Electron transport through molecules: Self-consistent and non-self-consistent approaches

San-Huang Ke,[1,2] Harold U. Baranger,[2] and Weitao Yang[1]

[1]*Department of Chemistry, Duke University, Durham, North Carolina 27708-0354, USA*
[2]*Department of Physics, Duke University, Durham, North Carolina 27708-0305, USA*
(Received 23 November 2003; revised manuscript received 16 March 2004; published 24 August 2004)

A self-consistent method for calculating electron transport through a molecular device is developed. It is based on density functional theory electronic structure calculations under periodic boundary conditions and implemented in the framework of the nonequilibrium Green function approach. To avoid the substantial computational cost in finding the *I-V* characteristic of large systems, we also develop an approximate but much more efficient non-self-consistent method. Here the change in effective potential in the device region caused by a bias is approximated by the main features of the voltage drop. As applications, the *I-V* curves of a carbon chain and an aluminum chain sandwiched between two aluminum electrodes are calculated—two systems in which the voltage drops very differently. By comparing to the self-consistent results, we show that this non-self-consistent approach works well and can give quantitatively good results.

I. INTRODUCTION

In recent years, electron transport through molecules sandwiched between metallic electrodes has been attracting increasing attention both for fundamental reasons and because it may form the basis of a future molecular electronics technology.[1–5] Experimentally, it is difficult to precisely manipulate or even measure the atomic structure of the molecule-electrode contacts. Therefore, neither the influence of atomic structure on transport through the devices nor a path to improved performance is clear. As a result, the ability to calculate the atomic and electronic structure as well as the transport properties of electrode-molecule-electrode systems is important and useful in this field.

Electron transport through nanoscale molecular devices differs significantly from that through macroscopic semiconductor heterostructures. In the latter, the effective-mass approximation is generally successful because of the periodic lattice structure and large electron wavelength. In contrast, in a molecular device a carrier electron will be scattered by only a few atoms whose particular arrangement, then, matters a great deal. Consequently, the effective-mass approximation breaks down, and the electronic structure of the molecular device must be taken into account explicitly. For this purpose, methods based on density functional theory (DFT) are sufficiently accurate and efficient.[6,7] Conventional DFT methods, however, deal with either closed molecular systems (in quantum chemistry) or periodic solids (in solid state physics), neither of which is applicable to molecular transport. Thus one needs to develop a DFT approach suitable for a system which is open, infinite, nonperiodic, and nonequilibrium (if the bias voltage is nonzero).

One way to do this was suggested by Lang *et al.*[8–11] By using the jellium model for the two metallic electrodes of an electrode-molecule-electrode system, they mapped the Kohn-Sham equation of the system into the Lippmann-Schwinger scattering equation and solved for the scattering states self-consistently. They then calculated the current by summing up the contributions from all the scattering states, following a Landauer-type approach.[12] In this way, both the conductance and *I-V* characteristics of the system can be investigated. The use of the jellium model for electrodes is convenient and simple but limited: it cannot include the effects of different contact geometries and surface relaxation, for instance. It also cannot deal with directional bonding such as in semiconductors and transition metals. As a result, the molecule-electrode charge transfer, which is one of the key factors affecting transport, may not be quantitatively correct.[13]

Another way to develop the desired DFT approach is to use the nonequilibrium Green function (NEGF) method.[14,15] The required open and nonequilibrium conditions can be treated rigorously, at least in a formal sense. This method is also closely related to the Landaur approach[12] and has proven to be powerful for studying electron transport through nanoscale devices. Therefore, by combining the NEGF method with conventional DFT-based electronic structure methods used in quantum chemistry or solid-state physics, the coherent transport properties of an electrode-molecule-electrode system can be determined fully self-consistently from first principles. A further advantage of the NEGF+DFT combination is that the atomic structure of the device region and the metallic electrodes are treated explicitly on the same footing. As has been mentioned, the molecule-electrode interaction will induce charge transfer between them and atomic relaxation of their contact—both have a significant effect on electron transport. As a result, the division of the system into the molecule and the electrodes is not meaningful anymore, and some parts of the electrodes must be included into the device region to form an "extended molecule."

Based on this combined NEGF+DFT method, there are several successful implementations[16–20] for molecular conduction and extensive theoretical results in the recent literature.[16–23] According to the way of treating the extended molecule, the semi-infinite leads, and their couplings in a lead-molecule-lead (LML) system, these implementations can be roughly divided into two categories.

In one category people adopted a cluster geometry for all the subsystems of a LML system or for the extended molecule with the leads treated by a tight binding approach (for example, Refs. 16, 23, and 24). It is then convenient to em-

ploy well-established quantum chemistry code (like Gaussian or DMol) to do the electronic structure calculation for the subsystem(s). However, there are potential problems with these treatments for strong molecule-lead couplings: in this case it is obviously necessary to include large parts of the leads into the extended molecule so that the strong molecule-lead interaction can be fully accommodated. To eliminate the artificially introduced surface effects an even larger system is needed, which is usually difficult to deal with by a quantum chemistry code. So in practice only several (or even only one) lead atoms are attached to the molecule to form an extended molecule (for example, Refs. 23 and 24). In this case, significant artificial surface effects are inevitable, the contact atomic relaxation cannot be included, and an accurate molecule-lead coupling is not available. In addition, there may be artificial scattering at the interface between the tight-binding part of the lead and the DFT part of the lead (included in the extended molecule).

In the other category (Refs. 18 and 19), people adopted periodic boundary conditions (PBC) (as in solid state physics) with large parts of the leads included in the extended molecule, so that the interaction between the molecule and its images will be screened off by the metallic lead in between. In this case all the potential problems mentioned above will be absent and the whole LML system becomes nearly perfect in geometry and all the subsystems are treated exactly on the same footing. Two examples of successful implementations adopting the PBC are the TranSiesta package[18] and the MCDCAL package.[19] On the other hand, a drawback is introduced by PBC: when a bias is applied, the Hartree potential must jump unphysically between unit cells. This has previously been addressed by having an independent solution of the Poisson equation.

In this paper, we first develop a fully self-consistent NEGF+DFT method with PBC, which has small but important differences from the two previous implementations. The advantage of our method is that it is simple while still rigorous: the nonequilibrium condition under a bias is fully included in the NEGF part, and, as a result, we do not need to make changes in the conventional electronic structure part. So it is straightforward to combine with any electronic structure method that uses a localized basis set. More importantly, in this way the problem of the unphysical jumps in the Hartree potential is avoided.

A shortcoming of the full self-consistent (SC) NEGF+DFT approach is the large computational effort involved, especially for large systems, large bias voltages, or cases where many bias voltages need to be calculated as for *I-V* characteristics. As a result, a non-self-consistent (non-SC) method with much higher efficiency and useful accuracy is highly desirable. As a step toward this goal, we also construct an approximate but much more efficient non-SC method in which the change in self-consistent effective potential in the device region caused by a bias is approximated by the main features of the voltage drop.

As an application of our approach, in this paper we do calculations by combining it with a very efficient electronic structure package SIESTA.[26] The *I-V* curves of two systems with different typical voltage drop behaviors—a carbon or aluminum chain sandwiched between two aluminum
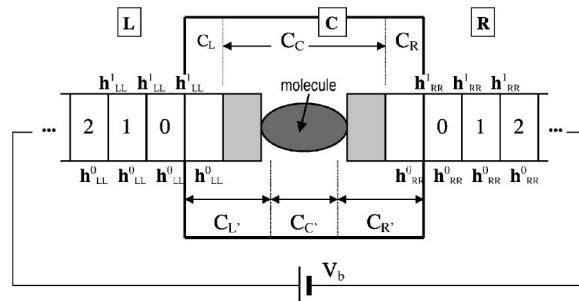

FIG. 1. Schematic drawing of a system containing a molecule sandwiched between two metallic electrodes (leads L and R). The region C is formed by including some parts of L and R so that the C_C part (extended molecule) is charge neutral and the C_L and C_R parts have bulk properties. Because of the use of a localized basis set, the leads L and R can be divided into principal layers (denoted by numbers 0, 1, 2, …). $C_{L'}$, $C_{C'}$, and $C_{R'}$ denote the parts used in the present non-SC approach (see Sec. II E). Their interface is called X in the text. $\mathbf{h^0}$'s and $\mathbf{h^1}$'s are the Hamiltonian matrices within and between the principal layers, respectively.

electrodes—are calculated. Our self-consistent results are in good agreement with those from previous calculations.[18,32] By comparison to the self-consistent results, we examine the validity of the non-SC approach, showing that this approach works quite well and can give quantitatively nearly correct answers.

