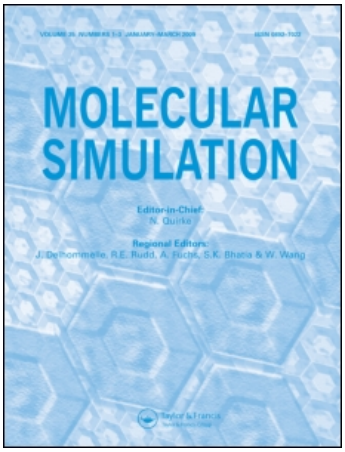

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Convergence Acceleration Scheme for Self-consistent Orthogonal-basis-set Electronic Structure Methods

D. A. Areshkin[a]; O. A. Shenderova[a]; J. D. Schall[a]; D. W. Brenner[a]
[a] Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC, USA

Online publication date: 26 October 2010

To cite this Article Areshkin, D. A. , Shenderova, O. A. , Schall, J. D. and Brenner, D. W.(2003) 'Convergence Acceleration Scheme for Self-consistent Orthogonal-basis-set Electronic Structure Methods', Molecular Simulation, 29: 4, 269 — 286
To link to this Article: DOI: 10.1080/0892702031000092197
URL: http://dx.doi.org/10.1080/0892702031000092197

Molecular Simulation, 2003 Vol. 29 (4), pp. 269–286

Convergence Acceleration Scheme for Self-consistent Orthogonal-basis-set Electronic Structure Methods

D.A. ARESHKIN*, O.A. SHENDEROVA, J.D. SCHALL and D.W. BRENNER

Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC 27695-7907, USA

(Received July 2002; In final form December 2002)

A new self-consistent convergence acceleration scheme that is a variant of the Newton–Raphson algorithm for non-linear systems of equations is presented. With this scheme, which is designed for use with minimal orthogonal basis set electronic structure methods, the conventional Newton–Raphson scaling with respect to the number of atoms is enhanced from quartic to cubic. The scheme is demonstrated using a self-consistent environment-dependent tight binding model for hydrocarbons that allows an efficient and reasonably precise simulation of charge density distortions due to external electric fields, finite system sizes, and surface effects. In the case of a metallic system, self-consistency convergence starts at a high fictitious temperature, typically 1500 K. As the electron density approaches the self-consistent configuration the temperature is decreased. Typically, seven to nine iterations are required to achieve self-consistency in metallic systems to a final temperature of 300 K. For systems with a finite band gap the convergence may start at the target temperature so that temperature reduction is unnecessary, and typically two iterations are needed to achieve self-consistency. The convergence algorithm can handle extremely high applied fields and is very robust with respect to initial electron densities.

Keywords: Convergence acceleration scheme; Newton–Raphson algorithm; Self-consistent orthogonal-basis-set electronic structure method; HOMO–LUMO gap

P.A.C.S: 31.15.Ne; 31.15.-p. 71.15.-m; 73.22.-f

INTRODUCTION

An efficient convergence scheme is an essential part of any self-consistent (SC) electronic structure method. At present there is no universal convergence acceleration algorithm that fits all possible situations, and therefore a variety of algorithms have been developed. These various schemes differ in the types of systems for which they can be efficiently used, the scaling with respect to the number basis functions, and whether convergence is guaranteed. Some of the properties of existing SC convergence algorithms are summarized in Table I. All methods employing scalar function minimization have quadratic scaling per iteration if an orthogonal basis set is used. Scaling does not account for eigenproblem solving that always scales cubically for the systems considered here. Therefore the total number of flops per SC iteration is $O(N^3)$ plus the appropriate table value which indicates the price of "charge mixing".

Methods for convergence acceleration generally fall into two categories. The first employs a variational principle and minimizes the total energy or other scalar target function that has a global minimum at the ground state [1–6]. Widely-used methods of this type include level shifting [2,6], direct inversion in the iteration subspace (DIIS) [2], relaxed constraints algorithms [2,3], and second-order minimization methods based on variable metrics algorithms by Broyden-Fletcher-Goldfarb-Shanno (BFGS) [4,5,7] or Davidson-Fletcher-Powell (DFP) [7]. The second category encompasses algorithms that minimize charge density or potential deviations from their self-consistent values [8–10] by solving a system of non-linear equations. These methods involve the evaluation of either an exact [8] or an approximate [9,10] Jacobian to solve a system of non-linear equations for charge density components by using a Newton–Raphson algorithm.

Each convergence acceleration method has its strengths and drawbacks. The convergence of level

*Corresponding author. E-mail: denis@eos.ncsu.edu

ISSN 0892-7022 print/ISSN 1029-0435 online
DOI: 10.1080/0892702031000092197

TABLE I Properties of major SC convergence acceleration algorithms. The last column gives the number of flops required to calculated the input charge density for the $(k+1)$th iteration provided the kth iteration eigenproblem has already been solved

Method category	*Method*	*Convergence*	*Number of iterations*	*Number of iterations for metallic systems*	*Scaling for a single iteration: Orthogonal/ Non-orthogonal basis*
Scalar function minimization	Level shifting	Guaranteed for large level shift parameter	Medium or Large	May converge slowly	$O(N^2)/O(N^3)$
	DIIS	May diverge	Medium	–	$O(N^2)/O(N^3)$
	Variable Metrics BFGS or DFP	Guaranteed	Medium	Second order approximation is not efficient	$O(N^2)$
	RCA	Guaranteed for uniform well posed Systems	Medium	May be large	$O(N^2)/O(N^3)$
Solving a system of non-linear equations	Broyden	Guaranteed	Large $O(N)$	Large $O(N)$	$O(N^2)$
	Newton–Raphson for system of non-linear equations	Guaranteed	Small	Small	$O(N^4)$

shifting algorithms depends on the value of the level shift parameter. The level shift parameter is not *a priori* known and should be individually optimized for each particular problem. It can be chosen large enough to provide a sufficiently large convergence radius and to guarantee convergence from the given starting point. However, the level shift parameter cannot be chosen to be too large because the convergence rate is inversely proportional to the level shift parameter magnitude. For metallic systems the level shift parameter should be chosen sufficiently large to achieve enough separation between the HOMO and LUMO. That may result in inefficiencies in the first-order perturbation approach, and hence slow convergence.

Variable metrics methods use a second-order target function expansion for iterative minimization by Newton's method. To achieve $O(N^2)$ scaling an approximate rather than exact Hessian matrix is used. It is built up iteratively; each SC iteration is used for the approximate Hessian improvement. If the target function is quadratic the approximate Hessian converges towards its exact value after N iterations. Metallic systems, however, are a serious obstacle for this method. The reason is that due to the strong mixing between occupied and unoccupied levels, and the strong non-linearity of the Fermi function, one usually cannot perform a full step in the direction prescribed by the second-order method. Backtracking is used instead to find the line minimum along the prescribed direction. That results in an extra expense for the target function calculation, and hence deteriorates scaling, which becomes $O(N^3)$ instead of $O(N^2)$. Backtracking also makes the iterative Hessian improvement much less efficient.

Recently Cancès and Le Bris [2,3] introduced a new class of relaxed constraints algorithms (RCA). They provided a rigorous mathematical proof of convergence from any starting point provided the system is uniform well posed, i.e. the system has a finite HOMO–LUMO gap. As follows from the proof, the number of iterations towards self-consistency is inversely proportional to the HOMO–LUMO distance. Thus RCA, or at least its variants described in Refs. [2,3] may be inefficient when applied to metallic systems.

The DIIS algorithm is superseded by the RCA in a sense of robustness, but may be slightly faster than the simplest RCA variant called the optimal damping algorithm. DIIS still remains popular for mostly historical reasons; it was introduced almost a decade prior to RCA. We put no comments on "Number of iterations for metallic systems" in Table I because there is no proof of convergence for DIIS, and thus its behavior for metals cannot be predicted.

Solving a system of non-linear equations for charge density components with an exact Jacobian is more efficient in terms of the number of required iterations than scalar target function minimization. This is because the Newton–Raphson algorithm for the *system* of equations drives *each* charge density component during each iteration step towards its self-consistent value. Scalar target function minimization does not possess this property. While a scalar target function (e.g. total energy) is driven to its minimum, some of charge density components may deviate further at each step from their self-consistent values.

The exact evaluation of a Jacobian has a prohibitive $O(N^4)$ scaling that makes the advantageous number of iterations of limited practical value. On the other hand the cost of a single SC iteration, which is also used to improve an approximate Jacobian by the Broyden method [8,9], is an $O(N^2)$ flops operation [10]. However $O(N)$ iterations are required to build up an approximate Jacobian that is sufficiently close

to the real one. The Broyden algorithm is best suited for use in conjunction with $O(N)$ methods for energy minimization. However, the Broyden method is not a good candidate for metallic systems for which $O(N)$ energy minimization [11,12] cannot be efficiently applied because the localization range of Wannier-like orbitals is larger than the typical system size.

We have recently begun to explore the electronic structure of medium-sized (i.e. 100–1000 atoms) metallic and semiconducting carbon–hydrogen systems in applied fields using a SC tight-binding scheme in which corrections involving a block-diagonal sparse matrix are added to a tight-binding Hamiltonian matrix [13,14]. These systems, which include fullerene nanotubes, nanodiamond clusters, and hybrid nanotube-diamond structures, hold promise as nanoelectronic device components, field emitters, and sensors [14,15]. Unfortunately, for reasons mentioned above existing convergence acceleration schemes are not practical for systems of this type, and therefore a new convergence algorithm had to be developed that satisfies several requirements that are not met by the methods outlined above. First, large applied fields that may exceed 1.0 V/Å must be handled. This poses difficulties for selecting a "good" starting point that is usually chosen by means of semi-empirical methods, e.g. non-self-consistent TB, or even as a superposition of electron densities of neutral atoms. Thus a method is needed that is insensitive to the starting point, i.e. convergence is guaranteed regardless of the choice of initial electron density. Second, the method must be effective for both metallic and semiconducting systems. Finally, the total number of flops required to achieve self-consistency should scale with the system size no worse than $O(N^3)$. In addition, a convergence scheme that is applicable to transport problems is desirable, and thus the algorithm framework should in principle be extendible to non-equilibrium cases.

We demonstrate our scheme using an environment dependent tight binding (EDTB) methodology combined with self-consistent (SC) field corrections [16,17]. The EDTB approach effectively includes three-center integrals through the dependence of hopping integrals on their atomic environment, resulting in a method that in many cases can produce results that are superior to DFT schemes with the same number of basis functions per atom. The self-consistent corrections involve adding block-diagonal matrix elements ΔH to the tight-binding Hamiltonian matrix. The matrix ΔH is sparse, with its elements $\Delta H\alpha\beta$ being zero if indexes α and β do not belong to the same atom (though α and β may stand for different orbitals of the same atom for non-zero $\Delta H\alpha\beta$). The method for computing ΔH, which involves using an explicit minimal Gaussian basis set, is described elsewhere [13].

This paper is organized as follows. The Second Section contains a detailed discussion of the standard implementation of the Newton–Raphson convergence acceleration method. Third Section shows how SC convergence can be accelerated in semiconductors by applying the Newton–Raphson algorithm for non-linear system of equations. The basic idea of how $O(N^3)$ instead of $O(N^4)$ scaling can be achieved for the Newton–Raphson algorithm is demonstrated. Fourth Section describes the extension of the Newton–Raphson algorithm with $O(N^3)$ scaling for metallic systems. The Fifth Section demonstrates two sample applications of the convergence acceleration algorithm.

NEWTON–RAPHSON METHOD FOR NON-LINEAR SYSTEMS OF EQUATIONS

The Newton–Raphson method [1] belongs to the class of globally convergent methods, where convergence is guaranteed regardless of the initial charge density guess. The idea of the algorithm is the following. For the non-self-consistent set of Kohn-Sham equations, some input electron density ρ_{in} determines the Hamiltonian matrix H and hence output electron density

$$\rho_{\text{out}}(r) = \sum_{\substack{\alpha=1,\beta=1 \\ \{\alpha,\beta\}\in\text{Same Atom}}}^{N} \varphi_\alpha(r)\varphi_\beta(r)\sum_{i=1}^{N} f_i Ci_\alpha Ci_\beta \tag{1a}$$

$$f_i \equiv f[\varepsilon_i] = \left[1+\exp\left(\frac{\varepsilon_i-\mu}{kT}\right)\right]^{-1}.$$

Here φ_ζ is the ζth atomic orbital, Ci_ζ is the ζth component of the ith eigenvector of matrix $H = H[\rho_{\text{in}}]$, and ε_i is the ith Hamiltonian eigenvalue. Because orbital orthogonality is assumed, Eq. (1a) contains only the products with indexes α and β belonging to the same atom.

