Efficient wave-function matching approach for quantum transport calculations

Hans Henrik B. Sørensen and Per Christian Hansen
Informatics and Mathematical Modelling, Technical University of Denmark, Bldg. 321, DK-2800 Lyngby, Denmark

Dan Erik Petersen, Stig Skelboe, and Kurt Stokbro
Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen, Denmark
(Received 5 May 2008; published 29 May 2009)

The wave-function matching (WFM) technique has recently been developed for the calculation of electronic transport in quantum two-probe systems. In terms of efficiency it is comparable to the widely used Green's function approach. The WFM formalism presented so far requires the evaluation of all the propagating and evanescent bulk modes of the left and right electrodes in order to obtain the correct coupling between device and electrode regions. In this paper we will describe a modified WFM approach that allows for the exclusion of the vast majority of the evanescent modes in all parts of the calculation. This approach makes it feasible to apply iterative techniques to efficiently determine the few required bulk modes, which allows for a significant reduction of the computational expense of the WFM method. We illustrate the efficiency of the method on a carbon nanotube field-effect-transistor device displaying band-to-band tunneling and modeled within the semiempirical extended Hückel theory framework.

DOI: 10.1103/PhysRevB.79.205322 PACS number(s): 73.40.−c, 73.63.−b, 72.10.−d, 85.35.Kt

I. INTRODUCTION

Quantum transport simulations have become an important theoretical tool for investigating the electrical properties of nanoscale systems.[1–5] The basis for the approach is the Landauer-Büttiker picture of coherent transport, where the electrical properties of a nanoscale constriction are described by the transmission coefficients of a number of one-electron modes propagating coherently through the constriction. The approach has been used successfully to describe the electrical properties of a wide range of nanoscale systems, including atomic wires, molecules, and interfaces.[6–15] In order to apply the method to semiconductor device simulation, it is necessary to handle systems comprising many thousand atoms, and this will require new efficient algorithms for calculating the transmission coefficient.

Our main purpose in this paper is to give details of a method we have developed based on the wave-function matching (WFM) technique,[16–18] which is suitable for studying electronic transport in large-scale atomic two-probe systems, such as large carbon nanotubes or nanowire configurations.

We adopt the many-channel formulation of Landauer and Büttiker to describe electron transport in nanoscale two-probe systems composed of a left and a right electrode attached to a central device (see Fig. 1). In this formulation, the conduction $\mathcal{G}$ of incident electrons through the device is intuitively given in terms of transmission and reflection matrices, $\mathbf{t}$ and $\mathbf{r}$, that satisfy the unitarity condition $\mathbf{t}^\dagger\mathbf{t}+\mathbf{r}^\dagger\mathbf{r}=\mathbf{1}$ in the case of elastic scattering. The matrix element t_{ij} is the probability amplitude of an incident electron in a mode i in the left electrode being scattered into a mode j in the right electrode, and correspondingly r_{ik} is the probability of it being reflected back into mode k in the left electrode. This simple interpretation yields the Landauer-Büttiker formula[3]

$$\mathcal{G} = \frac{2e^2}{h}\mathrm{Tr}[\mathbf{t}^\dagger\mathbf{t}], \tag{1}$$

which holds in the limit of infinitesimal voltage bias and zero temperature.

To our knowledge, the WFM schemes presented so far in the literature require the evaluation of all the Bloch and evanescent bulk modes of the left and right electrodes in order to obtain the correct coupling between device and electrode regions. The reason for this is that the complete set of bulk modes is needed to be able to represent the proper reflected and transmitted wave functions. In this paper we will describe a modified WFM approach that allows for the exclusion of the vast majority of the evanescent modes in all parts of the calculation. The primary modification can be pictured as a simple extension of the central region with a few principal electrode layers. In this manner, it becomes advantageous to apply iterative techniques for obtaining the relatively few Bloch modes and slowly decaying evanescent modes that are required. We have recently developed such an iterative method in Ref. 19, which allows for an order of magnitude reduction of the computational expense of the WFM method in practice.

In this work, the proper analysis of the modified WFM approach is presented. The accuracy of the method is inves-

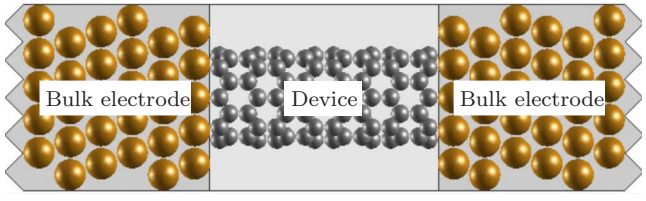

FIG. 1. (Color online) Schematic illustration of a nanoscale two-probe system in which a device is sandwiched between two semi-infinite bulk electrodes.

tigated and appropriate error estimates are developed. As an illustration of the applicability of our WFM scheme we consider a 1440 atom carbon nanotube field-effect transistor (CNTFET) device of 14 nm in length. We calculate the zero-bias transmission curves of the device under various gate voltages and reproduce previously established characteristics of band-to-band tunneling.[20] We compare directly the results of the modified WFM method to those of the standard WFM method for quantitative verification of the calculations.

The rest of the paper is organized as follows. The WFM formalism that is used to obtain **t** and **r** is introduced in Sec. II. In Sec. III we present our method to effectively exclude the rapidly decaying evanescent modes from the two-probe transport calculations. Numerical results are presented in Sec. IV and the paper ends with a short summary and outlook.

II. FORMALISM

In this section we give a minimal review of the formalism and notation that is used in the current work in order to determine the transmission and reflection matrices **t** and **r**. This WFM technique has several attractive features compared to the widely used and mathematically equivalent Green's function approach.[1,2] Most importantly, the transparent Landauer picture of electrons scattering via the central region between Bloch modes of the electrodes is retained throughout the calculation. Moreover, WFM allows one to consider the significance of each available mode individually in order to achieve more efficient numerical procedures to obtain **t** and **r**.

A. Wave-function matching

The WFM method is based on direct matching of the bulk modes in the left and right electrodes to the scattering wave function of the central region. For the most part this involves two major tasks: obtaining the bulk electrode modes and solving a system of linear equations. The bulk electrode modes can be characterized as either propagating or evanescent (exponentially decaying) modes but only the propagating modes contribute to $\mathcal{G}$ in Eq. (1). We may write $\mathcal{G}=(2e^2/h)T$, where

$$T=\sum_{kk'}|t_{kk'}|^2 \tag{2}$$

is the total transmission and the sum is limited to propagating modes k and k' in the left and right electrodes, respectively. Notice, however, that the evanescent modes are still needed in order to obtain the correct matrix elements $t_{kk'}$. We will discuss this matter in Sec. III C.