The arrangement of this paper is as follows. In Sec. II we give briefly a description of our implementation of the NEGF+DFT method. Because the basic formalism of the NEGF+DFT is well established,[16–19] we show only those formula useful for introducing the new features of our method. The present SC and non-SC approaches are explained in Sec. III. Section IV starts with results for a carbon chain and an aluminum chain sandwiched between two Al(001) electrodes. Our results are compared with previous results, and we discuss the validity of the non-SC approach by comparison to the self-consistent results. In Sec. V we summarize and conclude.

II. NEGF+DFT METHOD AND ITS IMPLEMENTATION

A. Modeling of real physical systems

Experimentally, a molecular device system consists of at least a molecular junction coupled with two metallic electrodes (leads L and R) under a bias V_b (two-terminal system). In some cases, there is also a gate terminal applying a gate voltage on the whole system (three-terminal system). Here we consider only the two-terminal system which is schematically shown in Fig. 1. An important consideration for modeling the real physical system is the charge transfer and atomic relaxation around the two molecule-lead contact regions. As a result, we have to include some parts of the metallic leads into the device region, forming an extended molecule. One obvious convergence criteria for the size of the extended molecule is its charge neutrality. Then the charge transfer and the potential disturbance caused by the molecule can be considered screened off outside the extended molecule region. In order to obtain good conver-

gence, we actually include a large part of each metallic lead into the C region, so that the layers adjacent to L and R (i.e., C_L and C_R parts in Fig. 1) have bulk properties. The total Hamiltonian of the system is

$$\hat{H} = \hat{H}_{LL} + \hat{H}_{CC} + \hat{H}_{RR} + \hat{H}_{LC} + \hat{H}_{CR}. \qquad (1)$$

Note that, here the leads L and R interact only through the molecular junction, so their direct interaction term $\hat{H}_{LR}$ vanishes (this can always be satisfied by using a localized basis set).

B. Localized basis set

When $\hat{H}$ is expanded in a basis set, generally only the matrix of $\hat{H}_{CC}$ (denoted $\mathbf{H}_{CC}$) is finite. However, consider a localized (but not necessarily orthogonal) basis set, by which we mean that the overlap between any two basis functions, $\phi_\mu(\mathbf{r}-\mathbf{R}_1)$ and $\phi_\nu(\mathbf{r}-\mathbf{R}_2)$, will be zero if they are separated far enough from each other, $\mathbf{S}_{\mu\nu} \equiv \langle \mu | \nu \rangle = 0$ if $|\mathbf{R}_1 - \mathbf{R}_2| >$ certain cutoff distance. In this case, the region C interacts directly only with finite parts of L and R, and the

nonzero part of the matrices $\mathbf{H}_{LC}$ and $\mathbf{H}_{CR}$ also become finite. Furthermore, we can divide the leads L and R into principal layers so that any principal layer interacts only with its two nearest neighbors (see Fig. 1). As a result, the matrices $\mathbf{H}_{LL}$ and $\mathbf{H}_{RR}$ have the following block tridiagonal form:

$$(\mathbf{H}_{LL})_{ij} = \begin{cases} \mathbf{h}_{LL}^0 & \text{if } i-j=0, \\ \mathbf{h}_{LL}^1 & \text{if } i-j=1, \\ (\mathbf{h}_{LL}^1)^\dagger & \text{if } j-i=1, \\ 0 & \text{if } |i-j|>1, \end{cases} \qquad (2)$$

where $\mathbf{h}_{LL}^0$ and $\mathbf{h}_{LL}^1$ are the Hamiltonian matrices within and between the principal layers, respectively, and i, j are principal layer indexes as shown in Fig. 1. Because the C_L and C_R parts which interact directly with L and R, have bulk properties, the nonzero part of $\mathbf{H}_{LC}$ ($\mathbf{H}_{CR}$) is just $\mathbf{h}_{LL}^1$ ($\mathbf{h}_{RR}^1$), as shown in Fig. 1.

In the localized basis and after the partition shown in Fig. 1, the matrix Green function $\mathbf{G}$ of the whole system, defined by $(E\mathbf{S}-\mathbf{H})\mathbf{G}(E)=\mathbf{I}$, satisfies

$$\begin{bmatrix} E\mathbf{S}_{LL}-\mathbf{H}_{LL} & E\mathbf{S}_{LC}-\mathbf{H}_{LC} & 0 \\ E\mathbf{S}_{LC}^\dagger-\mathbf{H}_{LC}^\dagger & E\mathbf{S}_{CC}-\mathbf{H}_{CC} & E\mathbf{S}_{CR}-\mathbf{H}_{CR} \\ 0 & E\mathbf{S}_{CR}^\dagger-\mathbf{H}_{CR}^\dagger & E\mathbf{S}_{RR}-\mathbf{H}_{RR} \end{bmatrix} \times \begin{bmatrix} \mathbf{G}_{LL} & \mathbf{G}_{LC} & \mathbf{G}_{LR} \\ \mathbf{G}_{CL} & \mathbf{G}_{CC} & \mathbf{G}_{CR} \\ \mathbf{G}_{RL} & \mathbf{G}_{RC} & \mathbf{G}_{RR} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{LL} & 0 & 0 \\ 0 & \mathbf{I}_{CC} & 0 \\ 0 & 0 & \mathbf{I}_{RR} \end{bmatrix}. \qquad (3)$$

The most important part of $\mathbf{G}$ is $\mathbf{G}_{CC}$, corresponding to the region C; from the above equation,

$$\mathbf{G}_{CC}(E) = \{E\mathbf{S}_{CC} - [\mathbf{H}_{CC} + \mathbf{\Sigma}_L(E) + \mathbf{\Sigma}_R(E)]\}^{-1}, \qquad (4)$$

where $\mathbf{\Sigma}_L(E)$ and $\mathbf{\Sigma}_R(E)$ are self-energies which incorporate the effect of the two semi-infinite leads L and R, respectively. $\mathbf{\Sigma}_L(E)$, for example, is defined by

$$\mathbf{\Sigma}_L(E) = (E\mathbf{S}_{LC} - \mathbf{H}_{LC})^\dagger \mathbf{G}_{LL}^0(E)(E\mathbf{S}_{LC} - \mathbf{H}_{LC}), \qquad (5)$$

where $\mathbf{G}_{LL}^0$ is the retarded Green function of the left semi-infinite lead. The latter is given in turn by

$$\mathbf{G}_{LL}^0(E) = (z\mathbf{S}_{LL} - \mathbf{H}_{LL})^{-1},$$

$$z = E + i\eta, \qquad (6)$$

where a typical value for the lifetime broadening η is about 1 meV.