Further we operate with uncompensated Mulliken populations

$$q_{\alpha\beta} = 2\sum_{i=1}^{N/2} Ci_\alpha Ci_\beta - q0_\alpha\delta_{\alpha\beta} \quad \text{if } T=0, \tag{1b}$$

$$q_{\alpha\beta} = 2\sum_{i=1}^{N} f_i Ci_\alpha Ci_\beta - q0_\alpha\delta_{\alpha\beta} \quad \text{if } T>0 \tag{1c}$$

rather than electron densities. Here $\delta_{\alpha\beta}$ is the Kronneker delta, and $q0_\alpha$ is the orbital Mulliken population in the bulk material for which a TB parameterization (in this case the EDTB parameterization) has been performed [13]. For example, $q0_\alpha = 1.2028$ if α stands for an s orbital in carbon. The quantity $q_{\alpha\beta}$ is the net orbital charge minus its equilibrium value, i.e. a measure of the deviation

from neutrality. Because indexes α and β belong to the same atom, matrix $q_{\alpha\beta}$ is sparse and further is treated as a double indexed vector. If we apply a small change $\Delta qIn_{\alpha\beta}$ to input vector $qIn_{\alpha\beta}$ it will result in a small change δH of the Hamiltonian H and a small change $\Delta qOut_{\alpha\beta}$ of output vector $qOut_{\alpha\beta}$. If $\Delta qIn_{\alpha\beta}$ is infinitesimally small, we can relate it linearly to $\delta H_{\alpha\beta}$. This in turn can be related to $\Delta qOut_{\alpha\beta}$ in a linear fashion using first order perturbation theory

$$\delta H_{\alpha\beta} = \sum_{\substack{\mu=1,\nu=1 \\ \{\mu,\nu\}\in \text{Same Atom}}}^{N} U_{\alpha\beta,\mu\nu}\Delta qIn_{\mu\nu}, \tag{2a}$$

$$\Delta qOut_{\alpha\beta} = 2 \sum_{\substack{\mu=1,\nu=1 \\ \{\mu,v\}\in \text{Same Atom}}}^{N} A_{\alpha\beta,\mu\nu}\,\delta H_{\mu\nu}. \tag{2b}$$

To make further calculations more convenient, we do not include the spin factor of 2 into matrix A. Due to the atomic orbital's orthogonality, the change of Hamiltonian matrix $\delta H_{\alpha\beta}$ is applied only to the elements with indexes α and β belonging to the same atom. That allows us to view $\delta H_{\alpha\beta}$ as a vector with the same length as $q_{\alpha\beta}$. If we want the self-consistency condition to be valid we must apply $\Delta qIn_{\alpha\beta}$ such that the output charge density $qOut_{\alpha\beta} + \Delta qOut_{\alpha\beta}$ equals $qIn_{\alpha\beta} + \Delta qIn_{\alpha\beta}$:

$$\begin{aligned} &qOut_{\alpha\beta} + \Delta qOut_{\alpha\beta} \\ &\quad = qOut_{\alpha\beta} + \sum_{\substack{\mu=1,\nu=1 \\ \{\mu,\nu\}\in \text{Same Atom}}}^{N} \mathfrak{B}_{\alpha\beta,\mu\nu}\Delta qIn_{\mu\nu} \\ &\quad = qIn_{\alpha\beta} + \Delta qIn_{\alpha\beta}. \end{aligned} \tag{3}$$

Matrix $\mathfrak{B}$ is a product of U and 2 A as defined by Eq. (2a). During each iteration step we use Eq. (3) to obtain the additional contribution of $\Delta qIn_{\alpha\beta}$ to the current iteration charge input vector $qIn_{\alpha\beta}$. Vector ΔqIn is a solution of a linear system

$$(E - \mathfrak{B})\Delta qIn = qOut - qIn, \tag{4}$$

where E is the identity matrix.

If the exchange energy is represented as a first order expansion over Mulliken population deviations from their bulk values, the matrix U is the same for each iteration. Its evaluation requires computation of Hartree and exchange integrals. These integrals can be evaluated analytically for Gaussian basis functions in $O(N^2)$ flops. The main computational burden is imposed by the evaluation of matrix A. As will be shown in "Improved Scaling for the Newton–Raphson Algorithm Section" the exact evaluation of A, which is an $O(N^4)$ operation, can be substituted by the approximate evaluation. In contrast to the Broyden method the approximate Jacobian $A_{\alpha\beta,\mu\nu} = \partial_{\delta H\mu\nu}\Delta qOut$ is not calculated iteratively. Instead a new value of $A = A[qIn]$ is computed during each iteration. Its evaluation requires about $4\,N^3$ flops. Remarkably, precision in a wide range is not related to computational workload. For sufficiently large systems the precision enhancement from 10^{-2} to 10^{-3} leads to $O(N^2)$ extra operations. According to our experience if the maximum deviation of approximate matrix A elements from their exact values is around 10^{-2}, any further precision enhancement does not accelerate the convergence.

The total number of flops per iteration for the cubically scaled Newton–Raphson algorithm is approximately $4/3N^3 + 4N^3 + N^3 + N^3 = (7 + 1/3)N^3$. The terms of that sum are, respectively, the cost for the eigenproblem solution, matrix A evaluation, evaluation of the $A{\cdot}U$ product, and solution of Eq. (4) by LU decomposition. To supersede the methods that perform scalar target function minimization, the number of iterations for the Newton–Raphson algorithm should be at least seven times less than for its competitors. We have not performed any tests that allow direct comparisons of different methods. Our goal is to accelerate convergence primarily for metallic systems; DIIS, RCA, and variable metrics methods require extra provisions to handle metals. While a comparison of the number of iterations for the molecules with a finite HOMO–LUMO gap cannot be extrapolated for metallic systems, it can be used for a very crude estimate of efficiency. The general rule valid for all methods is that the larger the HOMO–LUMO gap, the more probable is convergence (if the algorithm may diverge), and the smaller the number of iterations. Some insight into the comparative efficiencies of these methods can be drawn from convergence data given in Ref. [3] for the ODA applied to acetaldehyde. The HOMO–LUMO gap for acetaldehyde is 3.9 eV. Our SC-EDTB parametrization [13] is restricted to hydrocarbons; therefore we have to use a hydrocarbon substitute with a similar HOMO–LUMO gap. We choose a benzene molecule for that purpose. Acetaldehyde has strong polarization and its SC electron density substantially differs from the electron density obtained by core Hamiltonian diagonalization. We apply an external field of 1.0 V/Å in the plane of a benzene molecule to create a substantial difference between the SC and initial electron density that is also obtained by core Hamiltonian diagonalization. The SC-EDTB HOMO–LUMO gap for benzene in the applied field is 4.2 eV. We monitor convergence by the largest Mulliken population deviation, while Cancès and Le Bris [3] use total energy for that purpose. Because energy deviation from its ground state is quadratically proportional to charge density deviation

[18] 10^{-10} relative error in total energy is equivalent to 10^{-5} relative error in Mulliken population. The former is achieved by ODA after 24 iterations, while it takes 4 iterations for the Newton–Raphson algorithm to achieve a 10^{-5} Mulliken population convergence.

While the algorithm's relative speed remains an open question, one of the main reasons for using the Newton–Raphson scheme is its robustness with respect to the choice of initial electron density. However, the Newton–Raphson and variable metrics algorithms called "globally convergent" may still diverge. Global convergence means that regardless of starting point the minimization direction chosen in each iteration is correct. However the step size along that direction is chosen by using first order (Newton–Raphson) or second order (variable metrics) approximations. These approximations may not hold for rapidly changing functions, and their failure leads to a wrong step choice and possibly to the algorithm divergence. Convergence is guaranteed only if special backtracking checks are used to control the step size and reduce it if necessary. Because backtracking implies solving one or more extra eigenproblems we do not use it. We have performed convergence tests for semiconducting and metal systems under extremely high voltage. The first test described in "Examples: Hydrocarbon Nano-structures in Applied Field Section" employed a hydrogen passivated diamond nano-cluster with an applied voltage exceeding the cluster band gap by a factor of three. In the second test a field of 20 V was applied along a 75 Å long (6,0) metallic single-wall carbon nanotube. Perfect screening effects and field enhancement up to 2.0 V/Å were observed near the tube ends. In both cases convergence problems did not arise. Note that convergence is not guaranteed for DIIS, and is conditional for the RCA and level shifting algorithms.

Another reason for developing a Newton–Raphson based scheme is its applicability to non-equilibrium situations. None of the scalar target function minimization techniques will work for non-equilibrium cases because SC non-equilibrium electron densities do not correspond to a global energy minimum. At the same time the matrix $A_{\alpha\beta,\mu\nu} = \partial_{\delta H_{\mu\nu}} \Delta qOut$ can still be readily evaluated for non-equilibrium systems, which means that the analogue of the Newton–Raphson scheme presented here can be used for non-equilibrium studies.

IMPROVED SCALING FOR THE NEWTON–RAPHSON ALGORITHM

The key equation (2b) employed by the Newton–Raphson method relates the infinitesimally small change of the Hamiltonian matrix δH to the induced changes of Mulliken population components. We first consider this equation assuming zero temperature and a non-degenerate HOMO, which implies a finite HOMO–LUMO gap Δ. In "Examples: Hydrocarbon Nano-structures in Applied Field Section", it is extended to finite temperatures and metallic systems that may have degenerate HOMOs. The symbol $H0$ is used to denote the unperturbed Hamiltonian matrix, $C0$ denotes a matrix with columns that are $H0$ eigenvectors, and $C0i$ and $\varepsilon 0_i$ denote the ith column of $C0$ and the ith eigenvalue of $H0$, respectively. For zero temperature and a non-degenerate HOMO the component $q_{\alpha\beta}$ of the uncompensated Mulliken population vector is given by Eq. (1b). The variation of $q_{\alpha\beta}$ caused by the variation of the Hamiltonian matrix δH is given by

$$\Delta q_{\alpha\beta} = 2\sum_{i=1}^{N/2} C0i_\alpha (C0\, \varepsilon Inv_i C0^{\mathrm{T}}\, \delta H\, C0i)_\beta + 2\sum_{i=1}^{N/2} C0i_\beta (C0 \varepsilon Inv_i C0^{\mathrm{T}}\, \delta H\, C0i)_\alpha. \quad (5)$$

Here the superscript "T" denotes a transposition, and $(\ldots)_\zeta$ indicates the ζth component of the expression in the parenthesis. The symbol εInv_i stands for a diagonal matrix that has $(\varepsilon 0_i - \varepsilon 0_j)^{-1}$ at the jth position if $\varepsilon 0_i \neq \varepsilon 0_j$, and 0 otherwise. Equation (5) is similar to the one used by Brown [8] who first proposed using the Newton–Raphson algorithm for SC convergence acceleration. We provide a derivation of Eq. (5) in the Appendix. This is done for two reasons. First, it systematically handles the case of degenerate energy levels. Second, the intermediate equations obtained during the derivation are crucial for understanding the finite temperature case presented in "Examples: Hydrocarbon Nano-structures in Applied Field Section".