We assume a tight-binding setup for the two-probe systems in which the infinite structure is divided into principal layers numbered $i=-\infty,\ldots,\infty$ and composed of a finite central (C) region containing the device and two semi-infinite left (L) and right (R) electrode regions (see Fig. 2). The wave function is $\boldsymbol{\psi}_i(\boldsymbol{x})=\Sigma_j^{m_i}c_{i,j}\chi_{i,j}(\boldsymbol{x}-\mathbf{X}_{i,j})$ in layer i, where $\chi_{i,j}$ denotes localized nonorthogonal atomic orbitals and $\mathbf{X}_{i,j}$ are the positions of the m_i orbitals in layer i. We represent $\boldsymbol{\psi}_i(\boldsymbol{x})$ by

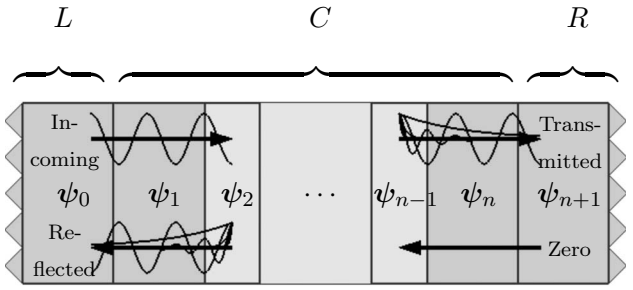

FIG. 2. Schematic representation of WFM applied to layered two-probe systems, where the central device region, consisting of layers $i=1,\ldots,n$, is attached to left and right semi-infinite electrodes. The incoming propagating mode from the left electrode is scattered in the central region and ends up as reflected and transmitted superpositions of propagating and evanescent modes.

a column vector of the expansion coefficients, given by $\boldsymbol{\psi}_i=[c_{i,1},\ldots,c_{i,m_i}]^T$, and write the wave function $\boldsymbol{\psi}$ extending over the entire system as $\boldsymbol{\psi}=[\boldsymbol{\psi}_{-\infty}^T,\ldots,\boldsymbol{\psi}_{\infty}^T]^T$. We also assume that the border layers 1 and n of the central region are always identical to a layer of the connecting electrodes.

We refer the reader to Refs. 16–18 and 21 for details on how to employ WFM to our setup. Here and in the rest of this paper, we will use the following notation for the key elements. The matrices $\boldsymbol{\Phi}_L^{\pm}=[\boldsymbol{\phi}_{L,1}^{\pm},\ldots,\boldsymbol{\phi}_{L,m_L}^{\pm}]$ contain in their columns the full set of m_L left-going (−) and m_L right-going (+) bulk modes $\boldsymbol{\phi}_{L,k}^{\pm}$ of the left electrode, and the diagonal matrices $\boldsymbol{\Lambda}_L^{\pm}=\mathrm{diag}[\lambda_{L,1}^{\pm},\lambda_{L,2}^{\pm},\ldots,\lambda_{L,m_L}^{\pm}]$ hold the corresponding Bloch factors.[22] If trivial modes with $|\boldsymbol{\phi}_{L,k}^{+}|=0$ or $|\boldsymbol{\phi}_{L,k}^{-}|=\infty$ occur they are simply rejected. We assume that all the evanescent bulk modes are (state-) normalized $\boldsymbol{\phi}_{L,k}^{\pm\dagger}\boldsymbol{\phi}_{L,k}^{\pm}=1$, while all the Bloch bulk modes are flux normalized[23] $\boldsymbol{\phi}_{L,k}^{\pm\dagger}\boldsymbol{\phi}_{L,k}^{\pm}=d_L/v_{L,k}^{\pm}$, where $v_{L,k}^{\pm}$ are the group velocities[15,24] and d_L is the layer thickness. Similarly for the right electrode the matrices $\boldsymbol{\Phi}_R^{\pm}$ and $\boldsymbol{\Lambda}_R^{\pm}$ are formed.

We also introduce the Bloch matrices[17] $\mathbf{B}_L^{\pm}=\boldsymbol{\Phi}_L^{\pm}\boldsymbol{\Lambda}_L^{\pm}(\boldsymbol{\Phi}_L^{\pm})^{-1}$ and $\mathbf{B}_R^{\pm}=\boldsymbol{\Phi}_R^{\pm}\boldsymbol{\Lambda}_R^{\pm}(\boldsymbol{\Phi}_R^{\pm})^{-1}$, which propagate the layer wave functions in the bulk electrode

$$\boldsymbol{\psi}_j^{\pm}=(\mathbf{B}^{\pm})^{j-i}\boldsymbol{\psi}_i^{\pm}, \tag{3}$$

where subscript L is implied for the left electrode $(i,j\leq 1)$ and R for the right electrode $(i,j\geq n)$. Notice that the first central region layer is defined for layer 1 and not layer 0, as is the case in Ref. 18.

As explicitly shown in Refs. 16–18, by fixing the layer wave functions coming into the C region (e.g., in our case $\boldsymbol{\psi}_1^{+}=\lambda_{L,k}^{+}\boldsymbol{\phi}_{L,k}^{+}$ and $\boldsymbol{\psi}_n^{-}=\mathbf{0}$) and matching the layer wave functions across the C region boundaries, the system of linear equations for the central region wave function $\boldsymbol{\psi}_C$ can be written as

$$(E\mathbf{S}_C-\mathbf{H}_C-\boldsymbol{\Sigma}_L-\boldsymbol{\Sigma}_R)\boldsymbol{\psi}_C=\mathbf{b}, \tag{4}$$

where E is the energy, $\mathbf{S}_C$ the overlap, and $\mathbf{H}_C$ the Hamiltonian matrix of the central region. In the following we discuss the terms, $\boldsymbol{\Sigma}_L$, $\boldsymbol{\Sigma}_R$, and $\mathbf{b}$, which arise from matching the boundary conditions with the electrode modes.

TABLE I. CPU times in seconds when using WFM for calculating $\mathbf{t}$ and $\mathbf{r}$ at 20 different energies inside $E \in [-2\ \text{eV}; 2\ \text{eV}]$ for various two-probe systems. The numbers of atoms in the central region (electrode unit cell) are indicated. The four rightmost columns show the CPU times spent for computing the electrode bulk modes with DGEEV and in this work vs solving the central region linear systems in Eq. (4) and the system with two extra principal layers on each side.

System	Atoms	Equation (4)	Equation (4) ($l=2$)	DGEEV	This work
Fe-MgO-Fe	27(6)	0.8	0.9	1.3	1.1
Al-C×7-Al	74(18)	0.4	0.6	3.6	1.6
Au-DTB-Au	102(27)	8.1	13.5	91.0	28.2
Au-CNT(8,0)×1-Au	140(27)	11.4	16.6	77.6	17.1
Au-CNT(8,0)×5-Au	268(27)	45.3	50.3	83.6	17.8
CNT(8,0)-CNT(8,0)	192(64)	7.0	11.9	129.0	19.4
CNT(4,4)-CNT(8,0)	256(64|64)	7.2	12.4	121.5	21.0
CNT(5,0)-CNT(10,0)	300(40|80)	24.7	31.5	113.3	22.6
CNT(18,0)-CNT(18,0)	576(144)	172.2	225.5	1362.2	253.3
CNTFET (see Fig. 6)	1440(160)	259.8	286.9	4633.0	372.3

The self-energy matrices, $\mathbf{\Sigma}_L$ and $\mathbf{\Sigma}_R$, arise from matching with the outgoing left and right electrode modes. They only have nonzero terms in the upper left and lower right corner blocks, respectively, and these elements can be calculated in terms of the Bloch matrices[16,17]

$$[\mathbf{\Sigma}_L]_{1,1} = \overline{\mathbf{H}}_{0,1}^{\dagger}[\overline{\mathbf{H}}_1 + \overline{\mathbf{H}}_{0,1}^{\dagger}(\mathbf{B}_L^-)^{-1}]^{-1}\overline{\mathbf{H}}_{0,1} \tag{5}$$

and

$$[\mathbf{\Sigma}_R]_{n,n} = \overline{\mathbf{H}}_{n,n+1}(\overline{\mathbf{H}}_n + \overline{\mathbf{H}}_{n,n+1}\mathbf{B}_R^+)^{-1}\overline{\mathbf{H}}_{n,n+1}^{\dagger}, \tag{6}$$

where we have introduced the overline notation $\overline{\mathbf{H}}_i \equiv E\mathbf{S}_i - \mathbf{H}_i$ and $\overline{\mathbf{H}}_{i,j} \equiv E\mathbf{S}_{i,j} - \mathbf{H}_{i,j}$. For the current setup, these matrices are *identical* to the self-energy matrices introduced in the Green's function formalism[1] (to within an infinitesimal imaginary shift of E) and may be evaluated by well-known recursive techniques[25,26] or constructed directly from the electrode modes using Eq. (6).