Because of the localized basis set, the nonzero part of $\mathbf{S}_{LC}$, $\mathbf{H}_{LC}$, $\mathbf{S}_{CR}$, and $\mathbf{H}_{CR}$ become finite (being $\mathbf{s}_{LL}^1$, $\mathbf{h}_{LL}^1$, $\mathbf{s}_{RR}^1$, and $\mathbf{h}_{RR}^1$). As a result, only the part of $\mathbf{G}_{LL}^0$ and $\mathbf{G}_{RR}^0$ corresponding to the 0th principal layer of the two leads (denoted $\mathbf{g}_{LL}^0$ and $\mathbf{g}_{RR}^0$) are needed for calculating the nonzero part of the self-energies,

$$\mathbf{\Sigma}_L(E) = (E\mathbf{s}_{LL}^1 - \mathbf{h}_{LL}^1)^\dagger \mathbf{g}_{LL}^0(E)(E\mathbf{s}_{LL}^1 - \mathbf{h}_{LL}^1). \qquad (7)$$

Our notation here follows that for the Hamiltonian: $\mathbf{s}$ and $\mathbf{g}$ are submatrices of the corresponding upper case matrices. $\mathbf{g}_{LL}^0$ and $\mathbf{g}_{RR}^0$ are simply the surface Green functions of the two semi-infinite leads. $\mathbf{g}_{LL}^0$, for example, can be calculated either by simple block recursion,

$$\mathbf{g}_{LL}^0(E) = [z\mathbf{s}_{LL}^0 - \mathbf{h}_{LL}^0 - (z\mathbf{s}_{LL}^1 - \mathbf{h}_{LL}^1)^\dagger \mathbf{g}_{LL}^0(E)(z\mathbf{s}_{LL}^1 - \mathbf{h}_{LL}^1)]^{-1}, \qquad (8)$$

or by a renormalization method[27] in terms of $\mathbf{s}_{LL}^0$, $\mathbf{s}_{LL}^1$, $\mathbf{h}_{LL}^0$, and $\mathbf{h}_{LL}^1$ which can be determined by separate DFT calculations for the two leads. For small lifetime broadening η (1 meV), we find that the renormalization method is much faster than simple block recursion. This is natural since n renormalization interations incorporate 2^n principal layers, while n recursions incorporate only n.

From $\mathbf{G}_{CC}(E)$, the projected density of states (PDOS) on the molecule (indicated by m) is given by

$$N_m(E) = -\frac{1}{\pi} \text{Im}\{\text{Tr}_m[\mathbf{G}_{CC}(E+i\eta) \cdot \mathbf{S}_{CC}]\}, \qquad (9)$$

where Tr_m means the trace is performed only on the molecular part of the matrix.
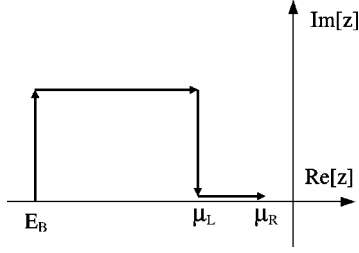
FIG. 2. Schematic drawing of the integration path in the complex energy plane used to calculate the density matrix [Eq. (14)]. E_B is the lowest energy of occupied states, and $\mu_{L,R}$ are the chemical potentials of the left and right leads, respectively ($\mu_L < \mu_R$ is assumed). Note that for energy window $[E_B, \mu_L]$ a complex contour integration is performed while for energy window $[\mu_L, \mu_R]$ a direct energy integration is performed by using a fine energy mesh and a very small imaginary part.

C. Current

The nonequilibrium Green function technique (NEGF)[15,28,29] provides a convenient way to calculate the current by post-processing a DFT calculation. The result is quite natural and intuitive: First, the basic assumption is that there is no energy relaxation within the molecular region. Then, following a Landauer-type point of view,[12,15] one divides the electrons in the molecule into two sets using scattering-wave states, those that came from the left lead and those that came from the right. The left-lead states are, of course, filled up to the chemical potential in the left lead, μ_L, while the right-lead states are filled up to μ_R. In equilibrium, the two chemical potentials are equal, and the current carried by the left-lead states is, of course, equal to that carried by the right-lead states. As a bias is applied, the balance between the two types of states is disrupted and current flows. As different states are populated because of the change in chemical potentials, the charge density in the molecule also changes. The potential profile must be solved self-consistently in order to get an accurate measure of the transmission. It is this self-consistency which is the time-consuming part of the calculation.

The expression for the steady-state current through the C region for applied bias V_b is

$$I(V_b) = -\frac{2e^2}{h} \int_{-\infty}^{+\infty} T(E, V_b)[f(E - \mu_L) - f(E - \mu_R)]dE, \tag{10}$$

where μ_L and μ_R are the chemical potentials, f is the Fermi function, and $T(E, V_b)$ is the transmission probability for electrons from the left lead to right lead with energy E under bias V_b. The transmission probability is related to Green functions by

$$T(E, V_b) = \text{Tr}[\mathbf{\Gamma}_L(E)\mathbf{G}_{CC}(E)\mathbf{\Gamma}_R(E)\mathbf{G}_{CC}^\dagger(E)], \tag{11}$$

where

$$\mathbf{\Gamma}_{L,R}(E) = i(\mathbf{\Sigma}_{L,R}(E) - [\mathbf{\Sigma}_{L,R}(E)]^\dagger) \tag{12}$$

reflect the coupling at energy E between the C region and the leads L and R, respectively.

The charge density corresponding to the above picture of left-lead states filled to μ_L and right-lead states filled μ_R can also be expressed in terms of Green functions. In particular, the density matrix of region C in the basis-function space is

$$\mathbf{D}_{CC} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dE[\mathbf{G}_{CC}(E)\mathbf{\Gamma}_L(E)\mathbf{G}_{CC}^\dagger(E)f(E - \mu_L) + \mathbf{G}_{CC}(E)\mathbf{\Gamma}_R(E)\mathbf{G}_{CC}^\dagger(E)f(E - \mu_R)] \tag{13}$$

$$= -\frac{1}{\pi} \int_{-\infty}^{+\infty} dE\, \text{Im}[\mathbf{G}_{CC}(E)f(E - \mu_L)] + \frac{1}{2\pi} \int_{-\infty}^{+\infty} dE[\mathbf{G}_{CC}(E)\mathbf{\Gamma}_R(E)\mathbf{G}_{CC}^\dagger(E)][f(E - \mu_R) - f(E - \mu_L)]. \tag{14}$$

Time-reversal symmetry ($\mathbf{G}_{CC}^\dagger = \mathbf{G}_{CC}^*$) was invoked in going from (13) to (14). The integrand of the first term of (14) is analytic [all poles of $\mathbf{G}_{CC}(E)$ are on real axis], so the integral can be evaluated easily by complex contour integration. However, the integrand of the second term is not analytic, so it must be evaluated by integrating very close to the real axis using a very fine energy mesh. The whole integration path[17] is shown in Fig. 2. Because we construct the region C such that C_L and C_R have essentially bulk properties, we can use the bulk density matrix for them.

The calculated density matrix is then output to the DFT part to calculate the electron density ρ and to construct a new $\mathbf{H}_{CC}$,

$$\rho(\mathbf{r}) = \sum_{\mu,\nu} \phi_\mu^*(\mathbf{r})\text{Re}[(\mathbf{D}_{CC})_{\mu\nu}]\phi_\nu(\mathbf{r}), \tag{15}$$

$$(\mathbf{H}_{CC})_{\mu\nu} = \langle\mu|\hat{T} + \hat{V}_{\text{ext}}(\mathbf{r}) + \hat{V}_H[\rho(\mathbf{r})] + \hat{V}_{\text{xc}}[\rho(\mathbf{r})]|\nu\rangle, \tag{16}$$

where $\hat{T}$ is the kinetic energy, and $\hat{V}_{\text{ext}}$, $\hat{V}_H$, and $\hat{V}_{\text{xc}}$ are the external, Hartree, and exchange-correlation potential energies, respectively. The new $\mathbf{H}_{CC}$ replaces the old, a new $\mathbf{D}_{CC}$ is calculated, and so on until $\mathbf{H}_{CC}$ or $\mathbf{D}_{CC}$ converges. Finally, the transmission function $T(E)$ can be calculated by Eq. (11).

One subtlety here is the different boundary conditions used in the Green function and DFT parts—open versus periodic, respectively. This means that some iteration must be done even at $V_b=0$. If the supercell of the DFT part has the same size as region C in Fig. 1, then the C_L and C_R parts will interact directly due to the periodic boundary condition. However, this interaction is absent in the calculation of the density matrix. The same problem exists also for $\mathbf{H}_{CC}$. So we need to do some translation work between the Green function and DFT parts: to add this interaction when we go from NEGF to DFT by using the density matrix elements between two adjacent principal layers, and to remove it when we go from DFT to NEGF by setting corresponding parts of $\mathbf{S}_{CC}$ and $\mathbf{H}_{CC}$ to zero. Generally, the supercell of the DFT part can be made larger than the size of the region C, especially for systems without a translational symmetry, because the DFT part is usually much cheaper than the Green function part.

