Novaes, Frederico D.; Silva, Antônio J. R. da; Fazzio, A.

Density Functional Theory Method for Non-Equilibrium Charge Transport Calculations: TRANSAMPA

Frederico D. Novaes, Antônio J. R. da Silva, and A. Fazzio
Instituto de Física, Universidade de São Paulo, CP 66318, 05315-970, São Paulo, SP, Brazil

We describe a procedure to calculate charge transport properties across a nanosystem. This scheme is based on a Green's Function formalism to treat a non-equilibrium problem, coupled to the Density Functional Theory to describe the electronic structure. As an illustration, we perform calculations for the charge transport across a (5,5) carbon nanotube with a vacancy.

Keywords: Charge Transport; Density Functional Theory; Non-Equilibrium Green's Function

I. INTRODUCTION

In the past few years it has been possible to measure the charge transport across nanometric systems, such as carbon nanotubes, metallic nanowires, and molecules[1]. All these systems have the same basic physical picture, with a central scattering region connected to a certain number of leads. Even though many fundamental concepts related to the understanding of how charge flows in such small systems are already well established[2], there are still quite a few challenges if one wants to quantitatively calculate a property such as an $I \times V$ curve.

Any theory with this predictive power should obey a number of requirements: i) a reliable description of the electronic structure properties of the atoms in the scattering region; ii) the treatment of the leads in the same footing as the scattering region; iii) it should have no adjustable parameters; iv) the self-consistent calculation of the charge redistribution within the scattering region due to the application of a voltage bias; v) it should do all this for a variety of different systems. Albeit with some drawbacks, as discussed later on, the Density Functional Theory Non-Equilibrium Green's Function (DFT-NEGF) approach[3–5] to charge transport calculations comes very close to satisfying a good number of these requirements.

One could say that DFT[6] is today the standard approach for ab initio electronic structure calculations of medium to large systems, including molecules, bulk solids, surfaces, nano-systems, etc. There are theorems that guarantee its solid foundation as an exact theory, in principle. To perform practical calculations, approximations are necessary. There are, however, excellent schemes that lead to reliable and predictive results for properties such as structural geometries, vibrational frequencies, energy barriers, and heats of reactions. All these properties are obtained from the total energy of the system, which is guaranteed to be a minimum of the density. The usual implementations of the DFT do not directly use an approximation of the total energy as a functional of the density, but employ a more reliable scheme that introduces an auxiliary set of one-particle-like equations, the Kohn-Sham (KS) equations. This scheme is introduced in order to calculate with acceptable accuracy the major part of the kinetic energy. It is important to stress that although the majority of the DFT calculations employ these single-particle KS equa-

tions, this does not mean that DFT is a mean-field theory. DFT for equilibrium situations is a full many-body calculation, and the only approximation is contained in the so-called exchange-correlation functional.

Even though these KS equations have an auxiliary role in the theory (the total charge density is obtained from the KS orbitals), their eigenvalues, *i.e.* the KS spectrum, are usually used as an estimate of the true excitation spectrum of the system, and experience shows that this is in general a good approximation, except for an underestimation of the gap in semiconductors and insulators (or the gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) in molecules). Moreover, the KS orbitals also provide a good approximation for the quasiparticle orbitals.

In DFT-NEGF the use of DFT to describe the electronic structure of the problem does not have the same solid foundation as in equilibrium calculations. The idea is to use the KS orbitals and eigenvalues, or in other words, the KS Hamiltonian, as a good description for the scattering-region Hamiltonian as well as for the leads Hamiltonian and for the coupling between the leads and the scattering region, *i.e.*, the system Hamiltonian is written as

$$H = \sum_{In} \varepsilon_{In} c_{In}^{\dagger} c_{In} + \sum_{\alpha} \varepsilon_{\alpha} d_{\alpha}^{\dagger} d_{\alpha} + \\ + \sum_{In\alpha} (V_{In,\alpha} c_{In}^{\dagger} d_{\alpha} + c.c), \quad (1)$$

where $c_{In}^{\dagger}(c_{In})$ creates (destroys) an electron in a KS orbital with KS energy ε_{In}, for the I lead ($I = L, R$), and $d_{\alpha}^{\dagger}(d_{\alpha})$ creates (destroys) an electron in a scattering region KS orbital with KS energy ε_{α}. The indices n and α generically represent all the necessary quantum numbers to label the eigenstates. As it can be seen, DFT is used here in a mean-field like spirit. There will be, therefore, limitations associated with this approximation. However, there is the enormous gain that the whole system is treated fully ab initio with a powerful method, and realistic systems can be described at an atomistic level.

We describe below the main ideas of the DFT-NEGF scheme (method) and the details of our implementation based on the SIESTA code[7], named TRANSAMPA [8]. We then present, as an illustration of the methodology, a calculation for the charge transport properties across a carbon nanotube with

a vacancy.

II. METHOD

A. Conductance Calculation

A given potential bias V applied between two electrodes will cause a current I to flow across the system.