The source term $\mathbf{b}$ arises from the incoming mode. Assuming an incoming mode from the left, we have $\mathbf{b} = [\mathbf{b}_1^T, \mathbf{0}^T, \ldots, \mathbf{0}^T]^T$ specified by the expression

$$\mathbf{b}_1 = -(\overline{\mathbf{H}}_{0,1}^{\dagger} + [\mathbf{\Sigma}_L]_{1,1}\mathbf{B}_L^+)\boldsymbol{\psi}_0, \tag{7}$$

where $\boldsymbol{\psi}_0$ is the incoming wave function.

For notational simplicity in the following sections, we leave out the implied subscripts L or R, indicating the left or right electrode, whenever the formalism is the same for both (e.g., for symbols $m, \lambda_k, \boldsymbol{\phi}_k, \mathbf{\Phi}^{\pm}, \mathbf{\Lambda}^{\pm}, \mathbf{B}^{\pm}, \mathbf{\Sigma}$, etc.).

B. Transmission and reflection coefficients

As a final step we want to determine the $\mathbf{t}$ and $\mathbf{r}$ matrices from the boundary wave functions $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_n$ that have been obtained by solving Eq. (4). When the incoming wave $\boldsymbol{\psi}_0$ is specified to be the kth right-going mode $\boldsymbol{\phi}_{L,k}^+$ of the left electrode, then $\boldsymbol{\psi}_n$ will be the superposition of outgoing right-transmitted waves. The kth column of the transmission matrix $\mathbf{t}_k$ is defined as the corresponding expansion coefficients in right electrode modes and can be evaluated by solving

$$\mathbf{\Phi}_R^+ \mathbf{t}_k = \boldsymbol{\psi}_n, \tag{8}$$

where $\mathbf{\Phi}_R^+$ is the $m_R \times m_R$ column matrix holding the right-going bulk modes of the right electrode (and here assumed to be nonsingular). Similarly the kth column of the reflection matrix $\mathbf{r}_k$ is given by

$$\mathbf{\Phi}_L^- \mathbf{r}_k = \boldsymbol{\psi}_1 - \lambda_{L,k}^+ \boldsymbol{\phi}_{L,k}^+, \tag{9}$$

where $\mathbf{\Phi}_L^-$ holds the left-going bulk modes of the left electrode. The flux normalization ensures that $\mathbf{t}^{\dagger}\mathbf{t} + \mathbf{r}^{\dagger}\mathbf{r} = \mathbf{1}$.

III. EXCLUDING EVANESCENT MODES

The most time-consuming task of the WFM method is often to determine the electrode modes, which requires solving a quadratic eigenvalue problem.[16] As examples, see the profiling results listed in Table I, where we have used the method to compute $\mathbf{t}$ and $\mathbf{r}$ for a selection of two-probe systems.[27] The CPU timings show that to determine the electrode modes by employing the state-of-the-art LAPACK eigensolver DGEEV is, in general, much more expensive than to solve the system of linear equations in Eq. (4). We expect this trend to hold for larger systems as well. Therefore, in the attempt to model significantly larger devices (thousands of atoms), it is of essential interest to reduce the numerical cost of the electrode modes calculation. We argue that a computationally reasonable approach is to limit the number of electrode modes taken into account, e.g., by excluding the least important evanescent modes. In this section, a proper technique to do this in a rigorous and systematic fashion is presented.

A. Decay of evanescent modes

The procedure to determine the Bloch factors λ_k and nontrivial modes $\boldsymbol{\phi}_k$ of an ideal electrode and subsequently char-

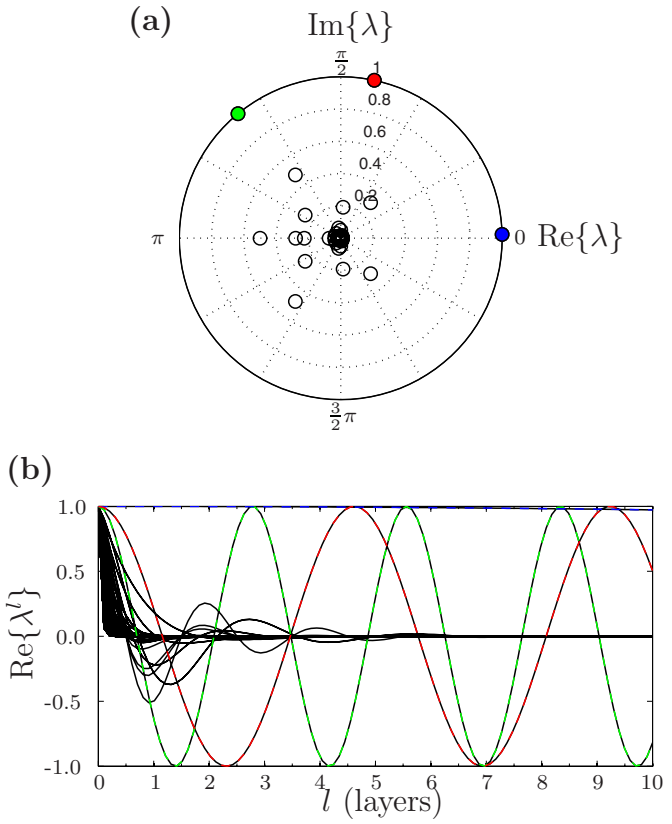

FIG. 3. (Color online) (a) Positions of the Bloch factors $\lambda_k(|\lambda_k| \leq 1)$ obtained for a bulk Au(111) electrode with 27 atoms per unit cell at $E=-1.5$ eV. (b) Amplitudes of the corresponding normalized electrode modes $\boldsymbol{\phi}_k$ moving through ten layers of the ideal bulk electrode. A total of 243 modes are shown of which three are propagating (colored/dashed) and the rest are evanescent (circles/black).

acterize these as right-going (+) or left-going (−) is well described in the literature.[16–18,28] We note that only the obtained propagating modes with $|\lambda_k|=1$ are able to carry charge deeply into the electrodes and thus enter the Landauer expression in Eq. (2). The evanescent modes with $|\lambda_k| \neq 1$, on the other hand, decay exponentially but can still contribute to the current in a two-probe system, as the "tails" may reach across the central region boundaries.

Consider a typical example of an electrode modes evaluation. We look at a gold electrode with 27 atoms in the unit cell represented by 9 (sp^3d^5) orbitals for each Au atom. Such a system results in 243 right-going and 243 left-going modes. Figure 3(a) shows the positions in the complex plane of the Bloch factors corresponding to the right-going modes (i.e., $|\lambda_k| \leq 1$) for energy $E=-1.5$ eV. We see that there are exactly three propagating modes which have Bloch factors located on the unit circle. The remaining modes are evanescent, of which many have Bloch factors with small magnitude very close to the origin.