III. NEW SELF-CONSISTENT AND NON-SELF-CONSISTENT APPROACHES

A. Bias voltage

For nonzero V_b, care must be taken to account for the effects of the bias voltage on the charge density. One way to proceed is to apply a constant field in the direction parallel to the leads within the supercell of the DFT calculation. Thus a linear drop is added to the external potential in Eq. (16), and the effect on the charge density follows from, for instance, solving the Poisson equation. This approach is not straightforward for periodic boundary conditions because of the artificial potential jumps at the two supercell boundaries in the lead direction. One way to eliminate the unphysical jumps is to use a larger supercell for the DFT calculation and replace the Hamiltonian of each of the regions near the potential jumps by the bulk one with a constant potential shift given by the bias voltage; this is implemented in the Transiesta program.[18]

Here we propose a different approach to handle the bias, one which is less obvious but turns out to be simpler in the end: The bias is included through the density matrix ($\mathbf{D}_{CC}$) in the Green function calculation instead of the potential ($\mathbf{H}_{CC}$) in the DFT part. Specifically, we calculate the density matrix by Eq. (13) under the boundary condition that there is a potential difference V_b between part C_L (together with the left lead) and part C_R (together with the right lead). This is done by shifting all the potentials related to the left (right) lead and the C_L (C_R) part by $-V_b/2$ ($+V_b/2$). Shifting the potential in a lead is equivalent to directly shifting the energy by the opposite amount, so

$$\mathbf{D}_{CC} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dE \left[\mathbf{G}_{CC}(E)\mathbf{\Gamma}_L\left(E + \frac{eV_b}{2}\right)\mathbf{G}_{CC}^\dagger(E)f(E-\mu_L) + \mathbf{G}_{CC}(E)\mathbf{\Gamma}_R\left(E - \frac{eV_b}{2}\right)\mathbf{G}_{CC}^\dagger(E)f(E-\mu_R) \right]. \tag{17}$$

Here the Green function $\mathbf{G}_{CC}(E)$ has all the potential shifts included,

$$\mathbf{G}_{CC}(E) = \left(E\mathbf{S}_{CC} - \left[\mathbf{H}_{CC} + \Delta\mathbf{H}_{C_L} + \Delta\mathbf{H}_{C_R} + \mathbf{\Sigma}_L\left(E + \frac{eV_b}{2}\right) + \mathbf{\Sigma}_R\left(E - \frac{eV_b}{2}\right) \right] \right)^{-1}, \tag{18}$$

where the C_L and C_R parts of $\mathbf{H}_{CC}$ are replaced by $\mathbf{h}_{LL}^0$ and $\mathbf{h}_{RR}^0$, and their potential shifts are

$$[\Delta\mathbf{H}_{C_L}]_{\mu\nu} = -\frac{eV_b}{2}[\mathbf{S}_{CC}]_{\mu\nu}, \ \forall \ \mu,\nu \in C_L, \tag{19}$$

$$[\Delta\mathbf{H}_{C_R}]_{\mu\nu} = +\frac{eV_b}{2}[\mathbf{S}_{CC}]_{\mu\nu}, \ \forall \ \mu,\nu \in C_R. \tag{20}$$

The actual computational process is exactly the same as that of (14) and Fig. 2, and self-consistency is achieved in the same way as for zero bias. Finally, in the calculation of the transmission, the potential shift present in the self-energies appears in $\Gamma_{L,R}$ as in (17),

$$T(E,V_b) = \mathrm{Tr}\left\{ \mathbf{\Gamma}_L\left(E + \frac{eV_b}{2}\right)\mathbf{G}_{CC}(E)\mathbf{\Gamma}_R\left(E - \frac{eV_b}{2}\right)\mathbf{G}_{CC}^\dagger(E) \right\}. \tag{21}$$

While the two approaches mentioned here—the linear external potential and the potential shift in the leads—give quite different results in the unphysical noninteracting limit, self-consistency ensures that they give the same result in physical cases. Our approach has the distinct technical advantage that the Green function and DFT calculations are completely disconnected, allowing the transport module to be easily combined with a wide variety of electronic structure codes.
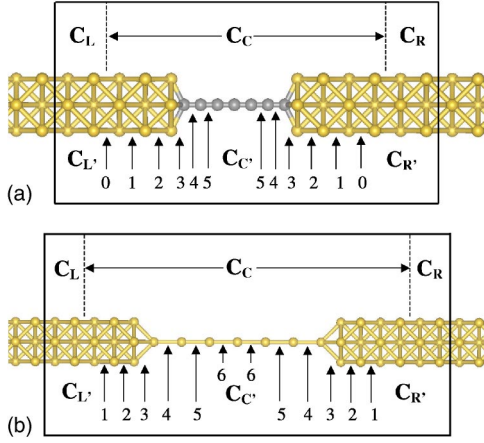
FIG. 3. (Color online) Systems calculated are (a) a chain of seven carbon atoms sandwiched between two Al leads in the (001) direction of bulk Al, (b) a chain of 7 Al atoms sandwiched between the same leads. The C-Al distance in (a) is 1 Å and the C-C distance is 1.323 Å. The Al-Al (chain) distance in (b) is 2.86 Å and the Al-Al (surface) distance is 2.025 Å (i.e., the two Al atoms at the ends of the chain are in their bulk positions). The notations for different parts are the same as in Fig. 1. The numbers 0, 1, 2, 3, 4, 5, 6 denote different interfaces, called the interface X, between $C_{L'}$ or $C_{R'}$ and $C_{C'}$ which are used in the present non-self-consistent approach.

B. Approximate non-self-consistent approaches

For large systems under large bias, the full SC approach described above is computationally very difficult. The longest part of a one-time calculation is finding the surface Green functions (for all the points in the energy mesh), even with the fast renormalization method. However, one only needs to do these calculations once and save the results. The major cost for a full self-consistent calculation is from finding $\mathbf{G}_{CC}(E)$ by Eq. (4) at the many energies needed for the density matrix, especially for the very fine mesh in the energy window $[\mu_L, \mu_R]$ (see Fig. 2).

To avoid the complex procedure and large computational effort of full self-consistency, here we propose an approximate non-SC approach in which the bias is included by (a) applying a potential shift $-eV_b/2$ $(+eV_b/2)$ to the left (right) lead through the self-energies as in Eq. (21), and (b) introducing potential shifts non-self-consistently into the region C. The main consideration is that it may be a good approximation to replace the change in self-consistent effective potential caused by a bias by the main features of the voltage drop. For instance, for conductive molecular devices, the bias voltage will drop mainly at the left and right contact regions if the contacts have low transparency or over the whole molecular region if the contacts are very transparent.