To estimate the number of flops with respect to the number of orbitals N required to evaluated Eq. (5), one needs to switch from a matrix notation to an explicit summation over matrix indexes. For brevity we consider only the first sum in Eq. (5). Here $(\varepsilon Inv_i)_{mm}$ is the mth diagonal element of εInv_i.

$$\sum_{i=1}^{N/2} C0i_\alpha (C0\, \varepsilon Inv_i\, C0^{\mathrm{T}}\ \delta H\, C0i)_\beta = \sum_{n=1}^{N}\sum_{s=1}^{N}\left[\sum_{i=1}^{N/2}\sum_{m=(N/2)+1}^{N} C0i_\alpha \right.$$

$$\left. \text{hskip24ptBigg} \times C0m_\beta (\varepsilon Inv_i)_{mm}\, C0m_n C0i_s\right] \delta H_{ns}. \quad (6)$$

We used the asymmetry of $(\varepsilon Inv_i)_{mm}$ with respect to indexes i and m to reduce the summation range over m from $\{1, N\}$ to $\{N/2+1, N\}$. Note that because $\varepsilon 0_i$ belongs to the occupied portion of the spectrum,

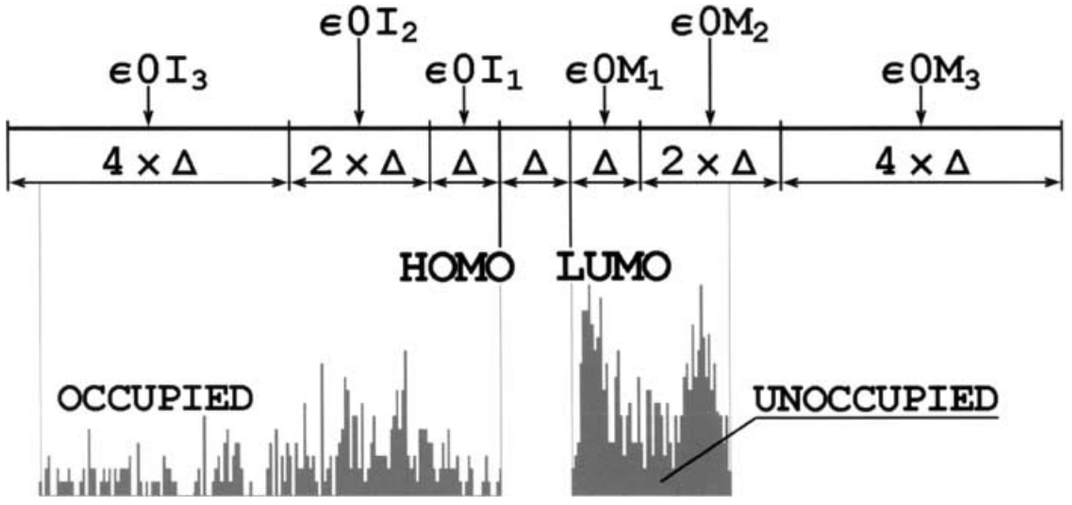

FIGURE 1 Sample semiconducor spectrum and energy axis partitioned for the use with Eq. (9). For better efficiency of approximation Eqs. (9)–(11) (un)occupied intervals can be shrunk proportionally to the size of the (un)occupied portion of the spectrum.

and $\varepsilon 0_m$ to the unoccupied portion of the spectrum, there is always a finite difference between these energies that is equal to or greater than Δ. Because δH is sparse, the double summation over indexes n and s requires order of N flops. If the system is composed solely of carbon atoms and we assume four orbitals per atom, there are ten distinguishable combinations of index pairs $\{\alpha, \beta\}$, and $\{n, s\}$ per atom. Therefore the expression in the parenthesis in Eq. (6) can be viewed as a square matrix (further on we refer it as matrix A) with dimensions $2.2N$, and double indexing $\{\alpha, \beta\}$, and $\{n, s\}$ in each dimension. The term $(\varepsilon Inv_i)_{mm}$ couples the summation over indexes i and m. Because the summation over i and m cannot be performed separately, N^2 flops are required to evaluate each entry of A, and $(2.5N)^2N^2$ flops are required for the evaluation of the entire matrix. To decouple the summation over i and m, and thus switch from N^4 to N^3 scaling, we substitute $(\varepsilon Inv_i)_{mm}$ by its power approximation. To demonstrate the basic idea let us consider the sample spectrum and its partitioning illustrated in Fig. 1. The centers of the energy intervals for occupied and unoccupied parts of the spectrum are marked as $\varepsilon 0I_x$ or $\varepsilon 0M_x$, respectively. For each given $\varepsilon 0_i$ and $\varepsilon 0_m$ which belong to the intervals with centers at $\varepsilon 0I_x$ or $\varepsilon 0M_y$, respectively, the value of

$$(\varepsilon Inv_i)_{mm} = \frac{1}{\varepsilon 0_i - \varepsilon 0_m} = \frac{1}{\varepsilon 0I_x - \varepsilon 0M_y} \times \frac{1}{1 + \frac{(\varepsilon 0_i - \varepsilon 0i_x) - (\varepsilon 0_m - \varepsilon 0m_y)}{\varepsilon 0I_x - \varepsilon 0M_y}} \quad (7)$$

can be approximated by a Taylor expansion over small parameter

$$\lambda = \frac{(\varepsilon 0_i - \varepsilon 0I_x) - (\varepsilon 0_m - \varepsilon 0M_y)}{\varepsilon 0I_x - \varepsilon 0M_y}. \quad (8)$$

As will be justified below, the third order Taylor expansion appears to be optimal. By virtue of the approximate representation for $(\varepsilon Inv_i)_{mm}$, matrix A can be expressed as

$$A \approx \sum_{i=1}^{N/2} \sum_{m=N/2+1}^{N} C0i_\alpha\, C0m_\beta\, C0m_n\, C0i_s \times \sum_{x,m\in Int_y}^{X} \sum_{y,i\in Int_x}^{Y} \sum_{j=0}^{3} \sum_{k=0}^{3-j} \Lambda_{xy,jk} (\varepsilon 0_i - \varepsilon 0I_x)^j \times (\varepsilon 0_m - \varepsilon 0M_y)^k. \quad (9)$$

Here $\Lambda_{xy,jk}$ is the Taylor expansion coefficient for Eq. (7) over parameter λ for energies $\varepsilon 0_i$ and $\varepsilon 0_m$ belonging to the spectral intervals centered at $\varepsilon 0I_x$ and $\varepsilon 0M_y$, respectively. The symbols X and Y are the number of energy intervals for occupied and unoccupied spectrum regions, respectively.

Spectral partitioning is an important element for the algorithm scaling improvement. Suppose the energy intervals are numbered by index $k = \{0, 1, 2\ldots\}$, which increases with the distance from the band gap for both occupied and unoccupied portions of the spectrum. To maintain the number of intervals X and Y reasonably small we increase the interval size exponentially with the increase of index k. For example, if the size of the intervals

adjacent to the band gap do not exceed the bandgap Δ, and the interval's size increases as 2^k (cf. Fig. 1), parameter λ never exceeds 1/2. In that case, if the third order power approximation is used, the maximum error of the Taylor expansion does not exceed 6%. At the same time, even for a small bandgap of approximately 0.4 eV the 2^k partitioning results in only 6 intervals that cover the entire range ($\sim$25 eV) occupied by the valence electrons in hydrocarbons. Therefore the summation limits X and Y in Eq. (9) are usually between 3 and 5.

The advantage of Eq. (9) over the original expression for A in Eq. (6) is that now the summation over indexes i and m can be done separately. First we evaluate the sums

$$SI_{x,j,\alpha s} = \sum_{\varepsilon 0_i \in Int_x} C0i_\alpha \, C0i_s (\varepsilon 0_i - \varepsilon 0I_x)^j,$$
$$SM_{y,k,\beta n} = \sum_{\varepsilon 0_m \in Int_y} C0m_\beta \, C0m_n (\varepsilon 0_m - \varepsilon 0M_y)^k. \tag{10}$$

The number of flops required to evaluate two the matrixes $SI_{x,j}$ and $SM_{y,k}$ for the given j and k and for the entire set of all possible x and y is N^3. Because j and k run from 0 to 3 the total number of flops is $4N^3$. Although it is convenient to precalculate $(\varepsilon 0_i - \varepsilon 0I_x)^j$ and $C0i_s(\varepsilon 0_i - \varepsilon 0I_x)^j$ for all i and s before computing $SI_{x,j}$, the evaluation of different matrix elements of $SI_{x,j}$ or $SM_{y,k}$ is still independent, and thus highly parallelizable. Given $SI_{x,j}$ and $SM_{y,k}$ the evaluation of A becomes an $O(N^2)$ process:

$$A_{\alpha\beta,ns} \approx \sum_{x=0}^{Y} \sum_{y=0}^{X} \sum_{j=0}^{3} \sum_{k=0}^{3-j} \Lambda_{xy,\,jk} SI_{x,j,\alpha s} SM_{y,k,\beta n}. \tag{11}$$

METALLIC SYSTEMS AT FINITE TEMPERATURES

At non-zero temperatures $q_{\alpha\beta}$ is given by Eq. (1c). Its variation $\Delta q_{\alpha\beta}$ is

$$\begin{aligned}\Delta q_{\alpha\beta} &\equiv \Delta q1_{\alpha\beta} + \Delta q2_{\alpha\beta} \\ &= 2\sum_{i=1}^{N} f(\varepsilon 0_i)\Delta(C0i_\alpha \, C0i_\beta) \\ &\quad + 2\sum_{i=1}^{N} \Delta f(\varepsilon 0_i) C0i_\alpha \, C0i_\beta.\end{aligned} \tag{12}$$

The term $\Delta q1_{\alpha\beta}$ can be obtained in the manner quite similar to those used to obtain Eq. (5) from Eq. (1b)

$$\Delta q1_{\alpha\beta} = 2 \sum_{\substack{n=1,\,s=1 \\ \{n,s\} \in \text{Same Atom}}}^{N} A_{\alpha\beta,ns} \, \delta H_{ns},$$

where

$$\begin{aligned}A_{\alpha\beta,ns} &= \sum_{i=1}^{N/2-KT-1} \sum_{m=1}^{N} C0i_\alpha \, C0m_\beta (\varepsilon Inv_i)_{mm} \, C0m_n C0i_s \\ &\quad + \sum_{i=1}^{N/2-KT-1} \sum_{m=1}^{N} C0i_\beta C0m_\alpha (\varepsilon Inv_i)_{mm} \, C0m_n C0i_s \\ &\quad + \sum_{i=N/2-KT}^{N/2+KT} \sum_{m=1}^{N} f(\varepsilon 0i) C0i_\alpha \, C0m_\beta (\varepsilon Inv_i)_{mm} \\ &\quad \times C0m_n C0i_s \\ &\quad + \sum_{i=N/2-KT}^{N/2+KT} \sum_{m=1}^{N} f(\varepsilon 0i) C0i_\beta C0m_\alpha \\ &\quad \times (\varepsilon Inv_i)_{mm} \, C0m_n C0i_s,\end{aligned} \tag{13}$$

where KT (not to be confused with thermal energy kT) is that absolute distance from the Fermi level (cf. Fig. 2) that marks the region where the Fermi factor $f(\varepsilon 0_i)$ is different from either 0 or 1. The summation range $\{N/2 - KT, N/2 + KT\}$ stands for all energy levels that are close enough to the Fermi energy to make the Fermi factor $1 > f(\varepsilon 0_i) > 0$. The summation limit $N/2 \pm KT \pm 1$ means that the summation is done to the last (from the first) level that is not influenced by temperature smearing. After cumbersome though straight-forward algebraic transformations that utilize the symmetry of δH, the asymmetry of $(\varepsilon Inv_i)_{mm}$ and the symmetry of $[f(\varepsilon 0_i) - f(\varepsilon 0_m)]$ $(\varepsilon Inv_i)_{mm}$ with respect to indexes i and m, one gets the final expression for $\Delta q1_{\alpha\beta}$. This expression can be applied to metallic systems at finite temperatures:

$$\begin{aligned}\Delta q1_{\alpha\beta} = 2 \sum_{\substack{n=1,\,s=1 \\ \{n,s\} \in \text{Same Atom}}}^{N} \delta H_{ns} (&ASym_{\alpha\beta,ns} \\ &+ ASym_{\beta\alpha,ns}).\end{aligned} \tag{14a}$$

