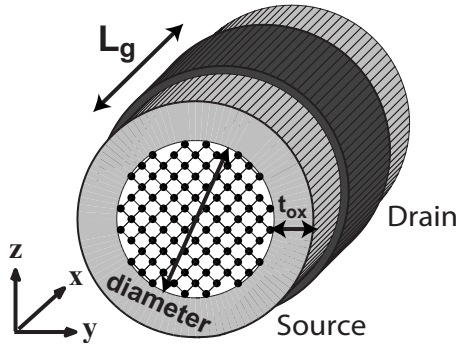

FIG. 1. Schematic view of a GAA circular Si NW FET with a diameter $d=2$, 2.5, and 3 nm. The transport direction x is aligned with the $\langle 100 \rangle$ crystal axis. The SiO_2 dielectric layer has a thickness $t_{ox}=1$ nm. The gate length L_g measures 15 nm while the source and drain contacts have a length of 12.5 nm and a donor doping concentration $N_D=1e20$ cm^{-3}.

Especially, in a two-terminal device (source and drain), current conservation implies that the current that enters the simulation domain at a port B_1 is the same as the current that leaves the structure at the other port B_2

$$I_d(B) = \frac{e}{\hbar} \sum_{\text{spin}} \sum_{nm} \int \frac{dE}{2\pi} \text{tr}[\mathbf{\Sigma}^{>B}_{nm}(E)\mathbf{G}^{<}_{mn}(E) - \mathbf{G}^{>}_{nm}(E)\mathbf{\Sigma}^{<B}_{mn}(E)]. \tag{16}$$

The current $I_d(B)$ is the current at the terminal $B \in (\text{source}, \text{drain})$, the position indices n and m run over all the atoms directly connected to the semi-infinite device contact B while $\mathbf{\Sigma}^{>B}_{nm}$ and $\mathbf{\Sigma}^{<B}_{nm}$ are the boundary self-energies at the port B. The trace operator "tr" runs over the different orbitals of the atoms n and m. The convergence criterion for the current becomes then $I_d(\text{drain})=I_d(\text{source})$. As a part of data postprocessing, it can be checked that indeed current is conserved throughout the device from one layer to the other.

Surrounding the self-consistent Born iterations between the Green's functions and the self-energies, there is the self-consistent coupling of the charge density and of the device electrostatic potential through Poisson equation.[40] Generally, the carrier density converges much faster than the current density (about 5–10 self-consistent Born iterations vs 20–50 for the current). Consequently, as long as the electrostatic potential is not stable, only the convergence of the charge density is fulfilled. Once that Poisson equation has converged, the criterion for the current density is satisfied too.

III. RESULTS

The structure of the Si GAA nanowire field-effect transistors considered in this work is depicted in Fig. 1. The diameter of the NWs is varied from 2 to 3 nm, the gate length L_g is set to 15 nm while the n-doped ($N_D=1e20$ cm^{-3}) source and drain extensions measure 12.5 nm, except for the nanowire with $d=2.5$ nm, where they are extended to 15 nm. Hence, the structure with $d=2$ nm is composed of 6497 atoms that with $d=2.5$ nm of 11 371 atoms and finally that with $d=3$ nm of 14 089 atoms. In the $sp^3d^5s^*$ tight-binding formalism without spin-orbit coupling, the $d=3$ nm NW is represented by an Hamiltonian matrix of size $N=140\,890$.

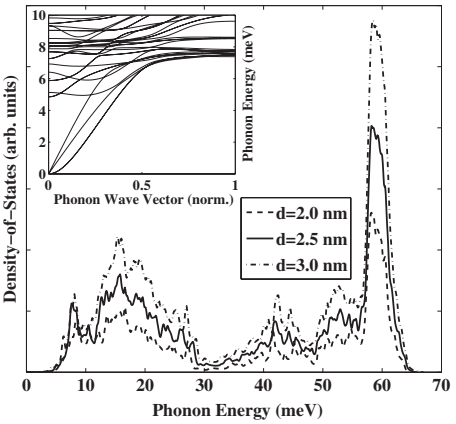

FIG. 2. Density of states of the confined phonons in a Si nanowire with a diameter $d=2.0$ nm (dashed line), $d=2.5$ nm (solid line), and $d=3.0$ nm (dashed-dotted line). The inset shows the lowest part of the phonon bandstructure in the nanowire with $d=2.5$ nm.

The drain current I_d flows along the x direction of the NW FETs, which is aligned with the $\langle 100 \rangle$ crystal axis, y and z are directions of confinement. The nanowire channel is isolated from the coaxial metal gate contact (work function $\phi_M=4.05$ eV) by a SiO_2 oxide layer of thickness $t_{ox}=1$ nm, characterized by a relative dielectric constant $\epsilon_R=3.9$, and assumed ideal (infinite band gap). The mechanical coupling between the nanowires and their environment is not taken into account in the simulations so that the transistors can be seen as free standing. All the simulations are performed at room temperature.