Figure 3(b) illustrates how the 243 left-going modes would propagate through ten successive gold electrode unit cells. The figure shows that the amplitudes of the three propagating modes are unchanged, while the evanescent modes are decaying exponentially. In particular, we note that

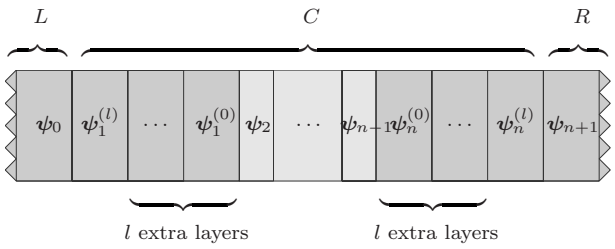

FIG. 4. Two-probe system in which the C region boundaries are expanded by l extra electrode layers.

the evanescent modes with Bloch factors of small magnitude are very rapidly decaying and vanishing in comparison to the propagating modes after only a few layers. In the following, we will exploit this observation and attempt to exclude such evanescent modes from the WFM calculation altogether. Formally this can be accomplished if only the electrode modes $\boldsymbol{\phi}_k$ with Bloch factors λ_k satisfying

$$\lambda_{\min} \leq |\lambda_k| \leq \lambda_{\min}^{-1} \tag{10}$$

are computed and subsequently taken into account, for a reasonable choice of $0<\lambda_{\min}<1$. Equation (10) is adopted as the key relation to select a particular subset of the available electrode modes (as recently suggested in Ref. 17).

B. Extra electrode layers

We will denote the mode, Bloch, and self-energy matrices from which the rapidly decaying evanescent modes are excluded with a tilde, i.e., as $\tilde{\boldsymbol{\Phi}}^{\pm}$, $\tilde{\mathbf{B}}^{\pm}$, and $\tilde{\boldsymbol{\Sigma}}$. The mode matrices holding the excluded modes are denoted by a mathring accent $\mathring{\boldsymbol{\Phi}}^{\pm}$, so that

$$\boldsymbol{\Phi}^{\pm} = [\tilde{\boldsymbol{\Phi}}^{\pm}, \mathring{\boldsymbol{\Phi}}^{\pm}] \tag{11}$$

is the assumed splitting of the full set. All expressions to evaluate the Bloch and self-energy matrices are unchanged as given in Sec. II [now $(\tilde{\boldsymbol{\Phi}}^{\pm})^{-1}$ merely represents the *pseudoinverses* of $\tilde{\boldsymbol{\Phi}}^{\pm}$]. However, since the column spaces of $\tilde{\boldsymbol{\Phi}}^{\pm}$ are not complete, there is no longer any guarantee that WFM can be performed so that the resulting self-energy matrices and, in turn, the solution $\boldsymbol{\psi}_C=[\boldsymbol{\psi}_1^T, \ldots, \boldsymbol{\psi}_n^T]^T$ of the linear system in Eq. (4), are correct. In addition, it is clear that errors can occur in the calculation of $\mathbf{t}$ and $\mathbf{r}$ from Eqs. (8) and (9) because the boundary wave functions $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_n$ might not be fully represented in the reduced sets $\tilde{\boldsymbol{\Phi}}_R^+$ and $\tilde{\boldsymbol{\Phi}}_L^-$.

In order to diminish the errors introduced by excluding evanescent modes, we propose to insert additional electrode layers in the central region (see Fig. 4). As illustrated in Sec. III A, this would quickly reduce the imprint of the rapidly decaying evanescent modes in the boundary layer wave functions $\tilde{\boldsymbol{\psi}}_1$ and $\tilde{\boldsymbol{\psi}}_n$, which means that the critical components outside the column spaces $\tilde{\boldsymbol{\Phi}}^{\pm}$ become negligible at an exponential rate in terms of the number of additional layers. We emphasize that the inserted layers may be "fictitious" in the sense that they can be accommodated by simple block-

Gaussian eliminations prior to the solving of Eq. (4) for the original system.

The above statements are confirmed by the following analysis. We expand the electrode wave functions in the corresponding complete set of bulk modes

$$\boldsymbol{\psi}_i^{\pm} = \boldsymbol{\Phi}^{\pm}\mathbf{a}_i^{\pm} = [\tilde{\boldsymbol{\Phi}}^{\pm}, \mathring{\boldsymbol{\Phi}}^{\pm}]\begin{bmatrix}\tilde{\mathbf{a}}_i^{\pm}\\ \mathring{\mathbf{a}}_i^{\pm}\end{bmatrix}, \tag{12}$$

where $\mathbf{a}_i^{\pm} = [\tilde{\mathbf{a}}_i^{\pm T}, \mathring{\mathbf{a}}_i^{\pm T}]^T$ are vectors that contain the expansion coefficients. In the particular case, where l extra electrode layers are inserted and the border layers of the C region are identical to the connecting electrode layers, the electrode wave functions entering the matching boundary equations will be

$$\boldsymbol{\psi}_1^{(l)-} = (\mathbf{B}_L^-)^{-l}\boldsymbol{\psi}_1^- = [\tilde{\boldsymbol{\Phi}}_L^-, \mathring{\boldsymbol{\Phi}}_L^-]\begin{bmatrix}(\tilde{\boldsymbol{\Lambda}}_L^-)^{-l}\tilde{\mathbf{a}}_1^-\\ (\mathring{\boldsymbol{\Lambda}}_L^-)^{-l}\mathring{\mathbf{a}}_1^-\end{bmatrix} \tag{13}$$

and

$$\boldsymbol{\psi}_n^{(l)+} = (\mathbf{B}_R^{\pm})^l\boldsymbol{\psi}_n^+ = [\tilde{\boldsymbol{\Phi}}_R^+, \mathring{\boldsymbol{\Phi}}_R^+]\begin{bmatrix}(\tilde{\boldsymbol{\Lambda}}_R^+)^l\tilde{\mathbf{a}}_n^+\\ (\mathring{\boldsymbol{\Lambda}}_R^+)^l\mathring{\mathbf{a}}_n^+\end{bmatrix} \tag{14}$$

using the definition $\mathbf{B}^{\pm} = \boldsymbol{\Phi}^{\pm}\boldsymbol{\Lambda}^{\pm}(\boldsymbol{\Phi}^{\pm})^{-1}$. This shows that the critical components outside the column spaces of $\tilde{\boldsymbol{\Phi}}_L^{\pm}$ and $\tilde{\boldsymbol{\Phi}}_R^{\pm}$ are given by coefficients $(\mathring{\boldsymbol{\Lambda}}_L^-)^{-l}\mathring{\mathbf{a}}_1^-$ and $(\mathring{\boldsymbol{\Lambda}}_R^+)^l\mathring{\mathbf{a}}_n^+$, respectively. If this set only consists of the most rapidly decaying of the evanescent modes according to Eq. (10), that is, $|\lambda_k| > \lambda_{\min}^{-1}$ for the diagonal elements of $\mathring{\boldsymbol{\Lambda}}_L^-$ and $|\lambda_k| < \lambda_{\min}$ for the diagonal elements of $\mathring{\boldsymbol{\Lambda}}_R^+$, where $\lambda_{\min}$ is less than 1, these coefficients always decrease as a function of l.

We conclude that WFM with the reduced set of modes approaches the exact case if additional electrode layers are inserted and the solution $\tilde{\boldsymbol{\psi}}_C$ obtained from Eq. (4) approaches the correct solution $\boldsymbol{\psi}_C$ accordingly.

C. Accuracy

As pointed out above, the exclusion of some of the evanescent modes from the mode matrices $\boldsymbol{\Phi}^{\pm}$ will introduce errors because the column spaces in $\tilde{\boldsymbol{\Phi}}^{\pm}$ are incomplete. In this section we will estimate how this will influence the accuracy of the calculated transmission and reflection coefficients in terms of the parameter $\lambda_{\min}$ and the number l of extra electrode layers.