Here $ASym_{\alpha\beta,ns} + ASym_{\beta\alpha,ns}$ is the symmetrized

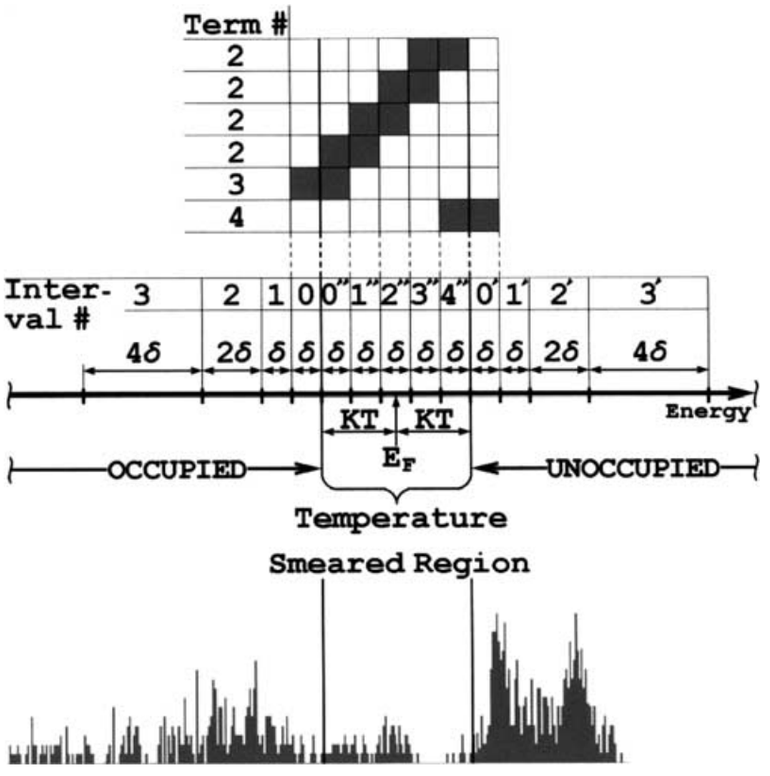

FIGURE 2 *Bottom*: Sample metallic spectrum. *Middle*: Partitioned energy axis. *Top*: Grayed squares denote spectral regions containing directly evaluated energy terms required to build matrix *Asym* (14b). The regions are sorted by terms in Eq. (14b). For example, if $\epsilon 0_i$ and $\epsilon 0_m$ belong respectively, to intervals $0''$ and 0 third term in sum (14b) must be evaluated directly. The number of directly evaluated terms is proportional to the area of shaded region. It can be made reasonably small by the appropriate choice of interval δ.

matrix $A_{\alpha\beta,ns}$

$$
\begin{aligned}
&ASym_{\alpha\beta,ns} \\
&= \sum_{i=1}^{N/2-KT-1} \sum_{m=N/2+KT+1}^{N} (\varepsilon Inv_i)_{mm} C0i_\alpha C0m_\beta C0m_n C0i_s \\
&+ \sum_{i=N/2-KT}^{N/2+KT} \sum_{m=i+1}^{N/2+KT} [f(\varepsilon 0_i) - f(\varepsilon 0_m)](\varepsilon Inv_i)_{mm} \\
&\times C0i_\alpha C0m_\beta C0m_n C0i_s \\
&- \sum_{i=N/2-KT}^{N/2+KT} \sum_{m=1}^{N/2-KT-1} [1 - f(\varepsilon 0_i)](\varepsilon Inv_i)_{mm} C0i_\alpha \\
&\times C0m_\beta C0m_n C0i_s \\
&+ \sum_{i=N/2-KT}^{N/2+KT} \sum_{m=N/2+KT+1}^{N} f(\varepsilon 0_i)(\varepsilon Inv_i)_{mm} C0i_\alpha \\
&\times C0m_\beta C0m_n C0i_s. \qquad (14b)
\end{aligned}
$$

To achieve N^3 scaling we use a strategy similar to that for semiconducting systems. The first term in Eq. (14b) can be handled in the same fashion as Eq. (6) for semiconducting system. The sole difference is that the role of the bandgap is now played by the $\{\mu - KT, \mu + KT\}$ region, which we refer to as the temperature smeared region or TSR. To handle the last three terms in an $O(N^3)$ fashion we divide the TSR into a number of smaller intervals. Shaded areas of the upper portion of Fig. 2 mark the regions of the spectrum that must be handled using direct evaluation. To make an illustration let us consider the terms of the second sum in Eq. (14b). If $\varepsilon 0_i$ and $\varepsilon 0_m$ belong to intervals $3''$ and $4''$ then this term should be evaluated exactly as stated by Eq. (14b). On the other hand, if the separation between $\varepsilon 0_i$ and $\varepsilon 0_m$ is sufficiently large, e.g. $\varepsilon 0_i$ and $\varepsilon 0_m$ belong respectively, to intervals $0''$ and $2''$, then a Taylor expansion for $(\varepsilon Inv_i)_{mm}$ can be used. Thus, the portion of the second sum in Eq. (14b) that includes all the terms with pairs $\varepsilon 0_i$ and $\varepsilon 0_m$ not belonging to the same or adjacent intervals can be evaluated using Eqs. (9)–(11). The third (fourth) sum in Eq. (14b) can also be handled by Eqs. (9)–(11) except the terms that have $\varepsilon 0_i$ belonging to interval 0 ($4''$), and $\varepsilon 0_m$ belonging to interval $0''$ ($0'$). These terms must be evaluated directly, i.e. by using Eq. (14b). The interval size can always be chosen small enough to make small the number of directly evaluated terms.

However, to keep the summation limits X and Y in Eq. (9) low, the interval size should not be made too small. In addition, the size of the intervals 1, 0 and 0′, 1′ adjacent to the TSR must have the same size as the intervals inside the TSR so that parameter λ in Eq. (8) does not exceed 1/2.

Let us take a detailed look at the energy dependent terms in Eq. (14b). The first sum energies $\varepsilon 0_i$ and $\varepsilon 0_m$ are at least $2\,KT$ apart, and hence $(\varepsilon Inv_i)_{mm}$ always equals the inverse difference between $\varepsilon 0_i$ and $\varepsilon 0_m$. The third (fourth) sum also has different $\varepsilon 0_i$ and $\varepsilon 0_m$ values because $\varepsilon 0_i$ belong to the TSR, and $\varepsilon 0_m$ to the occupied or unoccupied part of the spectrum. If $\varepsilon 0_i$ lies near the edges of the TSR and very close to $\varepsilon 0_m$, there may be some ambiguity when choosing the value for $(\varepsilon Inv_i)_{mm}$, because $\varepsilon 0_i$ and $\varepsilon 0_m$ may actually represent the energy of a single degenerate state, but may be slightly different due to numerical errors. However, this ambiguity does not have practical consequences because the factors $[1 - f(\varepsilon 0_i)]$ and $f(\varepsilon 0_i)$ in the third and the fourth sums in Eq. (14b) are zeros when $\varepsilon 0_i$ is close to $\varepsilon 0_m$. This is not the case for the energy dependent term of the second sum, which is of the order of unity when $\varepsilon 0_m \to \varepsilon 0_i$

$$\begin{aligned}
&\mathrm{Lim}_{\varepsilon 0_i \to \varepsilon 0_m} \\
&\quad \equiv \lim_{\varepsilon 0_i \to \varepsilon 0_m + 0} \left[f(\varepsilon 0_i) - f(\varepsilon 0_m) \right] (\varepsilon Inv_i)_{mm} \\
&\quad = \exp\left(\frac{\varepsilon 0_i + \mu}{kT}\right) \Bigg/ \left\{ \left[\exp\left(\frac{\varepsilon 0_i}{kT}\right) \right. \right. \\
&\qquad \left. \left. + \exp\left(\frac{\mu}{kT}\right) \right]^{-2} kT \right\}. \qquad (15)
\end{aligned}$$

For the case where $\varepsilon 0_m$ is *exactly* the same as $\varepsilon 0_i$, $(\varepsilon Inv_i)_{mm} = 0$. That may appear strange at a first. Suppose there is a multiply degenerate level with energy $\varepsilon 0_\zeta$, and $\mathrm{Set}\{\varepsilon 0\zeta\}$ is a set of eigenstate indices such that $\varepsilon 0_j = \varepsilon 0_\zeta \quad \forall\ j$ if $j \in \mathrm{Set}\{\varepsilon 0\zeta\}$. In that case the partial sum contained in the second term of Eq. (14b) is zero:

$$\begin{aligned}
&\Delta q1_{\alpha\beta}^{\mathrm{Partial}\{\varepsilon 0_\zeta\}} \\
&= \sum_{\substack{n=1,s=1 \\ \{n,s\} \in SameAtom}}^{N} \delta H_{ns} \sum_{i \in \mathrm{Set}\{\varepsilon 0_\zeta\}} \sum_{\substack{m \in \mathrm{Set}\{\varepsilon 0_\zeta\} \\ m=i+1}} \\
&\quad \times \left[f(\varepsilon 0_i) - f(\varepsilon 0_m) \right] (\varepsilon Inv_i)_{mm} (C0i_\alpha C0m_\beta C0m_n C0i_s \\
&\quad + C0i_\beta C0m_\alpha C0m_n C0i_s) = 0. \qquad (16)
\end{aligned}$$

However, if an infinitesimally small symmetry distortion is applied, a stepwise transition of $(\varepsilon Inv_i)_{mm}$ from 0 to $\mathrm{Lim}_{\varepsilon 0_i \to \varepsilon 0_m}$ occurs. It is intuitively obvious that the system response to the infinitesimally small distortion should also be infinitesimally small. The value of $\mathrm{Lim}_{\varepsilon 0_i \to \varepsilon 0_m}$, however, is of the order of unity. One may ask if there is any change to $\Delta q1_{\alpha\beta}^{Partial\{\varepsilon 0_\zeta\}}$ if a zero value of $(\varepsilon Inv_i)_{mm}$ is substituted by $Lim_{\varepsilon 0_i \to \varepsilon 0_m}$, but the states $C0i$ still remain essentially degenerate? The answer is *no*, i.e. $\Delta q1_{\alpha\beta}^{\mathrm{Partial}\{\varepsilon 0_\zeta\}}$ remains zero. That happens because

$$\sum_{\substack{n=1,s=1 \\ \{n,s\} \in SameAtom}}^{N} \delta H_{ns} C0m_n C0i_s = 0$$

$$\forall \{i,m\} \in \mathrm{Set}\{\varepsilon 0_\zeta\} \quad \text{if} \, i \neq m. \qquad (17)$$

To understand Eq. (17) one needs to recall that the SC correction δH refers to the same system geometry, and hence has the same symmetry as the Hamiltonian $H0$. Therefore δH cannot cause any split of degenerate levels. Instead all levels with index $j \in \mathrm{Set}\{\varepsilon 0\zeta\}$ shift as a whole by $\delta\varepsilon_j = \delta\varepsilon_\zeta$. Thus all eigenvalues $\delta\varepsilon_j$ of matrix $D_{j,k} \equiv C0j^{\mathrm{T}}\, \delta H\, C0k$, $\mathrm{Dim}(D) = \mathrm{Dim}(\mathrm{Set}\{\varepsilon 0_\zeta\})$ (cf. Eq. (A.9)) are the same. If the columns of matrix B are orthonormal eigenvectors Bi of matrix $D_{j,k}$, then

$$D = B^{\mathrm{T}}\, \mathrm{Diag}(\delta\varepsilon_\zeta) B = \mathrm{Diag}(\delta\varepsilon_\zeta). \qquad (18)$$

Here $\mathrm{Diag}(\delta\varepsilon_\zeta)$ is a diagonal matrix will all its diagonal elements equal to $\delta\varepsilon_\zeta$. Hence Eq. (18) is equivalent to Eq. (17). Note, however, that Eq. (14a) was derived without any special assumptions about δH symmetry. Indeed it is valid for any δH, provided it is infinitesimally small. In a sense, this means that the energy level shifts induced by δH must be smaller than the difference between any neighbor energy levels, including those that are degenerate but appear to be split due to the numerical errors. Fortunately this severe limitation does not apply if a Hamiltonian distortion δH does not change the Hamiltonian symmetry. It is convenient to imagine that the entire spectrum is broken into a number of finite-sized intervals, and spectral lines belonging to each interval are united in a single multiply degenerate virtual level. The major effect of applying a finite distortion δH possessing the same symmetry as $H0$ on such a transformed spectrum is that each multiply degenerate virtual level shits *almost as a whole* by some finite amount $\delta\varepsilon_i$.