The density of states (DOS) $g(E)$ of the phonon modes confined in the free-standing Si nanowire with $d=2.5$ nm is reported in Fig. 2. The inset represents the lowest acoustic bands of the phonon dispersion $E^{PH}(q)$ used to calculate $g(E)$

$$g(E) = \frac{1}{V_{uc}} \sum_{q} \delta[E - E^{PH}(q)], \tag{17}$$

where V_{uc} is the volume of a nanowire unit cell. The phonon DOS exhibits several peaks, especially around $E=15-16$ meV and around $E=58-59$ meV which is the energy of the optical phonons in bulk Si. This means that the electron-phonon interaction around these energies is stronger and the probability for an electron or hole to absorb or emit a phonon with these energies is higher.

The transfer characteristics I_d-V_{gs} at $V_{ds}=0.6$ V of Si GAA NW FETs with diameters $d=2$, 2.5, and 3 nm are shown in Figs. 3–5. The dashed gray lines with circles refer to ballistic transport simulations, the black lines with squares to simulations in the presence of electron-phonon scattering. The drain currents I_d are given on a logarithmic (left axis) and on a linear (right axis) scale.

The current reduction due to electron-phonon scattering is reported in Fig. 6 for the three nanowire FETs as a function of V_{gs} and at $V_{ds}=0.6$. It is defined as

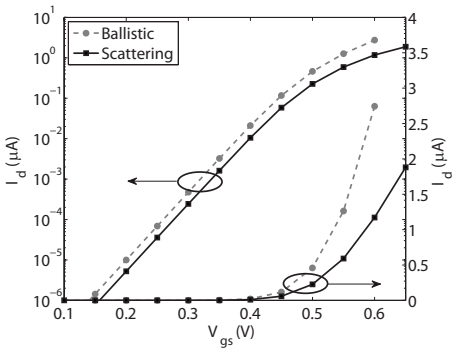

FIG. 3. Transfer characteristics I_d-V_{gs} at $V_{ds}=0.6$ V of a Si GAA NW FET with a diameter $d=2$ nm. The gray lines with circles refer to the ballistic current and the black lines with squares to the current with electron-phonon scattering.

$(I_{ball}-I_{scatt})/I_{ball}\times 100$, where I_{ball} and I_{scatt} are the ballistic and the dissipative currents, respectively. The arrows indicate the position of the transistor threshold voltage in the ballistic simulations. As shown in Figs. 3–5, the threshold voltage of the nanowire devices depends on the transport model, ballistic or dissipative, resulting in a short saturation region of the current reduction starting at the arrow positions.

In Fig. 7 the output characteristics I_d-V_{ds} of the nanowire with $d=2$ nm with (solid line with squares) and without (dashed line with circles) electron-phonon scattering are compared at $V_{gs}=0.5$ V. The current reduction due to scattering $(I_{ball}-I_{scatt})/I_{ball}\times 100$ is also given on the right y axis. It varies by less than 4% from $V_{ds}=0.05$ V to $V_{ds}=0.6$ V. From Figs. 3–7, three common characteristics shared by all the device are identified: (i) the ballistic current is larger than the current with scattering, (ii) the difference between the two currents becomes larger at high V_{gs}, and (iii) the inclusion of scattering helps the convergence of the Poisson equation up to higher V_{gs}. Each of these points is detailed below.

Electron-phonon scattering modifies the electrostatic potential of the NW FETs as illustrated in Fig. 8. In the source extension $(0\leq x\leq 15)$, the total number of electrons n_{tot} is composed of forward (n_F) and backward (n_B) moving charges. The forward moving charges are those injected at the source contact while the backward moving charges are

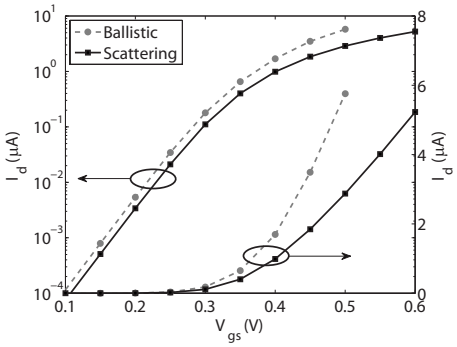

FIG. 4. Same as Fig. 3 but for a Si nanowire with a diameter $d=2.5$ nm.

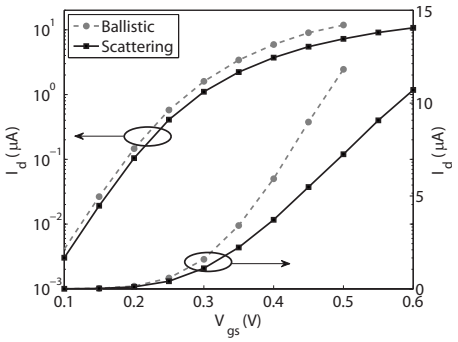

FIG. 5. Same as Figs. 3 and 4 but for a Si nanowire with a diameter $d=3$ nm.

those reflected back to the source contact or coming from the drain contact. At low gate biases $(V_{gs})n_F$ and n_B have almost the same magnitude, the doping concentration N_D divided by 2, most of the injected electrons being reflected at the source-to-drain potential barrier $(\rightarrow n_B=n_F)$, and charge neutrality being imposed by Poisson equation $(\rightarrow n_{tot}=n_F+n_B=N_D)$. The fundamental difference between the ballistic current and the current with scattering comes from the possibility for electrons flowing over the source-to-drain potential barrier to interact with a phonon and to be reflected back to the source. This effect reduces the current magnitude but it is no more important if it occurs after a distance l measured from the top of the potential barrier and called critical length for backscattering.[64] After this distance an electron interacting with a phonon has a very low probability to flow back to the source contact.