Consider first the accuracy of the transmission matrix $\mathbf{t}$ in the case of the extended two-probe system in Fig. 4. For a specific incoming mode k, we compare the correct result obtained with the complete set of modes [cf. Eq. (8)],

$$\mathbf{t}_k = \begin{bmatrix}\tilde{\mathbf{t}}_k\\ \mathring{\mathbf{t}}_k\end{bmatrix} = [\tilde{\boldsymbol{\Phi}}_R^+, \mathring{\boldsymbol{\Phi}}_R^+]^{-1}\boldsymbol{\psi}_n^{(l)+}, \tag{15}$$

to the result obtained with the reduced mode matrix (denoted by a prime),

$$\mathbf{t}_k' = \begin{bmatrix}\tilde{\mathbf{t}}_k'\\ \mathring{\mathbf{0}}'\end{bmatrix} = [\tilde{\boldsymbol{\Phi}}_R^+, \mathring{\mathbf{0}}]^{-1}\boldsymbol{\psi}_n^{(l)+}, \tag{16}$$

where $\mathring{\mathbf{0}}'$ represents the zero vector of size $\mathring{m}_R$ and $\mathring{\mathbf{0}}$ the zero matrix of size $m_R \times \mathring{m}_R$.

The important coefficients in $\mathbf{t}_k$ and $\mathbf{t}_k'$ for transmission calculations are the ones representing the Bloch modes which enter the Landauer-Büttiker formula in Eq. (2). Since these are never excluded they will always be located within the first $\tilde{m}_R$ elements, i.e., in $\tilde{\mathbf{t}}_k$ and $\tilde{\mathbf{t}}_k'$. It then suffices to compare these parts of the transmission matrix which we can do as follows.

From the properties of the pseudoinverse we are able to write the relation

$$(\tilde{\boldsymbol{\Phi}}_R^+)^{-1}[\tilde{\boldsymbol{\Phi}}_R^+, \mathring{\boldsymbol{\Phi}}_R^+] = [\tilde{\mathbf{I}}, (\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\mathring{\boldsymbol{\Phi}}_R^+], \tag{17}$$

where $\tilde{\mathbf{I}}$ is the identity matrix of order equal to the number of included modes $\tilde{m}_R$. Using the expression in Eq. (14) it then follows that

$$\tilde{\mathbf{t}}_k = (\tilde{\boldsymbol{\Lambda}}_R^+)^l\tilde{\mathbf{a}}_n^+ \tag{18}$$

and

$$\tilde{\mathbf{t}}_k' = \tilde{\mathbf{t}}_k + (\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\mathring{\boldsymbol{\Phi}}_R^+(\mathring{\boldsymbol{\Lambda}}_R^+)^l\mathring{\mathbf{a}}_n^+, \tag{19}$$

where the $\tilde{\mathbf{t}}_k'$ expression clearly corresponds to the correct coefficients $\tilde{\mathbf{t}}_k$ plus an error term.

We have already established in Sec. III B that the $(\mathring{\boldsymbol{\Lambda}}_R^+)^l\mathring{\mathbf{a}}_n^+$ factor in the error term will decrease as a function of l. We now show that the other term, $(\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\mathring{\boldsymbol{\Phi}}_R^+$ is independent of l, and consequently, that the error term in Eq. (19) must decrease as a function of l. To this end we look at the two-norm of $(\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\mathring{\boldsymbol{\Phi}}_R^+$, which satisfies

$$\|(\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\mathring{\boldsymbol{\Phi}}_R^+\|_2 \le \mathring{m}_R^{1/2}\|(\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\|_2, \tag{20}$$

since $\|\mathring{\boldsymbol{\Phi}}_R^+\|_2 \le \mathring{m}_R^{1/2}$ when all evanescent modes are assumed to be normalized. The norm $\|(\tilde{\boldsymbol{\Phi}}_R^+)^{-1}\|_2$ can be readily evaluated and depends on the set of modes included via the parameter $\lambda_{\min}$ but not on l. Thus, we conclude that the only term of Eq. (19) which depend on l is $(\mathring{\boldsymbol{\Lambda}}_R^+)^l\mathring{\mathbf{a}}_n^+$, and the error is therefore decreasing as function of l.

Writing Eq. (19) as $\tilde{\mathbf{t}}_k' = \tilde{\mathbf{t}}_k + \tilde{\boldsymbol{\epsilon}}_k$, where $\tilde{\boldsymbol{\epsilon}}_k$ holds the errors on the coefficients of the kth column, we further obtain that the total transmission T' can be expressed as

$$T' = T + \sum_{kk'}(\tilde{t}_{kk'}^*\tilde{\epsilon}_{kk'} + \tilde{\epsilon}_{kk'}^*\tilde{t}_{kk'} + |\tilde{\epsilon}_{kk'}|^2), \tag{21}$$

where T is the exact result and the summation is over the Bloch modes k and k' in the left and right electrodes, respectively.

For a first-order estimate of the error term in Eq. (21) we consider the worst case approximation, where all diagonal elements of $\mathring{\boldsymbol{\Lambda}}_R^+$ are equal to the maximum range $\lambda_{\min}$ of Eq. (10). This makes all elements $\tilde{\epsilon}_{kk'}$ proportional to $\lambda_{\min}^l$ and we arrive at the simple relation

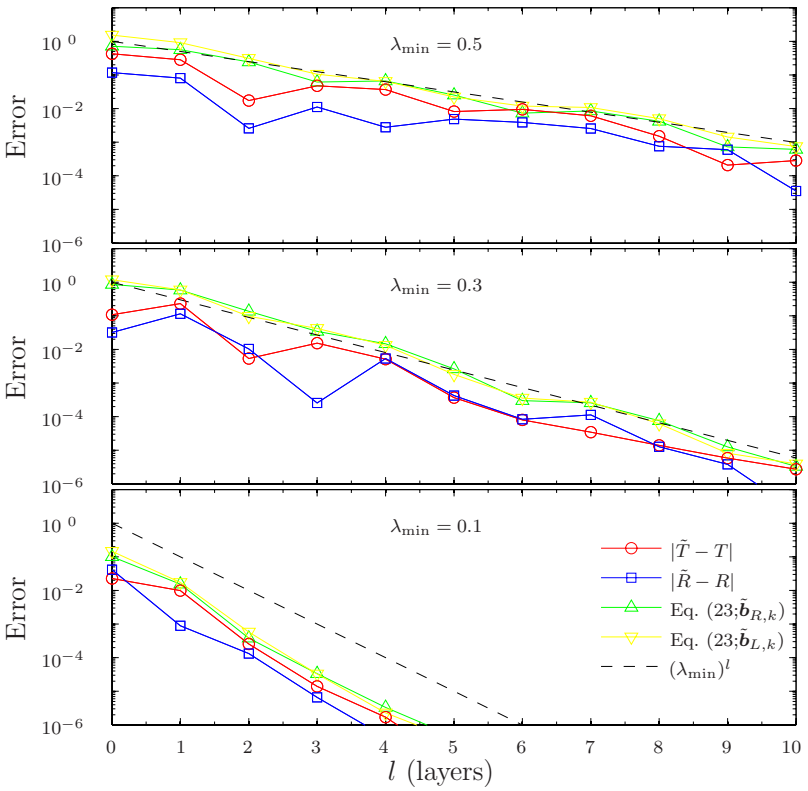

FIG. 5. (Color online) Error (absolute) in the calculated total transmission (circles/solid red lines) and reflection (squares/solid blue lines) coefficients T' and R' as a function of l. The panels show the cases of $\lambda_{\min}$ set to 0.5, 0.3, and 0.1, which corresponds to 3, 14, and 31 Au bulk modes (out of 243, see Fig. 3) taken into account, respectively. Dashed line indicates the first-order error estimate $\lambda_{\min}^{l}$. The upward-pointing and downward-pointing triangles (green and yellow lines) show error estimates obtained from Eq. (23).

$$|T' - T| \sim \lambda_{\min}^{l} + \mathcal{O}[(\lambda_{\min}^{l})^2], \tag{22}$$

which shows that the error decreases exponentially in terms of the number of extra layers l.