There is another major limitation imposed on $\delta\varepsilon_i$ that arises from the strongly non-linear derivative of the Fermi factor $f(\varepsilon 0_i)$ that is present in

the subexpression for $\Delta q2_{\alpha\beta}$ in Eq. (12):

$$\Delta q2_{\alpha\beta} = 2\sum_{i=N/2-KT}^{N/2+KT}\left(\partial_{\varepsilon 0_i} f(\varepsilon 0_i)\delta\varepsilon_i + \partial_\mu f(\varepsilon 0_i)\delta\mu\right) \times C0i_\alpha\, C0i_\beta. \quad (19)$$

Using the normalization condition

$$\sum_{i=1}^{N}\Delta f(\varepsilon 0_i) = \sum_{i=1}^{N}\partial_{\varepsilon 0_i} f(\varepsilon 0_i)(\delta\varepsilon_i - \delta\mu) = 0 \quad (20)$$

and noting that $\partial_{\varepsilon 0_i} f(\varepsilon 0_i) = -\partial_\mu f(\varepsilon 0_i)$ we rewrite Eq. (19) as

$$\Delta q2_{\alpha\beta} = 2\sum_{i=1}^{N}\partial_{\varepsilon 0_i} f(\varepsilon 0_i) \times \left(C0i_\alpha C0i_\beta - \frac{\sum_{\eta=1}^{N}\partial_{\varepsilon 0_\eta} f(\varepsilon 0_\eta) C0\eta_\alpha C0\eta_\beta}{\sum_{\xi=1}^{N}\partial_{\varepsilon 0_\xi} f(\varepsilon 0_\xi)} \right)\delta\varepsilon_i. \quad (21)$$

The quantity $\delta\varepsilon_i$ is given by Eq. (A.14), which is not linear with respect to δH because the coefficients Bi_j generally also depend on δH. Fortunately, due to the same symmetry of $H0$ and δH we can use Eq. (17) which makes Eq. (A.14) evolve into Eq. (A.15). Thus, $\Delta q2_{\alpha\beta}$ depends on δH in a linear fashion:

$$\Delta q2_{\alpha\beta} = 2\sum_{\substack{n=1,\,s=1\\ \{n,s\}\in Same\,Atom}}^{N} G_{\alpha\beta,ns}\,\delta H_{ns},$$

where

$$G_{\alpha\beta,ns} = \sum_{i=1}^{N}\partial_{\varepsilon 0_i} f(\varepsilon 0_i) \times \left(C0i_\alpha\, C0i_\beta - \frac{\sum_{\eta=1}^{N}\partial_{\varepsilon 0_\eta} f(\varepsilon 0_\eta) C0\eta_\alpha\, C0\eta_\beta}{\sum_{\xi=1}^{N}\partial_{\varepsilon 0_\xi} f(\varepsilon 0_\xi)} \right) \times C0i_n C0i_s. \quad (22)$$

Equations (20)–(22) are valid as far as a linear approximation for the Fermi factor $f(\varepsilon 0_i)$ remains valid. As a validity criteria for the latter we assume the inequality

$$0 \le f(\varepsilon 0_i) + \partial_{\varepsilon 0_i} f(\varepsilon 0_i)\delta\varepsilon_i \le 1, \quad (23)$$

which is true whenever $|\delta\varepsilon_i| \le kT$. Hence Eq. (22) can be used only if all levels inside the TSR are shifted by no more than kT. Usually this condition is not fulfilled during initial iterations. Therefore, at the nth iteration step, before updating qIn^n by adding ΔqIn^n obtained by solving Eq. (4), we evaluate $\delta\varepsilon_i^{\text{initial}}$:

$$\delta\varepsilon_i^{\text{initial}} = C0i^n\,\delta H_n C0i^n \quad \forall i :: \varepsilon 0_i \in \text{TSR}. \quad (24)$$

Here δH_n is related to ΔqIn^n by Eq. (2a), and $C0i^n$ is the ith eigenvector corresponding to the nth iteration Hamiltonian with input density qIn^n. If $\delta\varepsilon_i^{\text{initial}}$ satisfies the inequality

$$\text{Max}\left|\delta\varepsilon_i^{\text{initial}}\right| \le kT \quad \forall i :: i \in \text{TSR}. \quad (25)$$

We obtain the next iteration density by simply adding ΔqIn^n to qIn^n. Otherwise, if $kT = \tau\text{Max}|\delta\varepsilon_i^{initial}|$ and scaling coefficient $\tau < 1$, then

$$qIn^{n+1} = qIn^n + \tau\Delta qIn^n. \quad (26)$$

Applying $\tau\Delta qIn^n$ instead of ΔqIn^n rescales δH_n and hence makes valid condition (25). SC convergence starts at a non-physically high temperature to make the coefficient τ large and thus to reduce the number of iterations. Note, that the temperature cannot be chosen too high since it means a large TSR and therefore a large number of directly evaluated terms in Eq. (14b). After the solution has partially converged to satisfy Eq. (25) we start a temperature reduction. Temperature reduction. Temperature reduction is performed along with driving the solution towards self-consistency by using the same Newton–Raphson scheme. Suppose $C0$ represents a (non)-self-consistent solution for temperature kT and for some input density qIn. The first order difference between MP vectors $qOut$ evaluated for the same qIn, but at different temperatures kT and $kT - \delta kT$ is

$$\Delta Q_{\alpha\beta}^{\text{Direct}} \equiv \left(qOut_{kT-\delta kT} - qOut_{kT}\right)_{\alpha\beta} = 2\sum_{i=1}^{N}\left(-\partial_{kT} f(\varepsilon 0_i)\delta kT + \partial_\mu f(\varepsilon 0_i)\delta\mu\right) \times C0i_\alpha\, C0i_\beta. \quad (27)$$

The change of Fermi energy $\delta\mu$ is obtained from normalization condition similar to Eq. (20)

$$\sum_{i=1}^{N}\Delta f(\varepsilon 0_i) = -\sum_{i=1}^{N}\partial_{kT} f(\varepsilon 0_i)\delta kT + \sum_{i=1}^{N}\partial_\mu f(\varepsilon 0_i)\delta\mu = 0. \quad (28)$$

Substituting the expression for $\delta\mu$ in Eq. (27) we get

$$\Delta Q_{\alpha\beta}^{\text{Direct}} = -2\delta kT \sum_{i=1}^{N} \left[\partial_{kT} f(\varepsilon 0_i) - \partial_{\mu} f(\varepsilon 0_i) \frac{\sum_{j=1}^{N} \partial_{kT} f(\varepsilon 0_j)}{\sum_{j=1}^{N} \partial_{\mu} f(\varepsilon 0_j)} \right] \times C0i_{\alpha} C0i_{\beta}. \quad (29)$$

We assume that the linear approximation Eq. (29) remains valid as long as

$$0 \le f(\varepsilon 0_i, kT) + \partial_{kT} f(\varepsilon 0_i, kT)\delta kT \le 1$$

$$\forall \varepsilon 0_i \in \text{TSR} \equiv \left[\mu - KT, \mu + KT\right]. \quad (30a)$$

If $KT = \chi kT$, condition Eq. (30a) is equivalent to

$$|\delta kT| \le kT/\chi. \quad (30b)$$

If a temperature decrease is desired during a SC iteration, i.e. *qIn*, *qOut*, *C*0, and hence $\mathfrak{B}$ refers to temperature *kT*, but we want ΔqIn to be valid for $kT - \delta kT$, the SC condition is

$$qOut + \Delta Q^{\text{Direct}} + \mathfrak{B}\Delta qIn = qIn + \Delta qIn, \quad (31a)$$

and the analog of Eq. (4) is

$$(E - \mathfrak{B})\Delta qIn = qOut - qIn + \Delta Q^{\text{Direct}}. \quad (31b)$$

Temperature reduction becomes efficient only when the difference between *qIn* and *qOut* is small enough to make the scaling coefficient $\tau \ge 1$. Otherwise, reducing *kT* further reduces τ, and thus increases the number of iterations. Because τ is not *a priori* known, we need to solve Eq. (4) to obtain δH_n used in Eq. (24). Then, obtain the scaling coefficient τ, and if condition $\tau \ge 1$ is met decrease the temperature; if not proceed to the next iteration. Solution ΔqIn for Eq. (31b) can be viewed as a sum

$$\Delta qIn = \Delta qIn_{\text{SC}} + \Delta qIn_{\delta kT}, \quad (32)$$

where ΔqIn_{SC} is the solution of Eq. (4) accounting for driving *qIn* towards *qOut* at a constant temperature. The quantity $\Delta qIn_{\delta KT}$ is the solution of

$$(E - \mathfrak{B})\Delta qIn_{\delta kT} = \Delta Q^{\text{Direct}}. \quad (33)$$

The system of Eqs. (4) and (33) have the same left hand side. This allows using LU decomposition for solving Eq. (4), and then, if temperature reduction is possible, use only $O(N^2)$ extra flops to solve Eq. (33). Below is the algorithm summary.

(1) Choose an initial *KT* that determines the initial size of the TSR region, and initial *qIn*. We usually assume $qIn = 0$.
(2) For a given *qIn* obtain the SC and applied field corrections. Solve the eigenproblem to obtain *C*0, and $\varepsilon 0$.
(3) Partition the TSR region into smaller subregions to make the number of directly evaluated terms in Eq. (14b) reasonably small. Partition the rest of the energy spectrum in such a way that the interval size increases exponentially with the interval number (Fig. 2). Build the *ASym* matrix by using approximations Eqs. (9)–(11) to evaluate most of the terms in Eq. (14b). Terms accounting for $\varepsilon 0_m$ and $\varepsilon 0_i$ lying in the same or adjacent interval(s) are evaluated exactly.
(4) Evaluate the matrix *G* using Eq. (22).
(5) Evaluate the matrix $\mathfrak{B} = 2(A + G)U$. Matrix *U* can be precalculated before the first iteration, because it does not depend on *qIn*. The evaluation of matrix *U* is an $O(N^2)$ operation. To save memory space we choose to recalculate *U* during each iteration, rather than to keep it in memory.
(6) Solve system Eq. (4) for ΔqIn_{SC} using LU decomposition; keep *L* and *U* matrixes until step 7. Obtain δH and $\delta\varepsilon_i^{\text{Initial}}$ $\forall i :: \varepsilon 0_i \in \text{TSR}$ using Eq. (24). Obtain the scaling coefficient τ. If $\tau < 1$ use Eq. (26) and return to step 2. Otherwise proceed to step 7.
(7) If the inequality (25) is satisfied and temperature reduction is desired, decrease *kT* by no more than kT/χ. Obtain ΔQ^{Direct} from Eq. (27) and solve Eq. (33) for $\Delta qIn\delta_{kT}$ by using *L* and *U* matrixes from the previous step. Obtain the next step input density as a sum of *qIn*, ΔqIn_{SC} and $\Delta qIn_{\delta kT}$. Return to step 2 unless self-consistency and the required temperature are achieved.

Plotted in Fig. 3 are typical convergence rates for metallic and semiconducting systems. Typical numbers of iterations required to achieve 10^{-5} MP convergence is 2 for semiconductors and 10 for metals at $T = 300K$.

EXAMPLES: HYDROCARBON NANO-STRUCTURES IN APPLIED FIELD

To exemplify the method described in the previous sections we have calculated electron densities and Coulomb potentials for semiconducting and metallic (hydro)carbon structures in strong applied electric fields using an environment dependent tight binding (EDTB) methodology [16,17] combined with self-consistent (SC) corrections [13]. The typical number of carbon atoms in our simulations is about 450, and in all examples convergence starts with

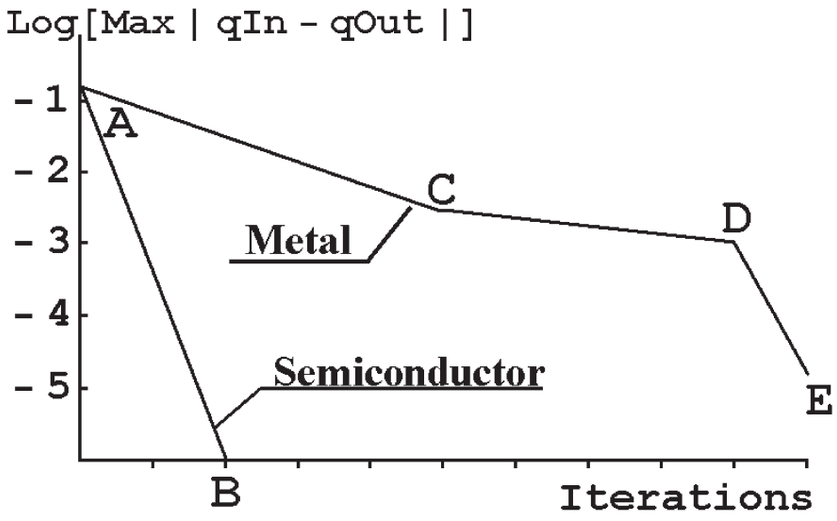

FIGURE 3 Illustration of typical SC-EDTB convergence. Line segment AC denotes convergence at constant (usually $kT = 0.25\,\text{eV}$) temperature and scaling coefficient $\tau < 1$. Once condition (25) is met at point *C*, the temperature reduction until $kT = 0.05$–$0.25\,\text{eV}$ is allowed. The final part DE corresponds to convergence at a constant low temperature and where the condition given by Eq. (25) is satisfied. Convergence for the semi-conducting system starts at low temperature. No temperature reduction is necessary in this case.

zero SC corrections ΔH. All simulations were done on a low-end workstation (500 MHz dual Pentium III Xeon).