At $V_{gs}=0.1$ V, the ballistic current $(I_{ballistic})$ exceeds the current with scattering $(I_{scattering})$ by a factor of 1.86 for a NW diameter with $d=2$ nm, 1.52 for $d=2.5$ nm, and 1.35 for $d=3$ nm. At $V_{gs}=0.5$ V the difference between the two currents increases to 1.98($d=2$ nm), 2.0($d=2.5$ nm), and 1.63($d=3$ nm). The fact that electron phonon plays a more important role in NWs with a small diameter tends to confirm the observation made in Ref. 10.

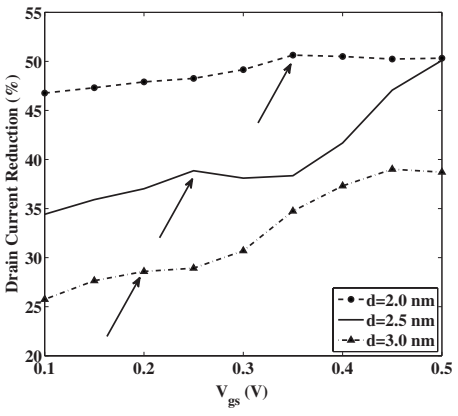

FIG. 6. Drain current reduction due to electron-phonon scattering at $V_{ds}=0.6$ V as a function of V_{gs} in Si nanowires with a diameter $d=2.0$ nm (dashed line with circles), $d=2.5$ nm (solid line), and $d=3.0$ nm (dashed-dotted line with triangles). The ballistic I_{ball} and dissipative I_{scatt} currents of Figs. 3–5 are used to calculate $(I_{ball}-I_{scatt})/I_{ball}\times 100$. The arrows indicate the threshold voltage in the ballistic simulations.

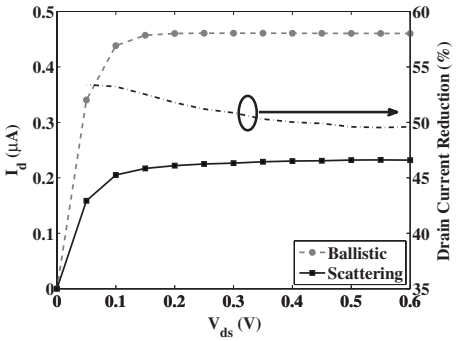

FIG. 7. Output characteristics I_d-V_{ds} at $V_{gs}=0.5$ V (left y axis) of the Si GAA NW FET with a diameter $d=2$ nm. The gray lines with circles refer to the ballistic current and the black lines with squares to the current with electron-phonon scattering. The drain current reduction due to scattering is plotted on the right y axis. It is defined as in Fig. 6.

The behavior of the electrostatic potential in the unrealistic ballistic transport regime is the main reason behind the current increase close to the transistor ON state ($V_{ds}=V_{gs}=0.6$ V). The lowering of the source-to-drain potential barrier, as observed in Fig. 8 at $V_{gs}=0.5$ V, reduces the amount of reflected electrons n_B to almost 0. To maintain charge neutrality in the source contact, the number of injected electrons n_F is artificially increased to the value of the donor concentration N_D by pushing down the conduction-band (CB) edge but the total number of electrons in the source $n_{tot}=n_F+n_B$ remains the same and is equal to N_D as illustrated in Fig. 9.

Since the ballistic current is directly proportional to the number of injected electrons at the source, it increases with n_F. When electron phonon is switched on, some of the electrons injected at the source contact are reflected back when they absorb or emit a phonon so that n_B does not completely vanish. It remains smaller than n_F but the conduction-band edge does not need to move down as deep as in the ballistic

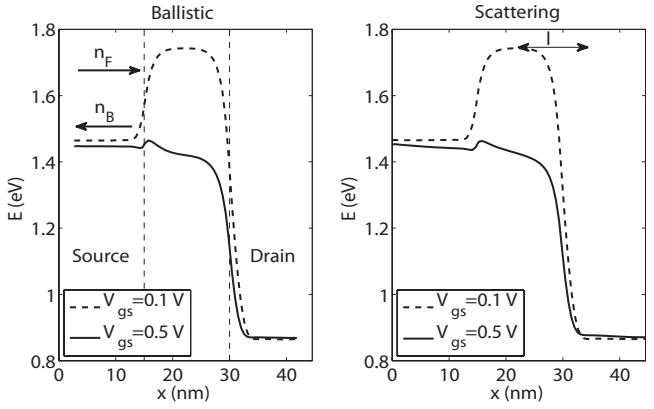

FIG. 8. Conduction-band edges at $V_{ds}=0.6$ V, $V_{gs}=0.1$ V (dashed lines) and $V_{gs}=0.5$ V (solid lines) in the nanowire with $d=2.5$ nm. The band edges in case of ballistic transport are shown in the left subplot, those with electron-phonon scattering in the right subplot. The variable $n_F(n_B)$ describes the electrons moving forward (backward) to (from) the gate and l is the critical length for backscattering (Ref. 64).