The relation between I and V is provided by the conductance G, given by

$$G = \frac{I}{V}. \tag{2}$$

A general expression for the current I has been obtained by Meir and Wingreen [9] within the Non-Equilibrium Green's Function formalism[10], where arbitrary interactions can occur within the scattering region. For the particular situation where the Hamiltonian describing the scattering region has a non-interacting form, the formula for the current is given by

$$I = \frac{e}{h} \int_{-\infty}^{\infty} T(E)\left(n_{FD}^{L}(E) - n_{FD}^{R}(E)\right) dE, \tag{3}$$

where $T(E)$ is the transmission function [4, 5, 9],

$$T(E) = Tr[\Gamma_L \mathbf{G}^r(E) \Gamma_R \mathbf{G}^a(E)], \tag{4}$$

with,

$$\Gamma_{L(R)}(E) = i\left(\Sigma_{L(R)}^r(E) - \left(\Sigma_{L(R)}^r(E)\right)^{\dagger}\right), \tag{5}$$

$$\Gamma_{L(R)}(E) = -2Im[\Sigma_{L(R)}^r(E)], \tag{6}$$

and we have assumed time reversal symmetry. In the above formulas, $n_{FD}^{L(R)}(E)$ is the Fermi-Dirac distribution function for the left (right) electrode, and the definitions and discussion regarding all the Green's functions and self-energies used above will be presented in the next section.

In the limit $V \to 0, T \to 0$, where T is the temperature, we assume a linear regime, with a mean value $T(\bar{E})$ in the *small* range of integration in (3),

$$I \to \frac{e}{h} T(\bar{E})\left(E_F^L - E_F^R\right) = \frac{e^2}{h} T(\bar{E}) V, \tag{7}$$

where $E_F^{L(R)}$ is the left (right) chemical potential. The conductance may, thus, be evaluated with an equilibrium calculation,

$$G = \frac{2e^2}{h} T(E_F), \qquad V \to 0, T \to 0, \tag{8}$$

considering spin degeneracy[11].

The main goal of the present implementation, as in previous ones [3–5, 12], is to devise a scheme to calculate the above Green's functions, and thus the current, from a fully ab initio method. We then use a DFT formalism. As already mentioned, in DFT the total energy and other properties are obtained via auxiliary one-particle equations, the Kohn-Sham equations. Both the resulting one-particle orbitals as well as the energies have no formal use in DFT. Albeit this caveat, they are usually considered to be a good approximation to the true quasi-particle wave-functions and excitation spectrum, respectively. Note, however, that there is the well known gap-problem within the DFT[13], as well as problems with the current exchange-correlation potentials, such as self-interaction errors[13].

In a mean-field spirit, the DFT formalism is used together with the NEGF to calculate the Green's functions and then the current. In the DFT formalism, the potential in the Kohn-Sham Hamiltonian depends on the electronic density. Thus, there is the need to solve these non-linear equations self-consistently. For equilibrium situations, the density matrix is traditionally obtained via the Kohn-Sham orbitals, through an appropriate application of either open (finite systems) or periodic boundary conditions, plus the self-consistent solution of the Kohn-Sham equations. The procedure presented below describes how to obtain the density matrix self-consistently, for both equilibrium and non-equilibrium situations, using Green's functions and open boundary conditions, for infinite systems.

The model consists of inserting two (large) electrodes (leads), to the left and to the right of a (nanoscale) region whose transport properties will be investigated (see Fig. 1). This partitioning in different regions, and the possibility to work with finite size matrices to simulate an open, non-periodic infinite system, is made possible by the use of localized basis sets[4]. Here we use all the developments that have already been implemented in the SIESTA code, in particular, localized atomic orbitals (which are strictly zero beyond certain cutoff radii). We then extend the code to be able to simulate open, non-periodic systems.
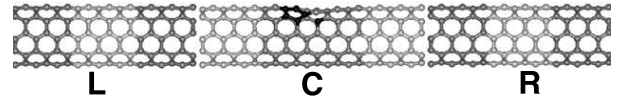

FIG. 1: Scheme for the geometrical structure used in the transport calculations, showing the left lead (L) and the right lead (R) connecting the scattering (central) region (C). The system displayed here is a (5,5) carbon nanotube, and there is a single vacancy in the scattering region. This procedure of splitting the system in three regions is a general feature of the DFT-NEGF implementation.

B. Green's function and the Density matrix

A periodic crystal can be described by a semi-infinite stack of principal layers (PL). A principal layer is defined as the smallest group of neighboring atomic planes such that only nearest-neighbor interactions exist between PLs [14]. Here, we extend this concept also for one-dimensional systems, such as carbon nanotubes, considering that a PL in this case is the

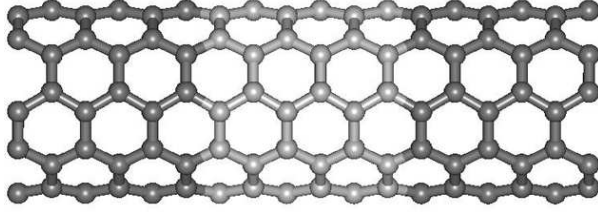smallest group of neighboring atoms, instead of infinite slabs (see Fig. 2).