In our approach, we introduce new parts on the left $(C_{L'})$ and right $(C_{R'})$ within the region C which extend from their respective leads to the molecular contacts, as shown in Fig. 1. We will denote by X the interface between the molecule $C_{C'}$ and the regions $C_{L'}$ plus $C_{R'}$. If the voltage drop around the left (right) contact is $\alpha V_b [(1-\alpha)V_b]$, then the potential shifts are applied in the following way (which we call the $\Delta \mathbf{H}1$ treatment):

$$\mathbf{H}'_{CC} = \mathbf{H}_{CC} + \left(\alpha - \frac{1}{2}\right) eV_b \mathbf{S}_{CC}, \qquad (22)$$

$$[\Delta \mathbf{H}_{C'_L}]_{\mu\nu} = -e\alpha V_b [\mathbf{S}_{CC}]_{\mu\nu}, \mu \text{ or } \nu \in C_{L'}, \qquad (23)$$

$$[\Delta \mathbf{H}_{C'_R}]_{\mu\nu} = e(1-\alpha)V_b [\mathbf{S}_{CC}]_{\mu\nu}, \mu \text{ or } \nu \in C_{R'}. \qquad (24)$$

Because the potential shift is applied to a matrix element when *either* orbital index is in the $C_{L',R'}$ part, the voltage drop will be slightly smeared around the interface X. To explicitly show the role of $\mathbf{S}_{CC}$ in (23) and (24), we also drop the potential using (called the $\Delta \mathbf{H}0$ treatment)
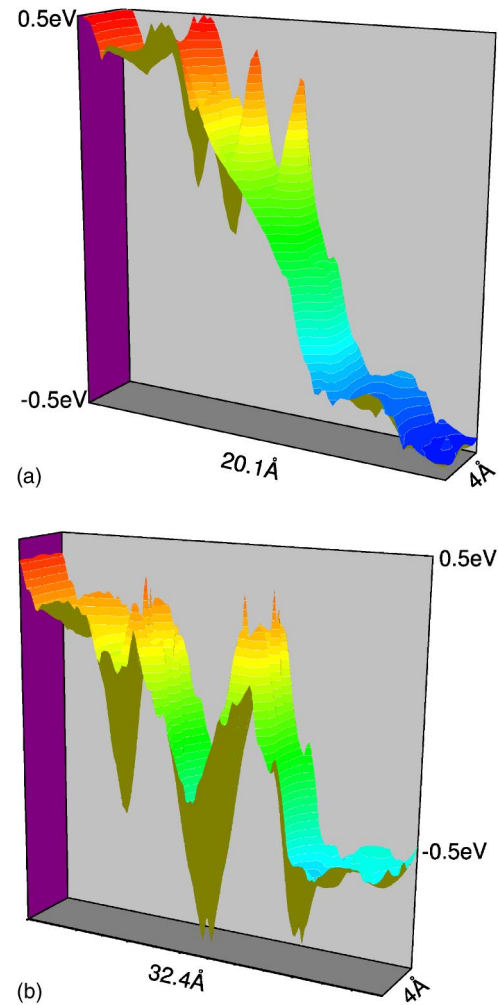

FIG. 4. (Color online) Voltage drop in a plane going through the atomic chain for the C_C region of (a) the C chain and (b) the Al chain, as shown in Fig. 3, for an applied bias of -1 V. Note that (1) in (a) the voltage drop mainly occurs around the right contact region while in (b) it occurs over the entire chain, and (2) the oscillation in (b) is much larger than that in (a). This is because the electrons in the Al chain are more free than that in the C chain; as a result, the polarization induced by the bias is larger in the Al chain.

$$[\Delta\mathbf{H}_{C'_L}]_{\mu\nu} = -e\alpha V_b[\mathbf{S}_{CC}]_{\mu\nu}, \mu, \nu \in C_{L'}, \qquad (25)$$

$$[\Delta\mathbf{H}_{C'_R}]_{\mu\nu} = e(1-\alpha)V_b[\mathbf{S}_{CC}]_{\mu\nu}, \mu, \nu \in C_{R'}. \qquad (26)$$

In either case, $\mathbf{G}_{CC}(E)$ is determined by an equation analogous to Eq. (18) in the SC case,

$$\mathbf{G}_{CC}(E) = \left(E\mathbf{S}_{CC} - \left[\mathbf{H}'_{CC} + \Delta\mathbf{H}_{C'_L} + \Delta\mathbf{H}_{C'_R} + \mathbf{\Sigma}_L\left(E + \frac{eV_b}{2} \right) \right. \right.$$
$$\left. \left. + \mathbf{\Sigma}_R\left(E - \frac{eV_b}{2} \right) \right] \right)^{-1}. \qquad (27)$$

The initial $\mathbf{H}_{CC}$ matrix here comes from a separate DFT calculation using a large L-C-R supercell. Finally, the current is calculated as usual through Eqs. (21) and (10).

So far we have considered the simplest case of low transparency contacts so that the voltage drops in only two places near the contacts. In this simplest case the α parameter here has the same role as the η parameter in Ref. 25. For a real device, the voltage drop may be much more complicated, and there may be several different voltage drops inside the device region. However, in our method all the factors affecting the voltage drop have been taken into account at the DFT level, and it is straightforward to generalize the present non-SC approach for these more complicated situations (see calculations for system B in Sec. IV) as long as the main features of the voltage drop are known. Actually, we will show later that results of this non-SC approach are not very sensitive to the choice of the voltage drop. If we assume that the form of the drop is not affected significantly by a change in the bias voltage itself, then the main features of the voltage drop in a system can be determined by a single self-consistent calculation using a relatively small bias voltage.

IV. APPLICATIONS: CHAINS OF CARBON OR ALUMINUM

We report calculations of I-V curves for two systems: a *carbon chain (system A)* or an *aluminum chain (system B)* sandwiched between two aluminum leads in the (001) direction of bulk Al. The structures are shown in Fig. 3. No further atomic relaxation is performed for simplicity and for direct comparison with previous results.

Our implementation of the transmission calculation is independent of the DFT part. Therefore, it can be easily combined with any DFT package that uses a localized basis set. As an application, here we combine it with the very efficient full DFT package SIESTA,[26] which adopts a LCAO-like and finite-range numerical basis set and makes use of pseudopotentials for atomic cores. In our calculations we adopt a single zeta (SZ) basis set. To check the convergence of the results, we also calculate the equilibrium transmission function using a single zeta plus polarization (SZP) basis set. The difference is found to be minor. The PBE version of the generalized gradient approximation[30] is adopted for the elec-

tron exchange and correlation, and optimized Troullier-Martins pseudopotentials[31] are used for the atomic cores. The initial density matrix of the region C is obtained from a separate DFT calculation using a large L-C-R supercell.

There are two main reasons for us to choose to study these systems. First, the transmission function $T(E)$ of system A has been calculated by both the TranSiesta[18] and MCDCAL[32] packages using a SZ basis set. So we can make a direct comparison to previous results. Second, systems A and B typify two different voltage drop behaviors (although both the C and Al chains are conductive). In system A, because the molecule-lead contact is a hetero-interface, the voltage drop will mainly occur around the two interfaces [see Fig. 4(a), note that the voltage drops around the two interfaces are actually asymmetric]. In contrast, in system B the molecule-lead contact is a homointerface, and furthermore the two Al atoms at the ends of the chain are at their bulk positions. So the voltage drop will occur over the entire Al chain in some way [see Fig. 4(b)]. Our purpose is to see whether our non-SC approach can handle these different conditions.

A. Transmission functions

In Fig. 5 we show the calculated transmission functions and PDOS (projected on the chains) for system A and B under zero bias voltage. As it can be seen, the transmission function generally follows the PDOS except for some localized states [for instance, around 4–5 eV in (a) and 1 eV in (b)] which are not coupled with the left or the right lead. Our result of $T(E)$ for system A is in very good agreement with the previous results from TranSiesta and MCDCAL packages[18,32] [see Fig. 6(a) of Ref. 18].

In Fig. 6 we show, for system A under a bias of 1.0 V (α=0.5), the calculated $T(E)$ by the SC approach and the non-SC approach ($\Delta\mathbf{H}1$ treatment) with different choices for the interface X between the $C_{L'}$ or $C_{R'}$ and $C_{C'}$ parts. The first thing we note is that our SC result for 1 V bias is in very good agreement with the previous self-consistent results[18,32] (see Fig. 6 of Ref. 18). When we change the interface X (denoted by the different numbers in Fig. 6) from deep in the leads to the contact regions (i.e., interface $X=1\rightarrow2\rightarrow3$), the non-SC result varies significantly. However, as X moves into the contact regions (i.e., interface $X=3$, or 4, or 5), the result becomes very close to the SC result and insensitive to the exact position of X. This result is just what we expect because in system A the bias voltage drops mainly around the hetero-interface contact regions. This can be regarded as an advantage of the present non-SC approach: its result is not strongly dependent on the technical choice.