For a higher-order estimate of the error, we directly monitor the error arising on the boundary conditions in terms of the coefficient vectors $\tilde{\boldsymbol{b}}_{L,k} \equiv (\tilde{\boldsymbol{\Phi}}_R^+)^{-1}(\boldsymbol{\psi}_1^{(l)+} - \lambda_{L,k}^+ \boldsymbol{\phi}_{L,k}^+)$ and $\tilde{\boldsymbol{b}}_{R,k} \equiv (\tilde{\boldsymbol{\Phi}}_R^-)^{-1}\boldsymbol{\psi}_n^{(l)-}$, where $\boldsymbol{\psi}_1^{(l)+}$ and $\boldsymbol{\psi}_n^{(l)-}$ are given by solving Eq. (4). When the boundary conditions are exactly satisfied, we have $|\tilde{\boldsymbol{b}}_{L,k}| = 0$ and $|\tilde{\boldsymbol{b}}_{R,k}| = 0$. In the case where the boundary conditions are not exactly satisfied, $\tilde{\mathbf{b}}_{R,k}$ represents the error on the left-going components within the right boundary layer in the same way that $\tilde{\boldsymbol{\epsilon}}_k$ represents the error on the right-going (transmitted) components. We would therefore expect the same orders of magnitude of $|\tilde{\mathbf{b}}_{R,k}|$ and $|\tilde{\boldsymbol{\epsilon}}_k|$ in an actual calculation for a given mode k. This suggests the following error estimate from Eq. (21):

$$|T' - T| \leq \sum_k (2|\tilde{\mathbf{t}}_k||\tilde{\boldsymbol{\epsilon}}_k| + |\tilde{\boldsymbol{\epsilon}}_k|^2) \sim \sum_k (2|\tilde{\mathbf{t}}_k||\tilde{\boldsymbol{b}}_{R,k}| + |\tilde{\boldsymbol{b}}_{R,k}|^2), \tag{23}$$

where all the vector norms (e.g., $|\tilde{\mathbf{t}}_k|^2 = \Sigma_{k'}|\tilde{t}_{kk'}|^2$) are assumed to be taken over the elements corresponding to Bloch bulk modes k' only.

Finally, we note without explicit derivation that similar arguments for the reflection matrix with columns $\tilde{\mathbf{r}}'_k = (\tilde{\boldsymbol{\Phi}}_L^-)^{-1}(\boldsymbol{\psi}_1^{(l)-} - \lambda_{L,k}^+ \boldsymbol{\phi}_{L,k}^+)$ and the total reflection coefficient R' result in the same accuracy expressions for $|R' - R|$ if we substitute $\tilde{\mathbf{t}}_k \to \tilde{\mathbf{r}}_k$ and $\tilde{\boldsymbol{b}}_{R,k} \to \tilde{\boldsymbol{b}}_{L,k}$ in Eqs. (22) and (23).

D. Example

To end this section, we exemplify the previous discussion quantitatively by looking at the Au(111) electrode described earlier and assuming a 128-atom (4 unit cells) device of zigzag (8,0) carbon nanotube (CNT) sandwiched between the gold electrodes (see the configuration in Fig. 1). For energy $E = -1.5$ eV, we have calculated the deviation between the total transmission obtained when all bulk modes are taken into account (T) and when some evanescent modes are excluded (T') as specified with different settings of $\lambda_{\min}$. Deviations are also determined for the corresponding total reflection coefficients (R and R'). Figure 5 shows the results as a function of l, together with the estimate $\lambda_{\min}^{l}$ of Eq. (22) and the estimate of Eq. (23) both for the transmission and reflection coefficients, where the higher order terms have been neglected.

We observe that the absolute error in the obtained transmission coefficients (red curves) and reflection coefficients (blue curves) is generally decreasing as a function of l, following the same convergence rate as $\lambda_{\min}^{l}$ (dashed line). Looking closer at results for neighbor l values, we see that the errors initially exhibit wavelike oscillations. This is directly related to the wave form of the evanescent modes that have been excluded [see the propagation of the slowest decaying black curves in Fig. 3(b)]. In other words, although the norm of the errors $|\tilde{\boldsymbol{\epsilon}}_k|$ are decreasing as a function of l, the specific error $\tilde{\epsilon}_{kk'}$ on a given (large) coefficient of $\tilde{t}'_{kk'}$ or $\tilde{r}'_{kk'}$ may increase, which means that the overall error term in Eq. (21) can go up. Fortunately this is only a local phenomenon with the global trend being rapidly decreasing errors.

Consider also the quality of the simple accuracy estimate of $\lambda_{\min}^{l}$ and the estimates expressed by Eq. (23) for the trans-

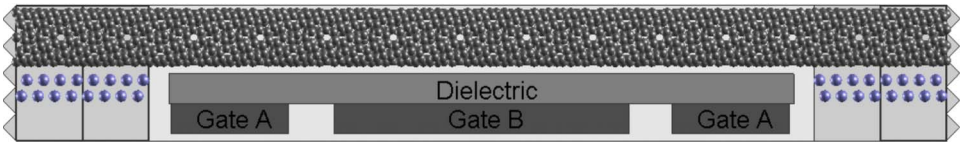

FIG. 6. (Color online) Schematic illustration of a carbon nanotube (8,4) band-to-band tunneling device. The carbon nanotube is positioned on Li surfaces next to an arrangement of three gates.

mission coefficients (green curves) and reflection coefficients (yellow curves), respectively. For relatively large $\lambda_{\min}$ all estimates are very good. However, for smaller values of $\lambda_{\min}$, only the latter two retain a high quality while the $\lambda^{l}_{\min}$ estimate tends to be overly pessimistic. It is important to remember that these estimates are by no means strict conditions but in practice give very reasonable estimates of the accuracy.

We note in passing that the results in the top panel of Fig. 5 correspond to using *only* the propagating Bloch modes in the transmission calculation. Still we are able to compute T and R to an absolute accuracy of three digits by inserting 2×5 extra electrode layers in the two-probe system. This is quite remarkable and shows promise for large-scale systems, e.g., with nanowire electrodes, for which the total number of evanescent modes available becomes exceedingly great.

IV. APPLICATION

In this section we will apply the developed method to a nanodevice consisting of a CNT stretched between to two metal electrodes and controlled by three gates. The setup is inspired by Appenzeller *et al.*[20] and we expect this particular arrangement to be able to display so-called band-to-band (BTB) tunneling, where one observes gate-induced tunneling from the valence band into the conduction band of a semiconducting CNT and vice versa.

We show the configuration of the two-probe system in Fig. 6. The device configuration contains ten principal layers of a CNT(8,4), having 112 atoms in each layer. The diameter of the tube and the thickness of the principal layer are 8.3 Å and 11.3 Å, respectively. The electrodes consist of CNT(8,4) resting on a thin surface of Li, where the lattice constant of the Li layers is stretched to fit the layer thickness of the CNT. The central region of the two-probe system comprises a total of 1440 atoms. An arrangement of rectangular gates is positioned below the carbon nanotube as indicated on the figure. In the plane of the illustration (length $\times$ height) the dimensions are as follows: dielectric 106×5 Å^2; gate A 20×5 Å^2; gate B 50×5 Å^2. We set $\epsilon = 4$ for the dielectric constant of the dielectric in order to simulate SiO_2 or Al_2O_3 oxides. All the regions are centered with respect to the electrodes so that the complete setup has mirror symmetry in the length direction. In the direction perpendicular to the illustration the configuration is assumed repeated every 19.5 Å as a supercell.