Nano-diamond Cluster with Hydrogenated Surface

We start with consideration of a nano-diamond cluster represented by an octahedron with (111) facets, and with the top and the bottom vertices truncated by (100) planes (see insert in Fig. 6). The cluster is composed of 435 carbon atoms. Its entire surface is passivated with 196 hydrogen atoms. Hydrogen passivation removes any of the metallic character of a clean diamond surface, and makes the entire system semiconducting (Fig. 4(a)). In the case when no external field is applied there are no energy levels in the vicinity of the Fermi level, and convergence acceleration proceeds under the scenario described in "Improved Scaling for the Newton–Raphson Algorithm Section" where the aufbau principle (1c) with $kT = 0.05\,\text{eV}$ is used. Only two SC iterations are required to achieve self-consistency in this case; after the second iteration the maximum difference between the input and output MP's becomes less than 10^{-4}. Figure 5 illustrates the Coulomb potential distribution inside the cluster. The potential peaks at the cluster surface are due to the C–H dipole layer. The eigenenergy spectrum in Fig. 4(a) can be used to read the electron affinity (EA), i.e. the position of the conduction band edge with respect to the vacuum level taken with the opposite sign. For the hydrogen passivated (111) surface the experimentally measured EA is negative, and equals $-1.4\,\text{eV}$ [19]. That value coincides well with the position of the lowest conduction band level in Fig. 4(a).

When the cluster is placed in a $0.2\,\text{V/Å}$ external field, the applied potential sweep across the cluster is about 3.0 V. This is less than the cluster band gap and therefore the applied field penetrates inside the cluster (Fig. 6) with the field strength decreased by approximately a factor of 2. The field attenuation is smaller than the zero frequency diamond dielectric constant $\varepsilon_{\text{Diamond}} = 5.7$. This discrepancy is partially due to the poor polarizability of the minimal basis set [13], and partially due to the small cluster size. Because the system spectrum has an appreciable band gap (Fig. 4(b)), convergence may start at the system temperature $kT = 0.05\,\text{eV}$, and therefore no temperature reduction is required. Self-consistency is reached after four iterations.

If the potential variation across the cluster exceeds its band gap, the potential of the valence band edge at the high potential region rises up to the conductance band edge at the low potential region. This can be verified by Fig. 4(c), which shows that the energy spectrum for the entire system does not have a HOMO–LUMO gap if the applied field $E = 2.0\,\text{V/Å}$. Band tilt causes mixing between conduction and valence states, and leads to long-range charge redistribution. In extremely high fields the cluster acts almost like an equipotential metallic particle, because the potential variation across the cluster cannot exceed its band gap (Fig. 7). Convergence starts at $kT = 0.25\,\text{eV}$, and after 7 iterations a maximum deviation of 0.015 MP convergence is achieved for this temperature. Temperature

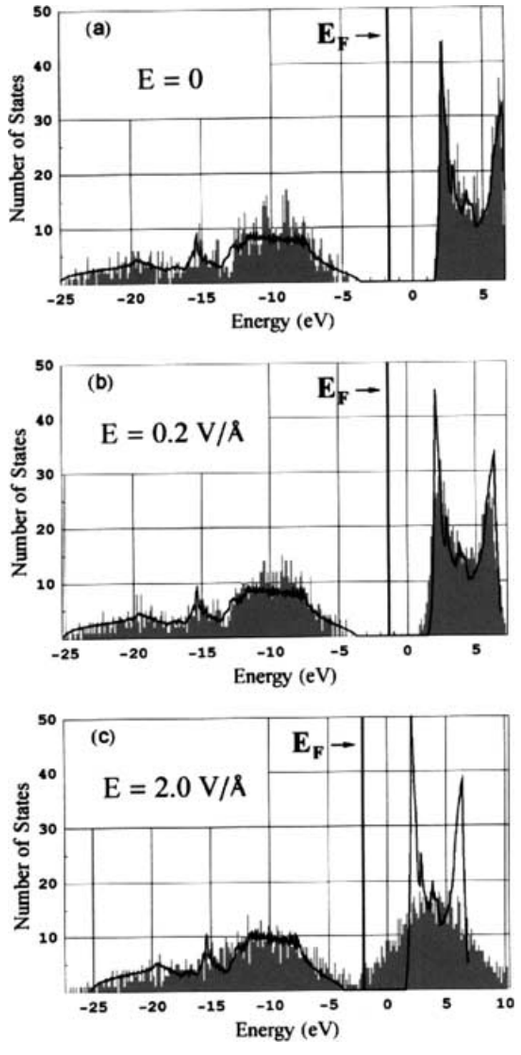

FIGURE 4 Eigenvalue spectra of a hydrogen passivated nano-diamond cluster in different applied field strengths. The solid line denotes the SC-EDTB spectrum for bulk diamond. The bulk spectrum is normalized to give the total number of states equal to the number of carbon valence electrons. The bulk spectrum was shifted by 1.4 eV, the average Coulomb potential experienced by carbon atoms in the cluster (cf. Fig. 5).

reduction to $kT = 0.05$ eV requires 5 extra iterations, and one iteration is needed to achieve 10^{-4} MP convergence at this low temperature. The number of iterations is above the average because of the extremely high applied field, which exceeds the field strength in the C–H dipole layer at the cluster surface (cf. Fig. 5).

Single Wall 9 × 0 Nano-tube

The following example demonstrates the applicability of the convergence acceleration method to a metallic system. A metallic 9×0 kinked fullerene nanotube was placed into a 0.2 V/Å electric field (Fig. 8); the field vector lies in the plane formed by two straight portions of the tube. The tube and the space inside it are equipotential. That is a well-known field screening effect featured by single wall metallic nano-tubes (e.g. [20]). The number of iterations needed to reach self-consistency is 10, which is slightly above the average due to the applied field.

Besides demonstrating field screening and moderate field enhancement at the kink region, this example is a good illustration of the question "why is *non*-equilibrium charge density is so important for quantum transport problems?" To calculate current through a mesoscopic structure connected to two conducting leads (Fig. 9(a)) the common practice [21–23] is to obtain the spectral density for this structure under equilibrium conditions. The current is then evaluated from the spectral density through a Green's functions formalism [21–25]. This scheme produces meaningful results for semiconducting systems, or when the resistance of the system is much higher than the resistance of the leads. The potential profile typical for that case is shown in Fig. 9(b). Suppose, however, that the "system" in Fig. 9(a) corresponds to the kink region of the tube in Fig. 8, and the leads are the straight portions of the kinked tube. Because the tube remains metallic in

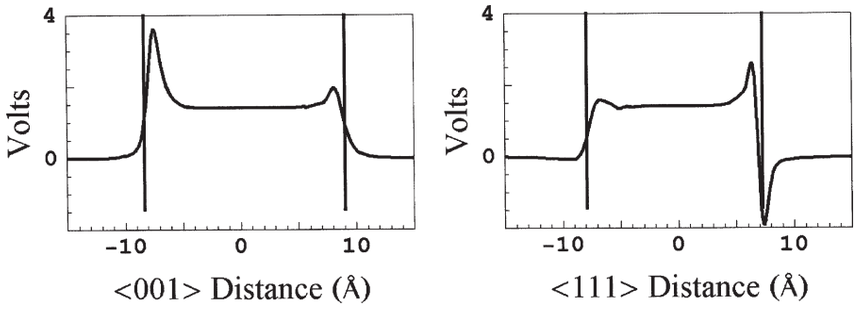

FIGURE 5 Coulomb potential profile along the ⟨001⟩ and ⟨111⟩ lines passing through the center of mass of the cluster. The vertical lines mark cluster facets.

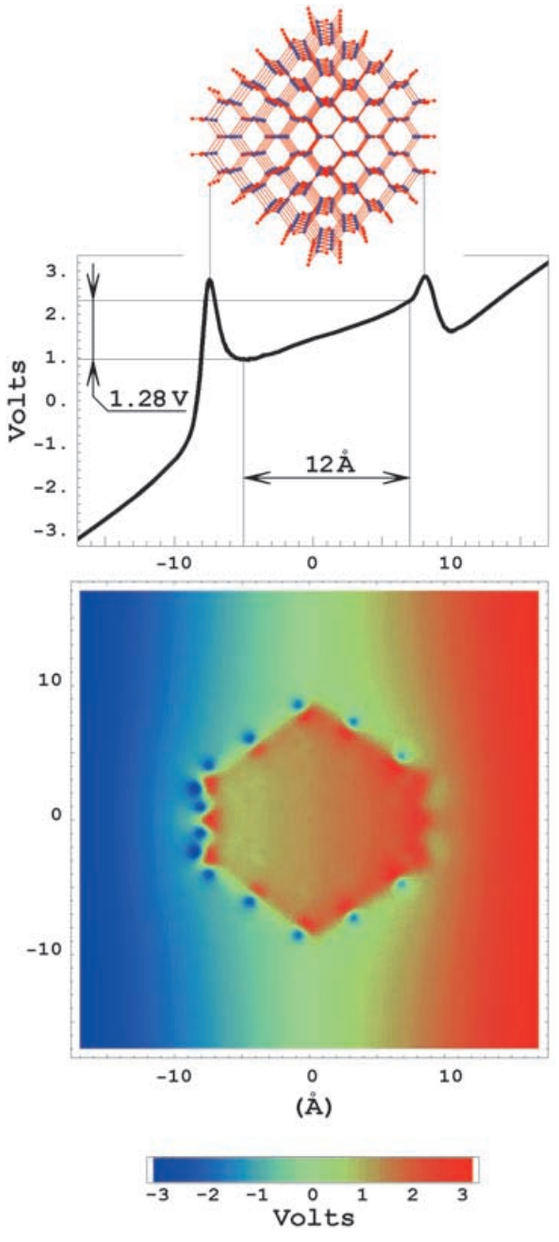

FIGURE 6 Electron potential in the vicinity and inside the cluster when 0.2 V/Å field is applied. *Top*: Hydrogen passivated cluster composed of 435 carbon and 196 hydrogen atoms. *Center*: Electron potential along the ⟨100⟩ line passing through the center of mass of the cluster and parallel to the applied field. *Bottom*: Electron potential in the (100) plane passing through the center of mass and oriented in the same way as the view plane of the top portion of the figure. The electric field vector lies in the view plane and directed from left to right.

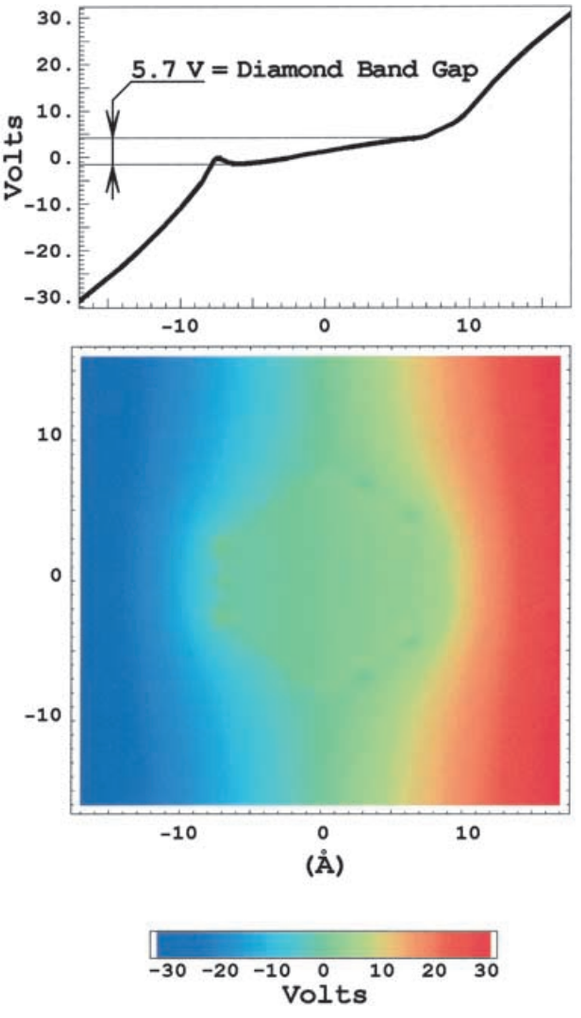

FIGURE 7 Same as Fig. 6 for 2.0 V/Å applied field.

the kink region, the equilibrium SC density calculation will result in the constant potential along the kink, and the non-physical voltage drop will occur only at the boundaries of the kink region. Schematically this potential profile is the same as shown in Fig. 9(b). The actual potential drop, however, occurs due to electron scattering at the kink, and has a single step structure shown in Fig. 9(c). This potential profile cannot be obtained by an equilibrium density calculation, because the scattering is essentially a non-equilibrium process.