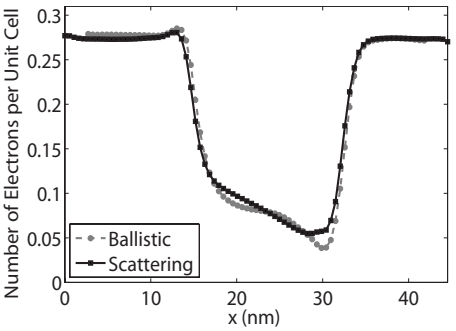

FIG. 9. Number of electrons per unit cell (ensemble of four atomic layers in the transport direction x) for the Si GAA NW FET with a diameter $d=2.5$ nm at $V_{ds}=0.6$ V and $V_{gs}=0.5$ V. The gray line with circles depicts the ballistic electron density and the black line with squares the carrier density in presence of electron-phonon scattering. The number of doping atoms per unit cell in the source and drain contacts is 0.274.

simulations. The conjunction of the backscattering effect and of the conduction-band shift explains the larger difference between the ballistic and scattering currents at high V_{gs}.

To determine the ballisticity $B=I_{scattering}/I_{ballistic}$ of the Si nanowire FETs, the effect of the artificial lowering of the conduction-band edge must be removed from $I_{ballistic}$. This is achieved by recalculating the ballistic current using the electrostatic potential obtained in the presence of electron-phonon scattering. Then, the only difference between $I_{ballistic}$ and $I_{scattering}$ arises from the backscattering mechanism[64] and not from unphysical artifacts. In doing so, at $V_{gs}=0.5$ V, $I_{ballistic}$ becomes larger than $I_{scattering}$ by a factor of 2.0 for the NW with $d=2$ nm, 1.56 for $d=2.5$ nm, and 1.42 for $d=3$ nm, which is comparable to the results obtained at $V_{gs}=0.1$ V (1.86, 1.52, and 1.35) and smaller than the values obtained in the pure ballistic simulations with the lowering of the CB edge (2.13, 2.0, and 1.63). Hence, one obtains $B_{2\ \mathrm{nm}}=50\%$, $B_{2.5\ \mathrm{nm}}=64\%$, and $B_{3\ \mathrm{nm}}=70\%$. The inclusion of interface roughness at the Si-SiO_2 interface will probably reduce the ballisticity of the FETs.[65]

For gate voltage V_{gs} beyond 0.5 V (0.6 V for the nanowire with $d=2$ nm), the Poisson equation does not converge anymore in the ballistic simulations. The mechanism that pushes down the conduction band, as explained above, becomes very sensitive to a small variation in the electrostatic potential which in turn starts to oscillate in the source extension region. The inclusion of dissipative scattering resolves this issue. At high V_{gs}, the transistor acts as a resistor, the electrostatic potential drops in the source and drain regions, as seen in Fig. 8, and the emission of phonons allows the electrons to fill energy states located below the conduction-band edge of the semi-infinite source contact. These states are inaccessible in the absence of dissipative scattering. Hence, the electrostatic potential remains stable and the self-consistent calculation of the electron density and of Poisson equation continues to converge.

Figure 10 shows the ballistic and scattering currents at $V_{gs}=0.5$ V and $V_{ds}=0.6$ V as a function of the position and energy in the NW FET with $d=2.5$ nm. The ballistic current is conserved for each injection energy E, as expected, while

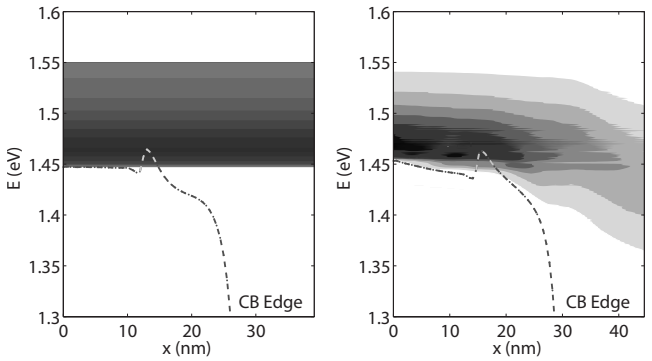

FIG. 10. Spectral current of the Si GAA NW FET with a diameter $d=2.5$ nm from Fig. 4 at $V_{gs}=0.5$ V and $V_{ds}=0.6$ V. The left subplot represents the ballistic current (homogeneous) and the right subplot the current with electron-phonon scattering. The dashed lines indicate the position of the CB edge.

the current with electron-phonon scattering follows the contour of the electrostatic potential, it flows at higher energy in the source than in the drain region due to phonon emission and only the sum of the current contributions from all the energies is conserved. Current conservation in the presence of electron-phonon scattering is demonstrated in Fig. 11.