Similar calculations of $T(E)$ for system B under V_b =1.0 V (α=0.5, $\Delta\mathbf{H}1$ treatment for non-SC) are shown in Fig. 7. Again, after the interface X is moved into the contact regions the different non-SC result become quite close. However, compared to the case of system A, the agreement with the SC result is not good, especially in the energy range around the (averaged) Fermi level: the transmission from the non-SC calculations (except for interface $X=5$) is noticeably larger than that from the SC calculation. This substantial dis-
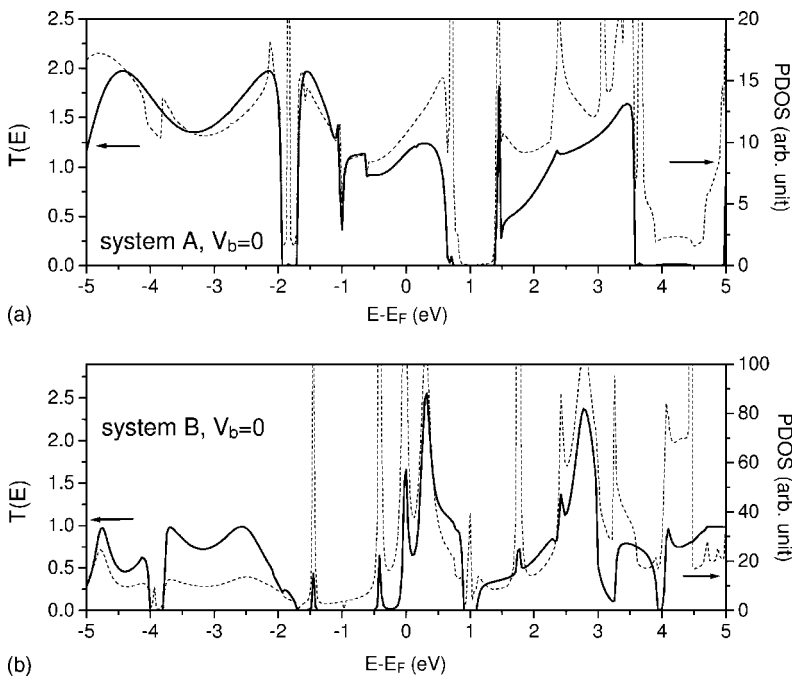
FIG. 5. Calculated transmission function $T(E)$ (solid line) and projected density of states (PDOS, projected on the chains, dashed line) by the self-consistent approach for (a) the carbon chain (system A) and (b) the aluminum chain (system B) under zero bias voltage. When $T(E)$ is significantly different from the PDOS, it means that a localized state exists at that energy.
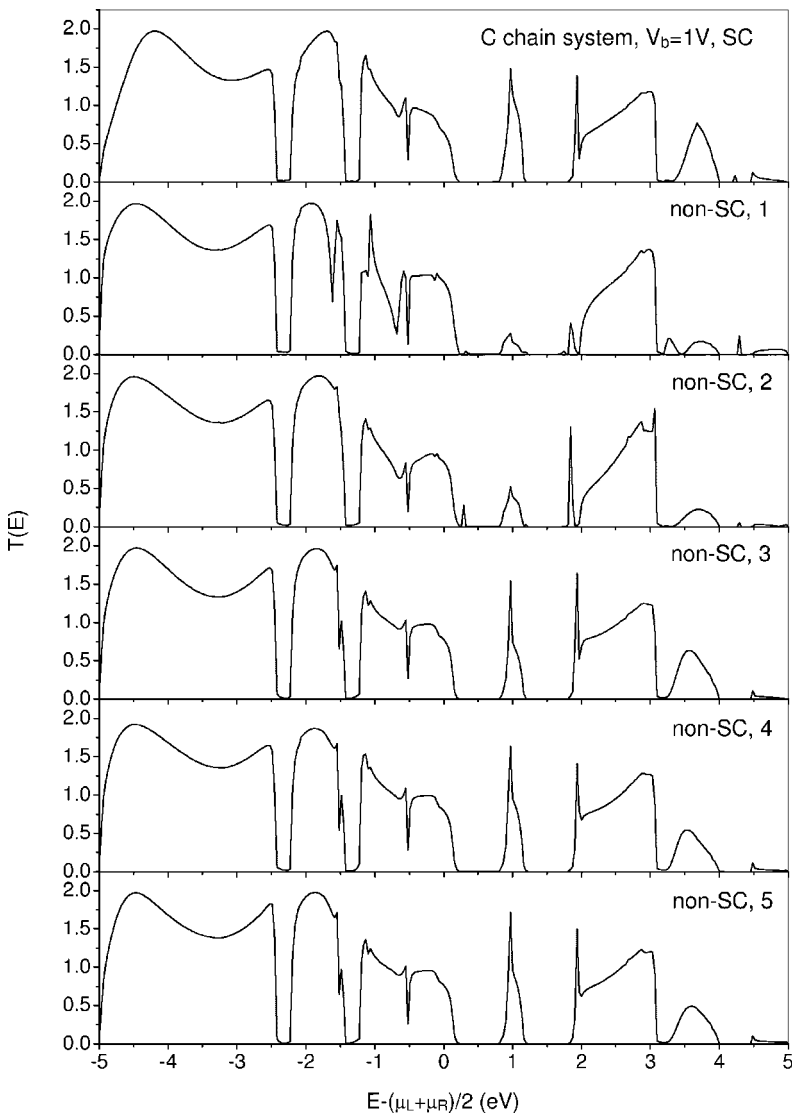

FIG. 6. Transmission function $T(E)$ for the carbon chain (system A) under 1 V bias (α $=0.5$) from the SC approach and the non-SC approach ($\Delta \mathbf{H}$1 treatment) with different choices for the interface X between $C_{C'}$ and the $C_{L'}$ or $C_{R'}$ parts (denoted by 1,...,5) as shown in Fig. 3(a). Note the similarity between the different non-SC results and the fully SC result once X is placed in the contact regions ($\geqslant 3$).
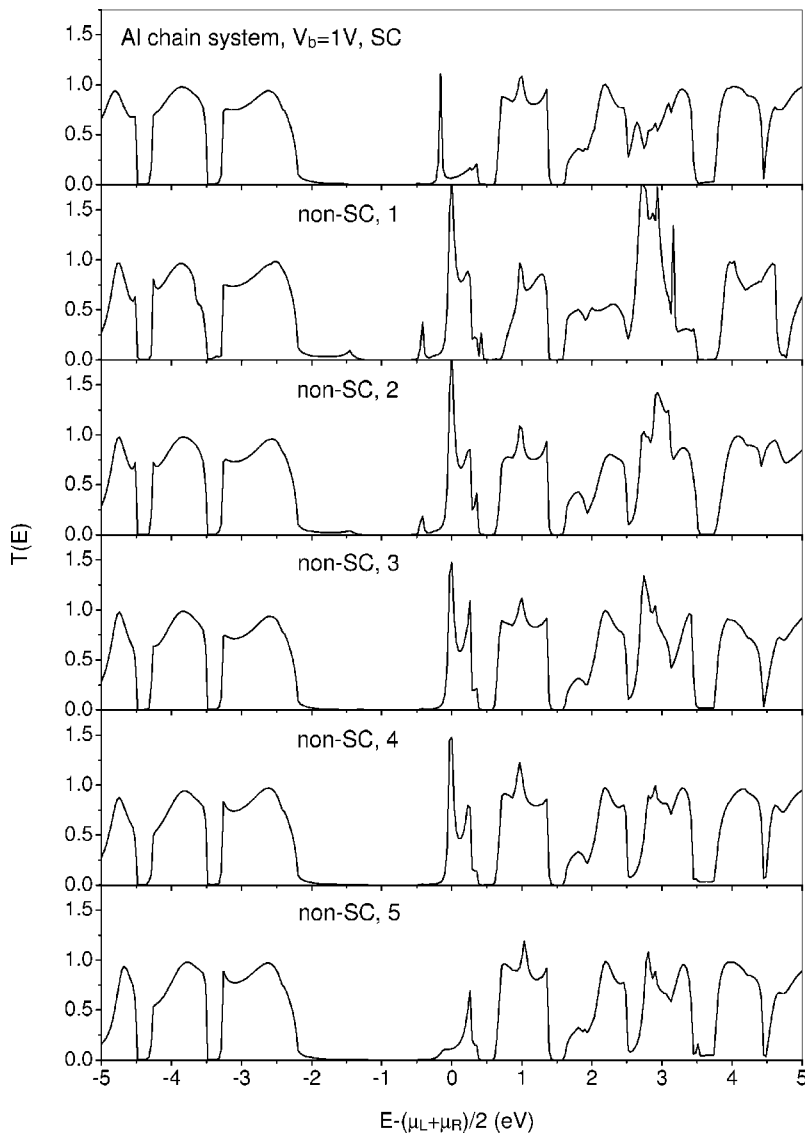
FIG. 7. Calculated transmission function $T(E)$ for the Al chain system under 1 V bias ($\alpha=0.5$, $\Delta\mathbf{H}1$ treatment for non-SC) by the SC and non-SC approaches. The notations are similar to those in Fig. 6.