CONCLUSIONS

We proposed a technique to improve the scaling of the Newton–Raphson algorithm for non-linear systems of equations. The improvement is based on decoupling the energy dependent term in first-order perturbation theory, and on partitioning the eigenvalue spectrum into exponentially increasing intervals. The Newton–Raphson algorithm with $O(N^3)$ scaling is most efficient for the sparse SC correction matrix, which implies an orthogonal basis set. The convergence technique is very

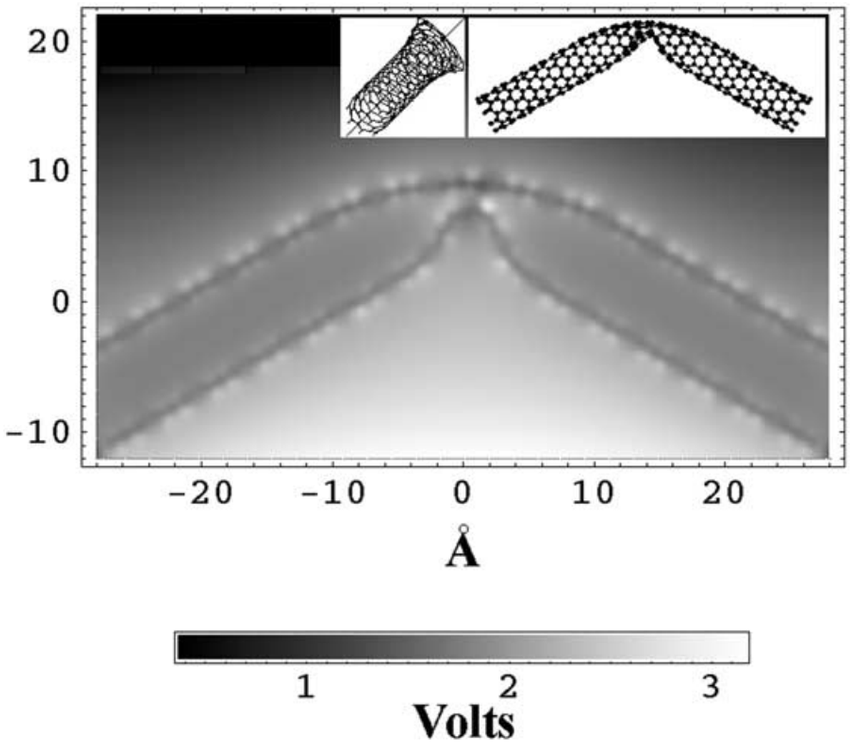

FIGURE 8 Electron potential in the cross-sectional plane passing through the axis of a kinked 9 × 0 single wall nano-tube when a 0.2 V/Å field is applied. The applied electric field vector lies in the view plane and directed downwards.

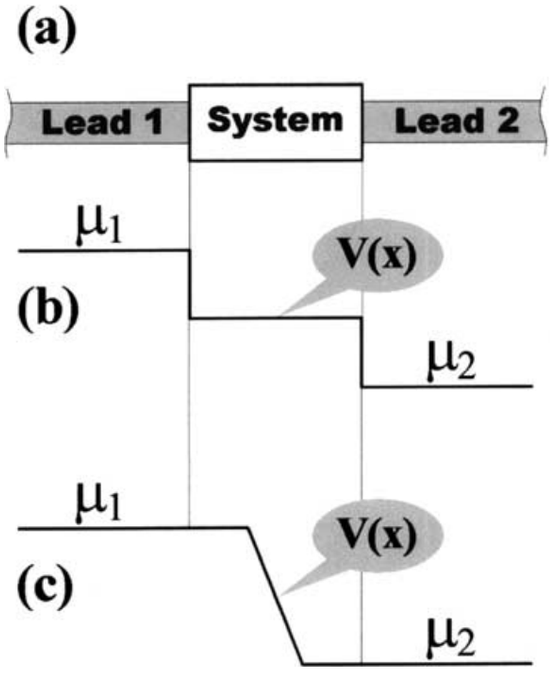

FIGURE 9 (a) Schematic layout considered in many transport problems. Only detailed description of the System is required, while Leads are described by adding the self-energies to the Hamiltonian of the System. The Leads have potentials μ_1 and μ_2. Suppose System is metallic and has one scattering center in the middle. The equilibrium approach does not allow potential variation inside the System. That results in the potential profile shown in portion (b). The true potential distribution (c), however, has only one step at the scattering region.

insensitive to the initial density, which may substantially differ from the SC configuration when the system is placed in a strong field. The technique can be efficiently applied both to metals and semiconductors.

A long-term goal of this work is to simulate *non-equilibrium* electron densities and hence the electric current in metallic or hybrid metal-semiconductor nano-structures for electronic device applications. The equilibrium density evaluation discussed here is not a subset of non-equilibrium problems, although as discussed above there are many issues common to both problems, and an equilibrium density calculation can be a convenient starting point for solving quantum transport problems. With proper changes, however, the Newton–Raphson algorithm can also be used for non-equilibrium SC density simulations, which is essential for transport problems in metallic systems.

Acknowledgements

This work was funded by the Office of Naval Research through a Multi-University Research Initiative.

APPENDIX: DERIVATION OF EQ. (5)

$$\Delta q_{\alpha\beta} = 2\sum_{i=1}^{N/2} C0i_\alpha (C0\,\varepsilon Inv_i\, C0^{\mathrm{T}}\ \delta H\, C0i)_\beta + 2\sum_{i=1}^{N/2} C0i_\beta \big(C0\varepsilon Inv_i C0^{\mathrm{T}}\,\delta H\, C0i\big)_\alpha. \quad (5)$$

We use "·" to explicitly denote a dot product. The perturbed Hamiltonian is denoted by H and its deviation from $H0$ by $\delta H = H - H0$. Matrixes having eigenvectors of H and $H0$ as their columns are denoted respectively as C and $C0$. The ith columns of these matrixes are denoted as Ci, and $C0i$, respectively, and the ζth components of these vectors are Ci_ζ and $C0i_\zeta$. $\varepsilon_i = \varepsilon 0_i + \delta\varepsilon_i$ stands for the ith eigenvalue of H. For the case of zero temperature and a non-degenerate HOMO, the change of Mulliken population component $q_{\alpha\beta}$ is

$$\Delta q_{\alpha\beta} = 2\sum_{i=1}^{N/2} Ci_\alpha\, Ci_\beta - 2\sum_{i=1}^{N/2} C0i_\alpha\, C0i_\beta \equiv 2\sum_{i=1}^{N/2} (C0i_\alpha + \Delta Ci_\alpha)\big(C0i_\beta + \Delta Ci_\beta\big) - 2\sum_{i=1}^{N/2} C0i_\alpha\, C0i_\beta. \quad (A.1)$$

Vector ΔCi can be represented as a linear combination of $C0j$. If vector $C0i$ corresponds to a degenerate level, there are some other $C0j$'s that belong to the same eigenenergy $\varepsilon 0_i$. We denote these vector indices as $\mathrm{Set}\{\varepsilon 0_i\}$. The quantity ΔCi can then be written as

$$\Delta Ci = \sum_{j=1,\, j\notin \mathrm{Set}\{\varepsilon 0_i\}}^{N} ai_j C0j + \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} bi_j C0j. \quad (A.2)$$

The coefficients ai_j and bi_j are of a different order of magnitude; $ai_j \sim \delta H \ll 1$ while $bi_j \sim 1$. Because any unitary transformation of $C0j$ also represents the equivalent subset of $H0$ eigenvectors, coefficients bi_j are not proportional to δH. Instead they are determined by symmetry considerations. Here is an example. Suppose $C01$ and $C02$ belong to the same eigenenergy. The small disturbance δH breaks the system symmetry and splits the originally degenerate state into two non-degenerate states, e.g.

$$C1 \to \frac{1}{\sqrt{2}}(C01 - C02) \quad \text{and} \quad C2 \to \frac{1}{\sqrt{2}}(C01 + C02).$$

The deviation of $C1$ from $C01$ is of the order of unity, and does not depend on the disturbance magnitude. Note that if $C01$ and $C02$ were initially chosen as given by Eq. (A.3), the disturbance δH would cause the split "along" $C01$ and $C02$. In that case the coefficients bi_j are small ($bi_j \sim \delta H^2$). Although it is convenient to have bi_j of the same order of magnitude as ai_j, it is possible to find the "right" combination of all degenerate level eigenvectors for only one particular δH. Because δH is not *a priori* known we cannot generally assume the smallness of bi_j. In the following, when performing the order of magnitude estimations we always assume $bi_j \sim 1$.

It is convenient to introduce the following notation:

- $\overline{abi}$: N-dimensional vector with ai_j on jth place if j does not belong to $\mathrm{Set}\{\varepsilon 0_i\}$ and bi_j if it does.
- $\overline{ai}$: N-dimensional vector with ai_j on jth place if j does not belong to $\mathrm{Set}\{\varepsilon 0_i\}$ and 0 if it does.
- $\overline{bi}$: N-dimensional vector with 0 on jth place if j does not belong to $\mathrm{Set}\{\varepsilon 0_i\}$ and bi_j if it does, i.e. $\Delta Ci = C0\cdot\overline{abi} = C0\cdot\overline{ai} + C0\cdot\overline{bi}$.
- $\overline{(1_i)}$: N-dimensional vector with 1 at jth positions and 0 at all others.

To obtain the expression for ΔCi we start with the equation for the ith eigenvector Ci of the disturbed Hamiltonian H:

$$(H0 + \delta H)\cdot(C0i + \Delta Ci) = (\varepsilon 0_i + \delta\varepsilon_i)(C0i + \Delta Ci). \quad (A.3)$$

Because $H0$ is symmetric, $C0$ is unitary and $C0\cdot C0^{\mathrm{T}} = E$. Multiplying the last equation in (A.3) from the left by $C0^{\mathrm{T}}$ and taking into account that $C0^{\mathrm{T}}\cdot H0 = \varepsilon 0\cdot C0^{\mathrm{T}}$ (here $\varepsilon 0$ is diagonal matrix composed of $H0$ eigenvalues) we obtain:

$$(\varepsilon 0 - E\varepsilon 0_i)\cdot\overline{abi} + C0^{\mathrm{T}}\cdot\delta H\cdot C0\cdot\overline{abi} = \delta\varepsilon_i C0^{\mathrm{T}}\cdot C0i + \delta\varepsilon_i\overline{abi} - C0^{\mathrm{T}}\cdot\delta H\cdot C0i. \quad (A.4)$$

Note that the matrix $(\varepsilon 0 - E\varepsilon 0_i)$ projects any vector $\overline{(1_j)}$ with index $j \in \mathrm{Set}\{\varepsilon 0_i\}$ to null space. Retaining only the first order of magnitude terms we can write Eq. (A.4) as

$$(\varepsilon 0 - E\varepsilon 0_i)\cdot\overline{ai} = \delta\varepsilon_i[\overline{(1_i)} + \overline{bi}] - C0^{\mathrm{T}}\cdot\delta H\cdot C0\cdot[\overline{(1_i)} + \overline{bi}]. \quad (A.5)$$

To proceed we need to have the equations that determine the coefficients $\{bi_1 \ldots bi_N\}$. If n is the degeneracy of the ith energy level, i.e. $n = \mathrm{Dim}(\mathrm{Set}\{\varepsilon 0i\})$, then for this given i there are n equations (A.5) with a zero left hand side (by definition $ai_j = 0$ if $j \in \mathrm{Set}\{\varepsilon 0i\}$). These n equations can be rewritten in the component form

$$0 = \big\{\big(C0^{\mathrm{T}}\cdot\delta H\cdot C0 - E\delta\varepsilon_i\big)\cdot\big[\overline{(1_i)} + \overline{bi}\,\big]\big\}_j \quad \forall j \in \mathrm{Set}\{\varepsilon 0_i\}. \quad (A.6a)$$