The ON current (I_d at $V_{gs}=V_{ds}=0.6$ V) of the NW FET with $d=2.5$ nm is given in Fig. 12. The potential drop in the source region is larger than at $V_{gs}=0.5$ V in Fig. 10 and regions with a high current concentration (indicated by black arrows) separated by regions with a lower current concentration become visible. The energy separation between these regions is about 15–16 meV, corresponding to one of the peaks of the phonon density of states shown in Fig. 2. The phonon DOS is larger at 58–59 meV than at 15–16 meV but since the electron-phonon coupling factor $\mathcal{V}^{ij}_{nlln}(\omega_{PH})$ in Eq. (14) is inversely proportional to the phonon frequency ω_{PH}, the strength of the electron-phonon interaction at $E=15-16$ meV is almost four times larger than at 58–59 meV. This explains the energy separation of the regions with a high current concentration.

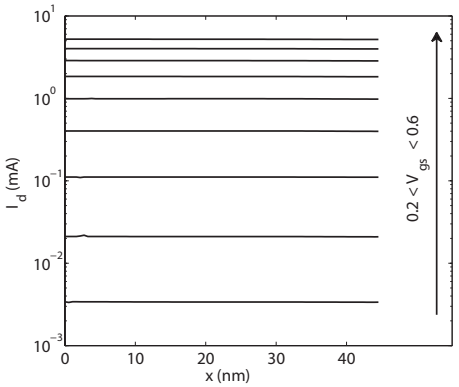

FIG. 11. Spatial current along the transport direction x of the Si GAA NW FET with a diameter $d=2.5$ nm and in the presence of electron-phonon scattering. The currents are calculated at $V_{ds}=0.6$ V and $0.2 \leq V_{gs} \leq 0.6$ V.

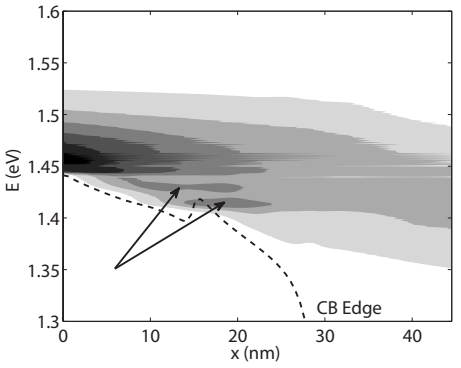

FIG. 12. Spectral current of the Si GAA NW FET with a diameter $d=2.5$ and electron-phonon scattering at $V_{gs}=V_{ds}=0.6$ V. The arrows indicate two regions with high current densities, separated by an energy difference of 15–16 meV.

IV. CONCLUSION AND OUTLOOK

Electron-phonon scattering has been demonstrated in a 3D, atomistic, full band, device simulator based on the nearest-neighbor $sp^3d^5s^*$ tight-binding model and the nonequilibrium Green's function formalism. Nanowire field-effect transistors with a diameter up to 3 nm, a length of 40 nm, and composed of more than 14 000 atoms have been simulated. The influence of electron-phonon scattering in the OFF and ON states of Si NW transistors is investigated. The overestimation of the drain current by ballistic transport simulations is more pronounced in the ON state due to an increase in the electrons that are injected at the source contact.

To further extend the dimensions of the 3D simulation domains, the numerical algorithm must be improved. The current parallelization of the RGF algorithm does not allow to distribute the work load over a large enough number of CPUs. More efficient approaches have been proposed by other groups[66,67] and will be tested in the future.

The capability of treating electron-phonon scattering will be of great interest to study certain types of devices such as band-to-band tunneling field-effect transistors (TFETs) which are based on the tunneling of electrons from the valence band of one contact into the conduction band of the other contact. Silicon might be the material of choice to fabricate such devices because it is compatible with the MOSFET technology and its properties are well understood. However, the indirect band gap of Si prevents direct tunneling to occur between its valence band and its conduction band. Momentum conservation requires that a phonon is absorbed or emitted during the tunneling process. At the same time, modeling TFETs demands a full-band approach to accurately describe the valence and conduction bands of the semiconductor material.

ACKNOWLEDGMENTS

The authors would like to thank Timothy Boykin for helpful discussions. This work was partially supported by NSF under Grant No. EEC-0228390 that funds the Network for Computational Nanotechnology, by NSF PetaApps under

Grant No. 0749140, by the Nanoelectronics Research Initiative through the Midwest Institute for Nanoelectronics Discovery, and by NSF through TeraGrid resources provided by the National Institute for Computational Sciences (NICS). This research also used resources of the National Center for Computational Sciences (NCCS) at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

APPENDIX: NUMERICAL IMPLEMENTATION

Parallel computing is the key element to solve Eqs. (1)–(4) and (14) in 3D nanowire structures. It would simply not be possible to simulate the devices described in Sec. III on a single processor, not even on a small cluster of 100 cores or less. All the results presented in this work were obtained on 3200 to 6400 cores and each simulation lasted between 36 and 100 h, depending on the device dimensions.