We have obtained the density matrix of the BTB device by combining the nonequilibrium Green's function formalism with a semiempirical extended Hückel model (EHT) using the parameterization of Hoffmann.[29] From the density matrix we calculate Mulliken populations on each atom and represent the total density of the system as a superposition of Gaussian distributions on each atom properly weighted by the Mulliken population. The width of the Gaussian is chosen to be consistent with CNDO parameters.[30] The electrostatic interaction between the charge distribution and the dielectrics and gates is subsequently calculated. The Hartree-type term is then included in the Hamiltonian and the combined set of equations are solved self-consistently. The resulting self-consistent EHT model is closely related to the work of Ref. 30, and a detailed description of the model will be presented elsewhere.[31]

In order to adjust the charge transfer between the CNT and the Li electrodes we add the term $\delta\epsilon S$ to the Li parameters. With an appropriate adjusted value of $\delta\epsilon$, the carbon nanotube becomes n-type doped. We adjust the value such that the average charge transfer from Li to the nanotube at self-consistency is $0.002e$ per carbon atom in the electrode. The Fermi energy is then located at -4.29 eV, which is 0.07 eV below the conduction band of the CNT(8,4).

In the following we fix $V_{\text{gate A}} = -2.0$ eV and vary the gate B potentials in the range $[-2 - 4$ eV$]$. Note that we report the gate potentials as an external potential on the electrons, and to translate the values into a gate potential of unit volts the values must be divided with $-e$.

In the left part of Fig. 7 we present the total self-consistent potential induced by the three gates on the carbon atoms in the CNT over the full extension of the device. For each configuration of the gate potentials the electrostatic potential is shown twice, i.e., by two curves with the same color displaced relative to each other with the energies of the valence-band and conduction-band edges, respectively. In this way the curves not only represent the electrostatic potential of the device but also the position of the valence- and conduction-band edges.

Along with this, in the right part of Fig. 7, we show the corresponding transmission spectrum $T(E)$ for four gate potentials $V_{\text{gate B}} = -2.0$, 1.0, 2.0, and 4.0 eV. When $V_{\text{gate B}} = -2.0$ eV the nanotube is largely unperturbed by the gate and the transmission coefficient is close to an ideal (8,4) CNT. We note that this is in agreement with *ab initio* calculations by Nardelli *et al.*,[32] which found that a two terminal (5,5) CNT device in a similar contact geometry showed a nearly ideal conductance spectrum. In addition, the calculated band gap of the (8,4) nanotube is 0.81 eV, which is in good agreement with the value of 0.96 eV obtained from *ab initio* density-functional calculations in the generalized gradient approximation.[33]

From Fig. 7 we see how the bands are shifted upwards by an increasing amount as the gate B potential is turned up. To begin with, e.g., for $V_{\text{gate B}} = 1$ eV, this results in lower conduction since the conduction band bends away from the

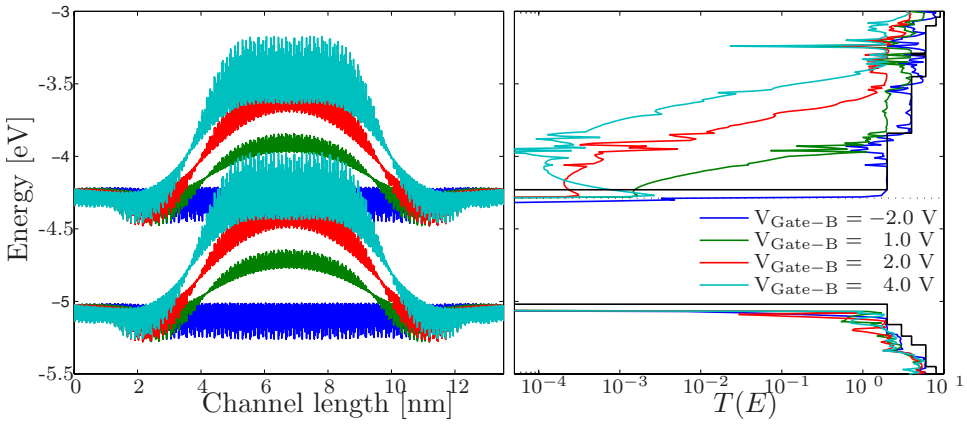

FIG. 7. (Color online) (Left panel) Representation of the electrostatic induced shift of the valence- and conduction-band edges along the length of the device for gate potentials $V_{\text{gate B}}=-2.0$, 1.0, 2.0, and 4.0 eV. (Right panel) The corresponding transmission spectrum. Dotted line shows the position of the Fermi level and the solid line shows the transmission coefficient for an ideal CNT(8,4).

Fermi-level and the Fermi-energy electrons need to tunnel through the central region. When the gate voltage is at $V_{\text{gate B}}=2$ eV, the valence band almost reaches the conduction band in which case BTB tunneling becomes possible. By increasing the gate voltage further, more bands become available for BTB tunneling and the effect is visible as a steady increase in the calculated transmission $T(E)$ just above the Fermi level.

The results for the Fermi-level transmission $T(E_F)$ corresponding to the $T=0$ K unit conduction G_0 are displayed with the black curve in Fig. 8. It shows an initial conductance for $V_{\text{gate B}}=-2.0$ V of the order of one, a subsequent drop by 4 orders of magnitude around $V_{\text{gate B}}=2.0$ V, and a final increase of 1 order of magnitude toward $V_{\text{gate B}}=4.0$ V. We also display the results for the room-temperature $T=300$ K conductance (red curve), which can be obtained from

$$G = \int dE T(E) \frac{e^{(E-E_F)/k_BT}}{(1+e^{(E-E_F)/k_BT})^2}. \tag{24}$$

The two conduction curves are similar, showing that the device is operating in the tunneling regime rather than the thermal emission regime.

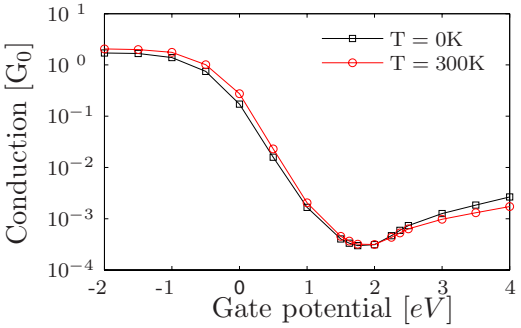

FIG. 8. (Color online) Conduction in units of the conductance quantum G_0 as a function of the gate B potential. In the calculations we use a dielectric constant of 4, $V_{\text{gate A}}=-2.0$ eV, and vary $V_{\text{gate B}}$ from −2.0 to 4.0 eV as indicated.

We next briefly comment on the comparison of the simulation to the experiment of Appenzeller *et al.*[20] In both cases the conduction curves have two branches, which we denote field emission (FE) and BTB. Initially, the conduction decreases with applied gate potential due to the formation of a barrier in the central region: this is the FE regime. For larger biases the conduction increases again due to BTB tunneling, this is the BTB regime. The experimental device displays thermal emission conduction and shows a corresponding subthreshold slope, S, of $k_BT\ln(10)/e \approx 60$ mV/dec in the FE regime. The theoretical device, on the other hand, displays tunneling conduction and has $S \approx 500$ mV/dec in the FE regime. In the BTB regime, the theoretical device has $S \approx 2000$ mV/dec, while the experimental device shows $S \approx 40$ mV/dec.

The very different behavior is due to the short channel length of the theoretical device. The central barrier has a length of ≈5 nm and at this length the electron can still tunnel through the barrier. We see that the short channel length not only affects the subthreshold slope of the FE regime, but also strongly influences the BTB regime. Works are in progress for a parallel implementation of the methodology, which will make it feasible to simulate larger systems and thereby investigate the transition from the tunneling to the thermal emission regime.