If we define $\overline{Bi} \equiv \overline{(1_i)} + \overline{bi}$ then Eq. (A.6a) can be transformed to

$$\sum_{k\in \mathrm{Set}\{\varepsilon 0_i\}} \left(C0j^{\mathrm{T}} \cdot \delta H \cdot C0k\right)\overline{Bi}_k - \delta\varepsilon_i \overline{Bi}_j = 0. \qquad \text{(A.6b)}$$

It is convenient to view Eq. (A.6b) as an eigenproblem. Non-zero coefficients Bi_k constitute an $n \times n$ matrix of eigenvectors for the symmetric matrix $D_{jk} \equiv C0j^{\mathrm{T}} \cdot \delta H \cdot C0k$ where $\{j, k\} \in \mathrm{Set}\{\varepsilon 0_i\}$. D is symmetric because δH is symmetric, and $C0i$ are real. Thus the Bi's satisfy the unitarity condition:

$$\sum_{k=1}^{N} Bi_k Bj_k = \delta_{ij} \quad \forall\{i, j\} \in \mathrm{Set}\{\varepsilon 0_i\}. \qquad \text{(A.7)}$$

Note that though k runs from 1 to N the actual number of non-zero terms in the sum (A.7) is n because $Bi_k = 0$ when $k \notin \mathrm{Set}\{\varepsilon 0_i\}$. Expression (A.7) will be used to transform (A.1) in the following way. First we rewrite Eq. (A.1) using Eq. (A.2)

$$\Delta q_{\alpha\beta} = 2\sum_{i=1}^{N/2} \left[C0\cdot\left(\overline{Bi} + \overline{ai}\right)\right]_\alpha \left[C0\cdot\left(\overline{Bi} + \overline{ai}\right)\right]_\beta - 2\sum_{i=1}^{N/2} \left[C0\cdot\overline{(1_i)}\right]_\alpha \left[C0\cdot\overline{(1_i)}\right]_\beta. \qquad \text{(A.8)}$$

Then let us show that

$$\sum_{i=1}^{N/2} \left[C0\cdot\overline{(1_i)}\right]_\alpha \left[C0\cdot\overline{(1_i)}\right]_\beta = \sum_{i=1}^{N/2} \left(C0\cdot\overline{Bi}\right)_\alpha \left(C0\cdot\overline{Bi}\right)_\beta. \qquad \text{(A.9)}$$

To do so we represent the sum over i (occupied states) as a double sum over all different occupied energy levels, and over all states that belong to the same degenerate energy level.

$$\begin{aligned}
&\sum_{i=1}^{N/2} (C0\cdot\overline{Bi})_\alpha (C0\cdot\overline{Bi})_\beta \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} (C0\cdot\overline{Bj})_\alpha (C0\cdot\overline{Bj})_\beta \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} \left(\sum_{k\in \mathrm{Set}\{\varepsilon 0_i\}} C0k_\alpha \overline{Bj}_k\right)\left(\sum_{n\in \mathrm{Set}\{\varepsilon 0_i\}} C0n_\beta \overline{Bj}_n\right) \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{k\in \mathrm{Set}\{\varepsilon 0_i\}} C0k_\alpha \sum_{n\in \mathrm{Set}\{\varepsilon 0_i\}} C0n_\beta \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} \overline{Bj}_k \overline{Bj}_n \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} C0k_\alpha \sum_{n\in \mathrm{Set}\{\varepsilon 0_i\}} C0n_\beta \delta_{kn}. \qquad \text{(A.10)}
\end{aligned}$$

The last line in Eq. (A.10) was obtained using Eq. (A.7). After summation over either k or n, and returning back to the single index summation over occupied states we obtain the left hand side of Eq. (A.9). Replacing the second sum in Eq. (A.8) by the right hand side of Eq. (A.9) and retaining only the first order of magnitude terms we obtain:

$$\Delta q_{\alpha\beta} = 2\left[\sum_{i=1}^{N/2} (C0\cdot\overline{Bi})_\alpha (C0\cdot\overline{ai})_\beta + \sum_{i=1}^{N/2} (C0\cdot\overline{ai})_\alpha (C0\cdot\overline{Bi})_\beta\right]. \qquad \text{(A.11)}$$

The next step is to get the expression for $\overline{ai}_j$ from Eq. (A.5). By definition if $j \in \mathrm{Set}\{\varepsilon 0_i\}\overline{ai}_j = 0$. For these j's both sides of Eq. (A.5) are zeros. All other $\overline{ai}_j$ can be obtained as

$$\overline{ai}_j = \left(\varepsilon Inv_i \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0\cdot\overline{Bi}\right)_j, \qquad \text{(A.12)}$$

where we introduced the diagonal matrix εInv_i which has $(\varepsilon 0_i - \varepsilon 0_j)^{-1}$ at the jth position if $\varepsilon 0_i \neq \varepsilon 0_j$, and 0 otherwise. Switching to a double summation and taking into account Eq. (A.12), the first sum in Eq. (A.11) can be transformed as follows

$$\begin{aligned}
&\sum_{\text{Different}\,\varepsilon 0_i} \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} \left(C0\cdot\overline{Bj}\right)_\alpha \left(C0\cdot\varepsilon Inv_i \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0\cdot\overline{Bj}\right)_\beta \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} \left(\sum_{k\in \mathrm{Set}\{\varepsilon 0_i\}} C0k_\alpha \overline{Bj}_k\right) \times \left[C0\cdot\varepsilon Inv_i \cdot C0^{\mathrm{T}} \cdot \delta H \cdot \left(\sum_{n\in \mathrm{Set}\{\varepsilon 0_i\}} C0n\overline{Bj}_n\right)\right]_\beta \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{k\in \mathrm{Set}\{\varepsilon 0_i\}} \sum_{n\in \mathrm{Set}\{\varepsilon 0_i\}} C0k_\alpha \times \left(C0\cdot\varepsilon Inv_i \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0n\right)_\beta \sum_{j\in \mathrm{Set}\{\varepsilon 0_i\}} \overline{Bj}_k \overline{Bj}_n \\
&= \sum_{\text{Different}\,\varepsilon 0_i} \sum_{k\in \mathrm{Set}\{\varepsilon 0_i\}} C0k_\alpha \left(C0\cdot\varepsilon Inv_i \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0k\right)_\beta \\
&= \sum_{i=1}^{N/2} C0i_\alpha \left(C0\cdot\varepsilon Inv_i \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0i\right)_\beta. \qquad \text{(A.13)}
\end{aligned}$$

In the second to last step we used Eq. (A.7). The second sum in Eq. (A.11) is handled in the similar manner. The final expression Eq. (5) can be

obtained after substitution of Eq. (A.13) into Eq. (A.11).

Finally, we derive the equation that relates matrix δH to the shift of the ith eigenstate energy level. We take the dot product of vector $\overline{Bj}$ and the expression in curly parenthesis in Eq. (A.6a). For $j \neq i$ we get the identity $0 = 0$ because vectors $\overline{Bi}$ and $\overline{Bj}$ are eigenvectors of matrix $D_{j,k} \equiv C0j^{\mathrm{T}} \cdot \delta H \cdot C0k$ ($j, k \in \mathrm{Set}\{\varepsilon 0_i\}$), and due to the orthonormality condition (A.7). If $j = i$ we obtain

$$\delta\varepsilon_i = \left[\overline{(1_i)} + \overline{bi}\right]^{\mathrm{T}} \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0 \cdot \left[\overline{(1_i)} + \overline{bi}\right]. \quad \text{(A.14)}$$

If the level degeneracy is unity, then Eq. (A.14) transforms to well known expression for the first order energy correction:

$$\delta\varepsilon_i = (1_i)^{\mathrm{T}} \cdot C0^{\mathrm{T}} \cdot \delta H \cdot C0 \cdot \overline{(1_i)} \equiv C0i^{\mathrm{T}} \cdot \delta H \cdot C0i. \quad \text{(A.15)}$$

References

[1] Pulay, P. (1982) "Improved SCF convergence acceleration", *J. Comput. Chem.* **3**, 556.

[2] Defranceshci, M. and Le Bris, C. (2000), Chapter 2, *Mathemaical Models and Methods for* ab-initio *Quantum Chemistry* (Springer-Verlag, Berlin).

[3] Cancès, E. and Le Bris, C. (2000) "Can we outperform the DIIS approach for electronic structure calculations?", *Int. J. Quantum Chem.* **79**, 82.

[4] Fischer, T.H. and Almlöf, J. (1992) "General methods for geometry and wave function optimization", *J. Chem. Phys.* **96**, 9768.

[5] Chaban, G., Schmidt, M.W. and Gordon, M.S. (1997) "Approximate second order method for orbital optimization of SCF and MSCF wavefunctions", *Theor. Chem. Acc.* **97**, 88.

[6] Saunders, V.R. and Hillier, I.H. (1973) "A 'level shifting' method for converging closed shell Hartree-Fock wave functions", *Int. J. Quantum Chem.* **7**, 699.

[7] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) *Numerical Recipes in C. The Art of Scientific Computing*, Second Edition (Cambridge University Press, New york, NY).

[8] Brown, T.H. (1968) "Quadratically convergent iteration procedure for self-consistent calculations", *J. Chem. Phys.* **49**, 2291.

[9] Broyden, C.G. (1965) "A class of methods for solving nonlinear simultaneous equations", *Math. Computation* **19**, 577.

[10] Vanderbilt, D. and Louie, S.G. (1984) "Total energies of diamond (111) surface reconstructions by a linear combination of atomic orbitals method", *Phys. Rev. B* **30**, 6118.

[11] Ordejon, P., Artacho, E. and Soler, J.M. (1996) "Self-consistent order-N density-functional calculations for very large systems", *Phys. Rev. B* **53**, 10441.

[12] Mauri, F. and Galli, G. (1994) "Electronic-structure calculations and molecular-dynamics simulations with linear system-size scaling", *Phys. Rev. B* **50**, 4316.

[13] Areshkin, D.A., Shenderova, O.A., Schall, J.D. and Brenner, D.W. (Unpublished) "Self-consistent tight binding adapted for large metallic hydro-carbon systems. Application to field emission from nano-diamond clusters".

[14] Brenner, D.W., Shenderova, O.A., Areshkin, D.A. and Schall, J.D. (2002) "Atomic modeling of carbon-based nanostructures as a tool for developing new materials and technologies", *Computer Modeling Engng. Sci.* **3**, 643.

[15] Bernholc, J., Brenner, D., Nardelli, M., Buongiorno Meunier, V. and Roland, C. (2002) "Mechanical and electrical properties of nanotubes", *Ann. Rev. Mat. Res.* **32**, 3476.

[16] Tang, M.S., Wang, C.Z., Chan, C.T. and Ho, K.M. (1996) "Environment-dependent tight-binding potential model", *Phys. Rev. B* **53**, 979.

[17] Tang, M.S., Wang, C.Z., Chan, C.T. and Ho, K.M. (1996) "Environment-dependent tight-binding potential model", *Phys. Rev. B* **54**, 10982.

[18] Harris, J. (1985) "Simplified method for calculating the energy of weekly interacting fragments", *Phys. Rev. B* **31**, 1770.

[19] Ristein, J. (2000) "Electronic properties of diamond surfaces—blessing or curse for devices?", *Diamond Relat. Mater.* **9**, 1129.

[20] Lou, L., Nordlander, P. and Smalley, R.E. (1995) "Fullerene nanotubes in electric fields", *Phys. Rev. B* **52**, 1429.

[21] Derosa, P.A. and Seminario, J.M. (2000) "Electron transport through single molecules: scattering treatment using density functional and Green function theories", *J. Chem. Phys.* **105**, 471.

[22] Bernholc, J., *et al.* (2000) "Large-scale applications of real-space multigrid methods to surfaces, nanotubes, and quantum transport", *Phys. Status Solidi B* **217**, 685.

[23] Roland, C., Nardelli, M.B., Wang, J. and Guo, H. (2000) "Dynamic conductance of carbon nanotubes", *Phys. Rev. Lett.* **84**, 2921.

[24] Datta, S. (2000) "Nanoscale device modeling: the Green's function method", *Superlattices Microstruct.* **28**, 253.

[25] Datta, S. (1995) *Electronic transport in mesoscopic systems* (Cambridge University Press, New york, NY).