Due to the energy coupling no embarrassingly parallelization of Eq. (14) is possible. To keep an efficient distribution of the energy points (ideally, there are $N_{E,local}=2$ energy points per CPU) the calculation of Eqs. (1)–(4) and (14) is separated in two steps: (1) for each new self-consistent Born iteration, all the $\mathbf{\Sigma}_{nn}^{<e\text{-}ph}(E)$ and $\mathbf{\Sigma}_{nn}^{>e\text{-}ph}(E)$ are first computed from Eq. (14). This requires a lot of communication since each energy point E needs about $2\times 30\times N_A$ $\mathbf{G}_{nn}^{\gtrless}(E')$ blocks from the previous self-consistent Born iteration. The factor 2 originates from the connection to higher and lower energies E', 30 is the number of different phonon energies that are typically considered while N_A is the number of atoms in the device structure. The N_A blocks $\mathbf{G}_{nn}^{\gtrless}(E')$ of size $t_B\times t_B$ are cast into a single variable $\mathcal{G}^{\gtrless}(E')$ of size $N_A t_B\times t_B$ and are not individually transferred from one CPU to the other but in one single step. Each of the 60 $\mathcal{G}^{\gtrless}(E')$ matrices are usually stored on 60 different CPUs. In other words, each core must send its Green's functions to 60 different CPUs and receive data from 60 others. The most efficient solution to manage the communication issues consists in using the nonblocking MPI_Isend function to send the data and MPI_Receive to collect them.[68] For a given phonon energy ω_{PH}, each CPU sends its $N_{E,local}\mathcal{G}^{\gtrless}(E)$ matrices to the CPU(s) containing the energies $E\pm\hbar\omega_{PH}$. After all the CPUs have sent their data, they can at any time receive the matrices $\mathcal{G}^{\gtrless}(E\mp\hbar\omega_{PH})$ they need to evaluate Eq. (14) for the $N_{E,local}$ energy points they take care of. (2) Once that each CPU has solved Eq. (14), Eqs. (1)–(4) can be computed in an almost embarrassingly parallel way, with limited interprocessor communication, as described in Ref. 40.

For a typical energy grid of $N_E=1500-2000$ energy points, a total number of 750–1000 CPUs can be used per simulation, assuming that Eqs. (1)–(4) are treated on a single core using a RGF approach.[56] However, the RGF algorithm can be slightly modified to efficiently parallelize on two CPUs. In its standard form,[56] the recursive calculation of the Green's functions starts at one corner of the Hamiltonian matrix, goes on till the other corner of the matrix, and moves back to the original corner. An alternative is to start the calculation from both corners of the matrix, each of them being handled by a different CPU, to move toward the center of the matrix, to exchange information there through MPI_Send and MPI_Receive, and to go back to the two matrix corners. A speed-up factor close to 2 can be obtained following this approach. Two additional cores can be used in Eqs. (1)–(4), one of them calculating $\mathbf{G}_{nn}^{R}(E)$ and $\mathbf{G}_{nn}^{<}(E)$, the other $\mathbf{G}_{nn}^{R}(E)$ and $\mathbf{G}_{nn}^{>}(E)$. This is not an ideal parallelization since $\mathbf{G}_{nn}^{R}(E)$ is calculated twice but it allows a speed up of 1.5. In summary, the domain decomposition of the Hamiltonian matrix and the separation of $\mathbf{G}_{nn}^{<}(E)$ and $\mathbf{G}_{nn}^{>}(E)$ lead to a speed-up factor of 3 on 4 cores as compared to one single CPU.

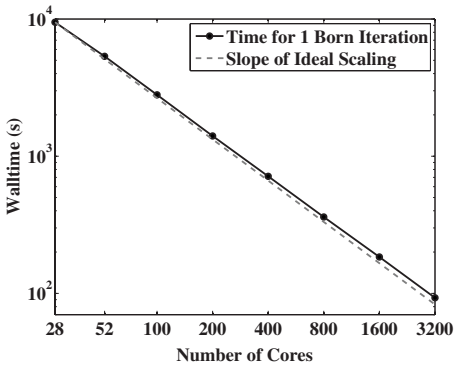

FIG. 13. Parallel execution time on a CRAY XT4 for the calculation of one Born iteration, i.e., for the solution of Eqs. (1)–(4) and (14) for all the energies E on 28–3200 cores. The test structure is a nanowire with $n=2$ nm and $N_E=1571$ energy points. Each energy point is treated by four CPUs. Typically, 20–50 Born iterations are required to obtain the device current.