agreement originates in the difference between the self-consistent effective potential [Fig. 4(b)] and the non-self-consistent one assumed in the non-SC approach.

B. *I-V* curve of system A

In order to show effects of different voltage drops and the difference between the $\Delta\mathbf{H}0$ and $\Delta\mathbf{H}1$ treatments [see Eqs. (23)–(26)] for potential shifting, we give in Fig. 8 the calculated *I-V* curves for system A from the non-SC approach with interface $X=3$ compared to the SC result. We do the non-SC calculations for three different voltage drops ($\alpha=0.2$, 0.5, 0.8) for the $\Delta\mathbf{H}1$ treatment. In addition, for $\alpha=0.8$ we do the non-SC calculation with the $\Delta\mathbf{H}0$ treatment. Among the three different voltage drops, the result for $\alpha=0.2$ is in poor agreement with the SC result while those for $\alpha=0.5$ and 0.8 are in good agreement. This makes sense in view of the main features of the voltage drop in Fig. 4(a): the bias voltage will mainly drop around the left (right) contact for a positive (negative) bias. However, the small difference in *I-V* curve between $\alpha=0.5$ and 0.8 indicates that the

I-V characteristics is actually not sensitive to the exact change in voltage drop. By comparing the SC result and the two non-SC results for $\alpha=0.8$, we see that the $\Delta\mathbf{H}1$ treatment improves the result markedly. This is understandable because voltage drops in real physical systems are not sharp step functions but are somewhat smeared out.

C. *I-V* curves of systems A and B: comparison of different approaches

I-V curves for systems A and B calculated by different approaches are given in Figs. 9 and 10, respectively. For the non-SC approach, $\alpha=0.5$ and the $\Delta\mathbf{H}1$ treatment are always adopted. It turns out that for small bias voltages ($V_b<0.3$ V) all the different treatments (the SC, the non-SC with different interface X), all give similar results, i.e., the effect of the different voltage drops is very small. Along with the increase of bias voltage, the difference among the different calculations becomes more and more significant.

For the carbon chain (system A), because the molecule-electrode contact is a hetero-interface and therefore the volt-
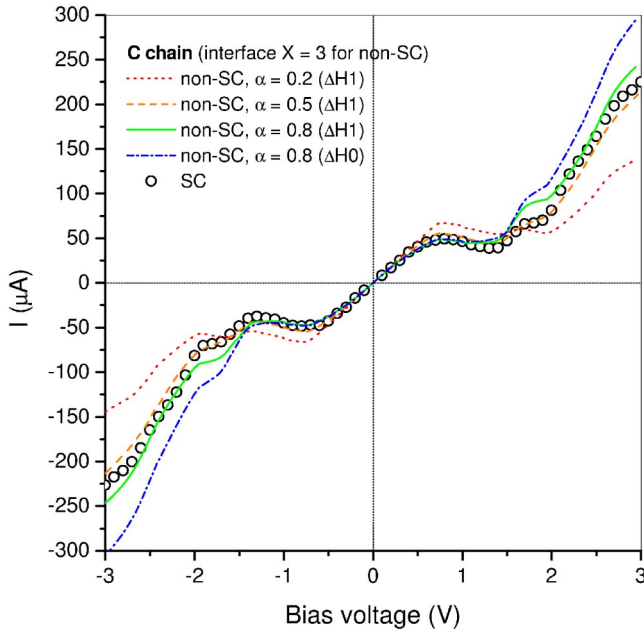
FIG. 8. (Color online) Calculated *I-V* curves for the C chain system by the SC approach and the non-SC approach with interface *X*=3 [as shown in Fig. 3(a)]. Three different voltage drops are considered α=0.2, 0.5, 0.8. The $\Delta\mathbf{H}0$ and $\Delta\mathbf{H}1$ treatments are adopted for α=0.8. The simple non-SC approach does very well using α=0.5 or 0.8 and the $\Delta\mathbf{H}1$ treatment.

age drop occurs mainly around the contact regions, the SC result and the non-SC results with the interface *X* located around the two contact regions are in good agreement. As for the transmission function in Fig. 6, once the interface *X* is within the contact regions the result is quite insensitive to the
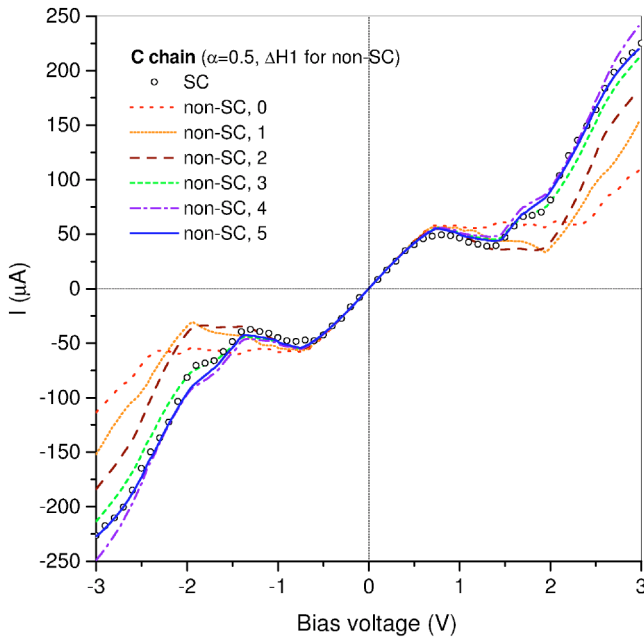

FIG. 9. (Color online) Calculated *I-V* curves for the C chain system by the SC approach and the non-SC approach with different interface *X* indicated by the numbers. α=0.5 and $\Delta\mathbf{H}1$ treatment are adopted for the non-SC calculations.
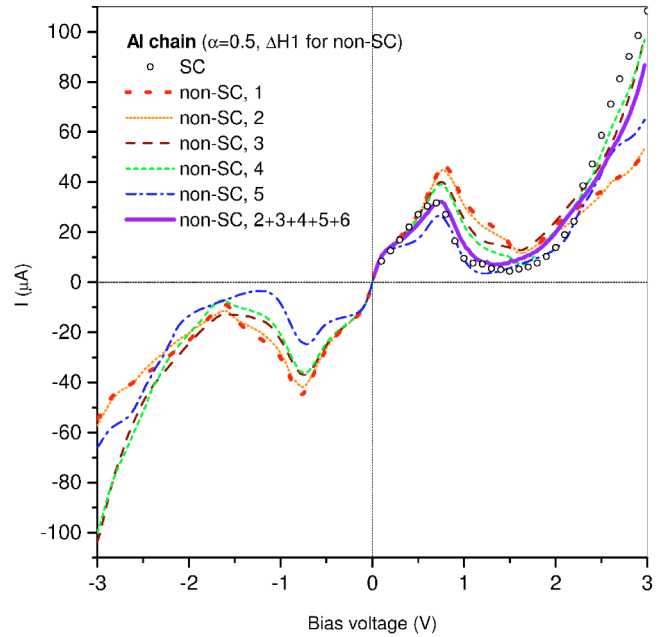

FIG. 10. (Color online) Calculated *I-V* curves for the Al chain system by the SC and non-SC approaches. The notations are similar to those in Fig. 9. 2+3+4+5+6 means a combined interface *X* [see Fig. 3(b)] and each interface bears 1/10 of a bias voltage (see Sec. III C).

technical choice. This indicates that for systems made of conductive molecular junctions coupled with metallic electrodes through hetero-interfaces, the present non-SC approach works quite well and can give nearly *quantitatively* correct answers.