All the above results have been calculated with the modified WFM method using parameters $\lambda_{\text{min}}=0.1$ and $l=1$. Thus, the results present a nontrivial application of the method. To verify the transmission results in Fig. 7 we present a comparison to the standard WFM method in Fig. 9. The figure shows that the transmissions curves are identical to about three significant digits. The CPU time required for calculating a complete transmission spectrum for Fig. 7 is (∼3 h), while the corresponding calculation presented in Fig. 9 with the standard WFM method took (∼35 h). Thus, the overall time saving achieved with the method was therefore more than an order of magnitude. The results in Table I indicate that similar time savings can be expected for other systems with nontrivial electrodes.

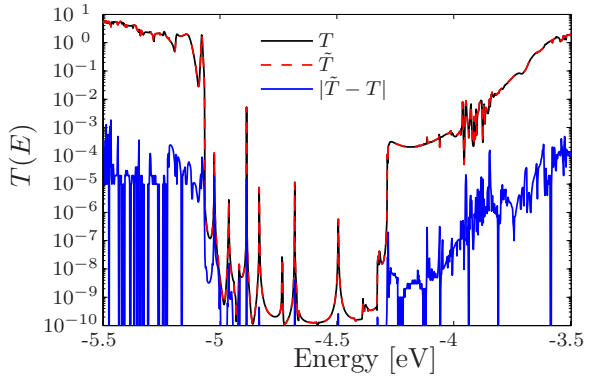

FIG. 9. (Color online) Transmission coefficients T and $\tilde{T}$ calculated with the standard WFM method (black solid) and the method of this work (red dashed), respectively, and the difference $|\tilde{T}-T|$ (blue line) as a function of energy E in the $V_{\text{gate B}}=2$ V case.

V. SUMMARY

We have developed an efficient approach for calculating quantum transport in nanoscale systems based on the WFM scheme originally proposed by Ando.[16] In the standard implementation of the WFM method for two-probe systems, all bulk modes of the electrodes are required in order to represent the transmitted and reflected waves in a complete basis. By extending the central region of the two-probe system with extra electrode principal layers, we are able to exclude the vast majority of the evanescent bulk modes from the calculation altogether. Our final algorithm is therefore highly efficient, and most importantly, errors and accuracy can be closely monitored.

We have applied the developed WFM algorithm to a CNTFET in order to study the mechanisms of band-to-band tunneling. The setup was inspired by Ref. 20, and the calculations display features that are also observed in the experiment. However, due to the short channel length the theoretical device operates in the tunneling regime, while the experimental device operates in the thermal emission regime.

By measuring the CPU times for calculating transmission spectra of the CNTFET two-probe system and comparing to the cost of the standard WFM method we have observed a speed up of more than a factor of 10. We see similar speed up for other nontrivial systems. We therefore believe that this is an ideal method to be used with *ab initio* transport schemes for large-scale simulations.

ACKNOWLEDGMENTS

This work was supported by the Danish Council for Strategic Research (NABIIT) under Grant No. 2106–04–0017, "Parallel Algorithms for Computational Nano-Science."

[1] S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, Cambridge, England, 2005).
[2] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).
[3] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).
[4] Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).
[5] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour, Science **278**, 252 (1997).
[6] S. V. Faleev, F. Léonard, D. A. Stewart, and M. van Schilfgaarde, Phys. Rev. B **71**, 195422 (2005).
[7] P. Pomorski, C. Roland, and H. Guo, Phys. Rev. B **70**, 115408 (2004).
[8] H. S. Gokturk, 5th IEEE Conference on Nanotechnology, Nagoya, Japan, 2005 (IEEE Xplore, 2005), Vol. 2, pp. 677–680.
[9] M. Stilling, K. Stokbro, and K. Flensberg, Mol. Simul. **33**, 557 (2007).
[10] A. Nitzan and M. A. Ratner, Science **300**, 1384 (2003).
[11] M. Di Ventra, S. T. Pantelides, and N. D. Lang, Phys. Rev. Lett. **84**, 979 (2000).
[12] K. Stokbro, J.-L. Mozos, P. Ordejon, M. Brandbyge, and J. Taylor, Comput. Mater. Sci. **27**, 151 (2003).
[13] N. D. Lang and P. Avouris, Phys. Rev. Lett. **84**, 358 (2000).
[14] B. Larade, J. Taylor, H. Mehrez, and H. Guo, Phys. Rev. B **64**, 075420 (2001).
[15] P. A. Khomyakov and G. Brocks, Phys. Rev. B **70**, 195402 (2004).
[16] T. Ando, Phys. Rev. B **44**, 8017 (1991).
[17] P. A. Khomyakov, G. Brocks, V. Karpan, M. Zwierzycki, and P. J. Kelly, Phys. Rev. B **72**, 035450 (2005).
[18] G. Brocks, V. M. Karpan, P. J. Kelly, P. A. Khomyakov, I. Marushchenko, A. Starikov, M. Talanana, I. Turek, K. Xia, P. X. Xu *et al.*, http://www.psi-k.org/newsletters/News_80/newsletter_80.pdf
[19] Hans Henrik B. Sørensen, P. C. Hansen, D. E. Petersen, S. Skelboe, and K. Stokbro, Phys. Rev. B **77**, 155301 (2008).
[20] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, Phys. Rev. Lett. **93**, 196805 (2004).
[21] H. H. B. Sørensen, Ph.D. thesis, Technical University of Denmark, 2008.
[22] Bloch's theorem (Ref. 24) $\boldsymbol{\psi}_i=\lambda_k\boldsymbol{\psi}_{i-1}$ for the ideal electrodes defines the phase factors $\lambda_k\equiv e^{iq_kd}$, where q_k is the complex wave number and d is the layer thickness, which are referred to as Bloch factors throughout this paper.
[23] When using the Landauer formula in Eq. (1) it is assumed that the electrode Bloch modes carry unit current in the conduction direction. This can be conveniently accommodated by flux normalizing the Bloch modes, i.e., $\boldsymbol{\phi}^{\pm}_{L,k}\rightarrow(d_L/v^{\pm}_{L,k})^{1/2}\boldsymbol{\phi}^{\pm}_{L,k}$, in the case of the left electrode (Ref. 34).
[24] N. W. Ashcroft and D. N. Mermin, *Solid State Physics* (Brooks-Cole, Belmont, MA, 1976).
[25] F. Guinea, C. Tejedor, F. Flores, and E. Louis, Phys. Rev. B **28**, 4397 (1983).
[26] M. P. López Sancho, J. M. López Sancho, and J. Rubio, J. Phys. F: Met. Phys. **15**, 851 (1985).
[27] We should point out that the metallic electrodes in the two-probe systems considered in Table I can be fully described by much smaller unit cells than indicated (often only a few atoms are

needed) and therefore the time spent on computing the bulk modes can be vastly reduced in these specific cases. For a general method, however, which supports CNTs, nanowires, etc., as electrodes, the timings are appropriate for showing the overall trend in the computational costs.

[28] P. S. Krstić, X.-G. Zhang, and W. H. Butler, Phys. Rev. B **66**, 205319 (2002).

[29] R. Hoffmann, J. Chem. Phys. **39**, 1397 (1963).

[30] F. Zahid, M. Paulsson, E. Polizzi, A. W. Ghosh, L. Siddiqui, and S. Datta, J. Chem. Phys. **123**, 064707 (2005).

[31] K. Stokbro (unpublished).

[32] M. B. Nardelli, J.-L. Fattebert, and J. Bernholc, Phys. Rev. B **64**, 245423 (2001).

[33] G. L. Zhao, D. Bagayoko, and L. Yang, Phys. Rev. B **69**, 245416 (2004).

[34] D. S. Fisher and P. A. Lee, Phys. Rev. B **23**, 6851 (1981).