Hence, a total of 3000–4000 CPUs can be used per bias point for an energy grid of $N_E=1500-2000$ energy points. The scaling performance of our approach for $N_E=1571$ points is demonstrated in Fig. 13. It is shown that the simulation time for one Born iteration [solution of Eqs. (1)–(4) and (14) for all the energy points] in a nanowire with a diameter of 2 nm decreases almost ideally from 28 to 3200 cores on a CRAY XT4 machine.[69] The number of CPUs per energy point is set to 4 [domain decomposition and separation of $\mathbf{G}_{nn}^{<}(E)$ and $\mathbf{G}_{nn}^{>}(E)$] for all the timing experiments reported in Fig. 13. This means that seven energy points are simultaneously treated when the total number of CPUs is equal to 28 and 800 when the total number of CPUs is 3200. Furthermore, several bias points can be simulated at the same time in a embarrassingly parallel way. In this paper, a maximum a 2 bias points has been simultaneously treated, leading to a total number of 6400 cores per simulation.

The scaling behavior of the simulation approach does not depend on the nanowire size or cross section and remains almost ideal for all cases. However, the simulation time increases as a function of the number of atoms in the device structure. The time to solve Eqs. (1)–(4) (standard RGF approach) increases as $O(N_{CS}^3)$, where N_{CS} is the number of atoms per nanowire unit cell[21] and as $O(N_{UC})$, where N_{UC} is the number of unit cells along the nanowire length. The time to compute the scattering self-energy in Eq. (14) linearly increases as a function of the total number of atoms N_A since only diagonal interactions are considered.