For the aluminum chain (system B) the molecule-electrode contact is a homo-interface and the two Al atoms at the ends of the chain are at their bulk positions; therefore, the voltage drop is not localized around the contact regions. Consequently, the result from the non-SC approach with the interface *X* located around the contact region is not in good agreement with the SC result. In order to further verify our analysis for system B, we generalize the present non-SC approach for a voltage drop occurring at multiple points: We use a combined interface *X*=2+3+4+5+6 in which each layer bears a voltage drop of $V_b/10$. Because of the role played by the overlap matrix in Eqs. (23) and (24), the resulting voltage drop will occur over the entire device region. The calculated *I-V* curve by the generalized non-SC approach is given in Fig. 10 by a violet solid line. The overall agreement with the SC result is remarkably improved, indicating that our analysis is reasonable.

D. Limitation of the present approaches

For finishing the discussion we would like to point out the cases where the present method will not work. Obviously, the present method is only valid for steady state coherent electron transport through metal-molecule-metal systems; there are basically two cases where our method does not work: (1) electron transport in Coulomb blockade regime, for both the SC and non-SC approaches, and (2) cases where the

main feature of voltage drop is sensitive to the value of the bias voltage itself, for the non-SC approach. In the first case, the contact barrier is so high that the molecule and the leads are essentially separated, and as a result, the molecular chemical potential is generally different from the Fermi energies of the leads even under zero bias. Because in our DFT+NEGF approach there is only one Fermi energy under zero bias, it will fail in this case. The second case is just the opposite to that assumed in our non-SC approach. We do not know at this moment what systems will have this behavior or whether such kind of systems exist. But this can be easily checked by doing self-consistent calculations for several different bias voltages within the bias range interested.

V. SUMMARY

A full self-consistent DFT-electronic-structure-based Green function method has been proposed and implemented for electron transport from molecular devices. Our method is simple and straightforward while strict. The implementation is very independent of the DFT electronic structure part; it can be easily combined with any electronic structure package using a localized basis set. In an effort to avoid the extremely burdensome computational cost for large systems or for I-V characteristic analysis, we developed an approximate non-self-consistent approach in which the change in effective potential caused by a bias in the device region of a system is approximated by the main features of the voltage drop.

As applications of our methods, we calculated the I-V curves for two different systems with different typical voltage drops: a carbon chain and an aluminum chain sandwiched between two aluminum electrodes. Our self-consistent results are in very good agreement with those from other calculations. For both systems the present non-SC approach can give results in good agreement with the self-consistent results, indicating that it is a good approximate method with high efficiency for I-V characteristic analysis[33] (more than one order of magnitude faster for moderate systems). It is straightforward to generalize this non-SC approach to deal with any kind of voltage drop situation.

[1] *An Introduction to Molecular Electronics*, edited by M. C. Petty and M. Bryce (Oxford University Press, New York, 1995).

[2] J. R. Heath and M. A. Ratner, Phys. Today **56** (5), 43 (2003).

[3] J. M. Tour, *Molecular Electronics* (World Scientific, Singapore, 2003).

[4] C. Joachim, J. K. Gimzewski, and A. Aviram, Nature (London) **408**, 541 (2000).

[5] A. Nitzan, Annu. Rev. Phys. Chem. **52**, 681 (2001).

[6] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[7] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[8] N. D. Lang, Phys. Rev. B **52**, 5335 (1995).

[9] N. D. Lang and Ph. Avouris, Phys. Rev. Lett. **84**, 358 (2000).

[10] M. Di Ventra, S. T. Pantelides, and N. D. Lang, Phys. Rev. Lett. **84**, 979 (2000).

[11] N. D. Lang and Ph. Avouris, Phys. Rev. B **64**, 125323 (2001).

[12] R. Landauer, Philos. Mag. **21**, 863 (1970).

[13] For instance, our test calculation for benzene molecule sandwiched between two Au electrodes showed that the molecule-electrode charge transfer depends significantly on the adsorption site of the molecule on the Au surface.

[14] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport, Optics of Semiconductors* (Springer-Verlag, Berlin, 1996).

[15] S. Datta, in *Electronic Transport in Mesoscopic Systems*, edited by H. Ahmed, M. Pepper, and A. Broers (Cambridge University Press, Cambridge, England, 1995).

[16] P. S. Damle, A. W. Ghosh, and S. Datta, Phys. Rev. B **64**, 201403 (2001).

[17] Y. Xue, S. Datta, and M. A. Ratner, Chem. Phys. **281**, 151 (2002).

[18] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).

[19] J. Taylor, H. Guo, and J. Wang, Phys. Rev. B **63**, 245407 (2001).

[20] C. Krzeminski, C. Delerue, G. Allan, D. Vuillaume, and R. M. Metzger, Phys. Rev. B **64**, 085405 (2001).

[21] J. Taylor, M. Brandbyge, and K. Stokbro, Phys. Rev. B **68**, 121101 (2003); Phys. Rev. Lett. **89**, 138301 (2002); S. K. Nielsen, M. Brandbyge, K. Hansen, K. Stokbro, J. M. van Ruitenbeek, and F. Besenbacher, *ibid.* **89**, 066804 (2002).

[22] W. Lu, E. G. Wang, and H. Guo, Phys. Rev. B **68**, 075407 (2003); C.-C. Kaun, B. Larade, and H. Guo, *ibid.* **67**, 121411 (2003); C.-C. Kaun, B. Larade, H. Mehrez, J. Taylor, and H. Guo, *ibid.* **65**, 205416 (2002); H. Mehrez, A. Wlasenko, B. Larade, J. Taylor, P. Grutter, and H. Guo, *ibid.* **65**, 195419 (2002); C. Roland, B. Larade, J. Taylor, and H. Guo, *ibid.* **65**, 041401 (2002).

[23] Y. Xue, S. Datta, and M. A. Ratner, J. Chem. Phys. **115**, 4292 (2001); Y. Xue and M. A. Ratner, Phys. Rev. B **68**, 115407 (2003); **68**, 115406 (2003); R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic, J. Appl. Phys. **92**, 3730 (2002).

[24] For example, J. Heurich, J. C. Cuevas, W. Wenzel, and G. Schon, Phys. Rev. Lett. **88**, 256803 (2002).

[25] S. Datta, W. Tian, S. Hong, R. Reifenberger, J. I. Henderson, and C. P. Kubiak, Phys. Rev. Lett. **79**, 2530 (1997).

[26] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).

[27] M. P. López Sancho, J. M. López Sancho, and J. Rubio, J. Phys. F: Met. Phys. **15**, 851 (1985).

[28] Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).

[29] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, J. Appl. Phys. **81**, 7845 (1997).

[30] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**,

3865 (1996).

[31] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[32] B. Larade, J. Taylor, H. Mehrez, and H. Guo, Phys. Rev. B **64**, 075420 (2001).

[33] After we submitted our paper, we noticed a very recent theoretical work [G. Stefanucci and C.-O. Almbladh, cond-mat/0310084 (unpublished)] in which, by building up a formally exact NEGF+TDDFT (abbreviation for time dependent DFT) scheme, the authors show that the steady state current in the long-time limit depends only on the asymptotic shape of the one-particle effective potential of TDDFT. Our numerical result from the non-SC approach is just consistent with this theoretical finding.