[1] International Technology Roadmap for Semiconductor 2007, http://www.itrs.net
[2] Y. Cui, L. J. Lauhon, M. S. Gudiksen, J. Wang, and C. M. Lieber, Appl. Phys. Lett. **78**, 2214 (2001).
[3] A. I. Persson, M. W. Larsson, S. Steinström, B. J. Ohlsson, L. Samuelson, and L. R. Wallenberg, Nature Mater. **3**, 677 (2004).
[4] A. B. Greytak, L. J. Lauhon, M. S. Gudiksen, and C. M. Lieber, Appl. Phys. Lett. **84**, 4176 (2004).
[5] T. Saito, T. Saraya, T. Inukai, H. Majima, T. Nagumo, and T. Hiramoto, IEICE Trans. Electron. **E85-C**, 1073 (2002).
[6] H. Majima, Y. Saito, and T. Hiramoto, Tech. Dig. - Int. Electron Devices Meet. **2001**, 733.
[7] Y. Cui, Z. Zhong, D. Wang, J. Wang, and C. M. Lieber, Nano Lett. **3**, 149 (2003).
[8] Y. Hayamizu, M. Yoshita, S. Watanabe, H. Akiyama, L. N. Pfeiffer, and K. W. West, Appl. Phys. Lett. **81**, 4937 (2002).
[9] N. Singh, A. Agarwal, L. K. Bera, T. Y. Liow, R. Yang, S. C. Rustagi, C. H. Tung, R. Kumar, G. Q. Lo, N. Balasubramanian, and D.-L. Kwong, IEEE Electron Device Lett. **27**, 383 (2006).
[10] S. D. Suk, M. Li, Y. Y. Yeoh, K. H. Yeo, K. H. Cho, I. K. Ku, H. Cho, W. Jang, D.-W. Kim, D. Park, and W.-S. Lee, Tech. Dig. - Int. Electron Devices Meet. **2007**, 891.
[11] R. E. Bank, D. J. Rose, and W. Fichtner, IEEE Trans. Electron Devices **30**, 1031 (1983).
[12] C. Jacoboni and L. Reggiani, Rev. Mod. Phys. **55**, 645 (1983).
[13] E. Polizzi and N. Ben Abdallah, J. Comput. Phys. **202**, 150 (2005).
[14] J. Wang, E. Polizzi, and M. S. Lundstrom, J. Appl. Phys. **96**, 2192 (2004).
[15] A. Martinez, M. Bescond, J. R. Barker, A. Svizhenko, M. P. Anantram, C. Millar, and A. Asenov, IEEE Trans. Electron Devices **54**, 2213 (2007).
[16] J. Wang, A. Rahman, A. Ghosh, G. Klimeck, and M. S. Lundstrom, IEEE Trans. Electron Devices **52**, 1589 (2005).
[17] A. Svizhenko, P. W. Leu, and K. Cho, Phys. Rev. B **75**, 125417 (2007).
[18] A. Lherbier, M. P. Persson, Y. M. Niquet, F. Triozon, and S. Roche, Phys. Rev. B **77**, 085301 (2008).
[19] A. Pecchia, L. Salamandra, L. Latessa, B. Aradi, T. Frauenheim, and A. Di Carlo, IEEE Trans. Electron Devices **54**, 3159 (2007).
[20] M. Luisier and G. Klimeck, IEEE Electron Device Lett. **30**, 602 (2009).
[21] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, Phys. Rev. B **74**, 205323 (2006).
[22] T. B. Boykin, M. Luisier, and G. Klimeck, Phys. Rev. B **77**, 165318 (2008).
[23] J. Schwinger, J. Math. Phys. **2**, 407 (1961).
[24] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (W. A. Benjamin, New York, 1962).
[25] L. V. Keldysh, Sov. Phys. JETP **20**, 1018 (1965).
[26] S. Datta, J. Phys.: Condens. Matter **2**, 8023 (1990).
[27] S. Jin, Y. J. Park, and H. S. Min, J. Appl. Phys. **99**, 123719 (2006).
[28] M. Frey, A. Esposito, and A. Schenk, Proceedings of the 38th European Solid-State Device Research Conference (ESSDERC), 2008 (unpublished), p. 258.
[29] A. Svizhenko and M. P. Anantram, Phys. Rev. B **72**, 085430 (2005).
[30] S. O. Koswatta, S. Hasan, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov, IEEE Trans. Electron Devices **54**, 2339 (2007).
[31] M. Pourfath, H. Kosina, and S. Selberherr, J. Comput. Electron. **6**, 321 (2007).
[32] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, J. Appl. Phys. **81**, 7845 (1997).
[33] C. Rivas, R. Lake, W. R. Frensley, G. Klimeck, P. E. Thompson, K. D. Hobart, S. L. Rommel, and P. R. Berger, J. Appl. Phys. **94**, 5005 (2003).
[34] U. Aeberhard and R. H. Morf, Phys. Rev. B **77**, 125343 (2008).
[35] T. Frederiksen, M. Paulsson, M. Brandbyge, and A.-P. Jauho, Phys. Rev. B **75**, 205413 (2007).
[36] A. Pecchia, G. Romano, and A. Di Carlo, Phys. Rev. B **75**, 035401 (2007).
[37] Y. Asai, Phys. Rev. B **78**, 045434 (2008).
[38] M. V. Fischetti, Phys. Rev. B **59**, 4901 (1999).
[39] K.-M. Liu, W. Chen, L. F. Register, and S. K. Banerjee, J. Appl. Phys. **104**, 114515 (2008).
[40] M. Luisier and G. Klimeck, in *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing*, Article 12 (IEEE Press, Piscataway, NJ, 2008).
[41] H. Haug and A. P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, 1989).
[42] P. Danielewicz, Ann. Phys. (N.Y.) **152**, 239 (1984).
[43] T. B. Boykin, G. Klimeck, and F. Oyafuso, Phys. Rev. B **69**, 115201 (2004).
[44] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[45] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B **57**, 6493 (1998).
[46] A. Wacker, Phys. Rep. **357**, 1 (2002).
[47] G. D. Mahan, *Many-Particle Physics* (Plenum, New York, 1990).
[48] N. Ashcroft and N. Mermin, *Solid State Physics* (Rinehart and Winston, New York, 1976).
[49] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
[50] R. Binder and S. W. Koch, Prog. Quantum Electron. **19**, 307 (1995).
[51] W. Schäfer and M. Wegener, *Semiconductor Optics and Transport Phenomena* (Springer, Berlin, 2002).
[52] Z. Sui and I. P. Herman, Phys. Rev. B **48**, 17938 (1993).
[53] T. Thonhauser and G. D. Mahan, Phys. Rev. B **69**, 075213 (2004).
[54] A. Valentin, J. See, S. Galdin-Retailleau, and P. Dollfus, J. Phys.: Condens. Matter **20**, 145213 (2008).
[55] A. K. Buin, A. Verma, and M. P. Anantram, J. Appl. Phys. **104**, 053716 (2008).
[56] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, J. Appl. Phys. **91**, 2343 (2002).
[57] R. Golizadeh-Mojarad and S. Datta, Phys. Rev. B **75**, 081301(R) (2007).
[58] G. Klimeck, R. Lake, C. L. Fernando, R. Chris Bowen, D. Blanks, M. Leng, T. Moise, Y. C. Kao, and W. R. Frensley, in *Quantum Devices and Circuits*, edited by K. Ismail, S. Bandyopadhyay, and J. P. Leburton (Imperial, London, 1996).
[59] A. Svizhenko and M. P. Anantram, IEEE Trans. Electron Devices **50**, 1459 (2003).
[60] T. B. Boykin, N. Kharche, and G. Klimeck, Phys. Rev. B **76**, 035310 (2007).
[61] Y. M. Niquet, D. Rideau, C. Tavernier, H. Jaouen, and X. Blase, Phys. Rev. B **79**, 245201 (2009).
[62] M. Luisier and A. Schenk, J. Comput. Theor. Nanosci. **5**, 1031

(2008).
[63] Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).
[64] M. S. Lundstrom, IEEE Electron Device Lett. **22**, 293 (2001).
[65] K. Rogdakis, S. Poli, E. Bano, K. Zekentes, and M. G. Pala, Nanotechnology **20**, 295202 (2009).
[66] S. Cauley, J. Jain, C.-K. Koh, and V. Balakrishnan, J. Appl. Phys. **101**, 123715 (2007).
[67] E. D. Petersen, S. Li, K. Stokbro, H. H. B. Sorensen, P. C. Hansen, S. Skelboe, and E. Darve, J. Comput. Phys. **228**, 5020 (2009).
[68] W. Gropp, E. Lusk, N. Doss, and A. Skjellum, Parallel Comput. **22**, 789 (1996).
[69] http://www.nccs.gov/computing-resources/jaguar/