FIG. 2: Illustration of a small region of a (5,5) carbon nanotube bulk, containing 3 principal layers (PL, see text for details), shown in alternate shades of gray. Note that each PL contains three bulk unit cells, composed of 20 carbon atoms. The size of the PLs is chosen such that, due to the strictly localized character of the basis functions, there is only coupling between nearest neighboring PLs. For example, in the (5,5) nanotube shown, the leftmost PL does not couple to the rightmost PL.

We may thus consider to be working with tri-diagonal matrices, where the *diagonal* elements are matrices representing a PL. Then, for an infinite system, a Hamiltonian (or Overlap - **S**) matrix, will have the general form,

$$\mathbf{H} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \cdots & H_{0,0} & H_{0,1} & 0 & 0 & \cdots \\ \cdots & H_{1,0} & H_{1,1} & H_{1,2} & 0 & \cdots \\ \cdots & 0 & H_{2,1} & H_{2,2} & H_{2,3} & \cdots \\ \cdots & 0 & 0 & H_{3,2} & H_{3,3} & \cdots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{9}$$

where each element of the above matrix is itself a matrix whose size is the number of basis set elements contained in a PL. For example, in the example shown in Fig. 2, each PL has three bulk unit cells. Each bulk unit cell has 20 C atoms, so that there are 60 C atoms in a PL. If we use 13 basis functions for each C atom, each PL has a total of 780 basis function, and this is the dimension of each one of these block matrices. All the $H_{I,I}$ matrices are identical, as are also the $H_{I,J}$ PL coupling matrices.

When we use Greek letters as matrix indices, they represent specific atomic orbitals, and when L, R or C are used, they represent whole regions, the left or right electrodes (that are infinite) and the central (scattering) region, respectively. Bold letters symbolize matrices. Here we emphasize that the term *Bulk* means a periodic infinite system, and that it may not be a three dimensional one (as the example in Fig. 2, an infinite carbon nanotube that simulates the leads).

1. Equilibrium Situation

We start by considering how to solve a DFT problem for an open system, still in an equilibrium situation, *i.e.*, when there is no potential bias between the left and right leads.

The usual approach for this type of situation is to use a supercell approximation, where the "defect" region, plus a bulk buffer region, are periodically repeated in space. This artificial periodicity usually provides quite accurate results. Care must, however, be taken in order to prevent (or at least reduce) that the spurious interactions between images affect the final results/conclusions. In some cases, such as carbon nanotubes, we have observed that it is quite hard to reduce these interactions to an acceptable level, since the periodically repeated defects introduce a periodic potential that opens up gaps at every band degeneracies at the Brillouin zone boundaries, which are mainly due to the zone folding due to the supercell approximation. In cases like these, the use of Green's functions to exactly treat an open boundary situation is quite advantageous. This type of application is a quite general use of Green's functions[15, 16], and is not particularly related to NEGF. Moreover, the construction outlined below can be used as part of a procedure to obtain the transmittance in the linear response regime.

We start with the formal definition of the retarded Green's function [15],

$$\left(E^+\hat{1} - \hat{H}\right)\hat{G}^r(E) = \hat{1}, \tag{10}$$

where

$$E^+ = \lim_{\delta \to 0^+} E + i\delta. \tag{11}$$

Representing the set of basis orbitals by $\{|\phi_\mu\rangle\}$, the *Dual* set, $\{|\phi^\mu\rangle\}$, is such that

$$\langle\phi^\mu|\phi_\nu\rangle = \delta^\mu_\nu. \tag{12}$$

Using the completness relation,

$$\hat{1} = \sum_\mu |\phi_\mu\rangle\langle\phi^\mu| = \sum_\mu |\phi^\mu\rangle\langle\phi_\mu|, \tag{13}$$

we obtain the matrix form of (10),

$$\sum_\lambda \left(E^+ S_{\mu,\lambda} - H_{\mu,\lambda}\right) G^r_{\lambda,\nu}(E) = \delta_{\mu,\nu}, \tag{14}$$

with

$$S_{\mu,\nu} = \langle\phi_\mu|\phi_\nu\rangle, \tag{15}$$

$$H_{\mu,\nu} = \langle\phi_\mu|\hat{H}_{KS}|\phi_\nu\rangle, \tag{16}$$

$$G^r_{\mu,\nu}(E) = \langle\phi^\mu|\hat{G}^r(E)|\phi^\nu\rangle, \tag{17}$$

where $\hat{H}_{KS}$ is the Kohn-Sham Hamiltonian. The matrices **H** and **S** are obtained using the SIESTA code[7], which is based on DFT.

Usually, for a periodic system, in order to obtain the solution of the equation

$$\hat{H}_{KS}|\psi_i\rangle = E_i|\psi_i\rangle, \tag{18}$$

one solves the generalized eigenvalue problem,

$$\sum_\lambda H_{\mu,\lambda} c_i^\lambda = E_i \sum_\lambda S_{\mu,\lambda} c_i^\lambda \tag{19}$$

with

$$c_i^\lambda = \langle \phi^\lambda | \psi_i \rangle. \tag{20}$$

From this, one can obtain the density matrix,

$$\rho_{\mu,\nu} = \sum_i c_i^\mu c_i^{\nu*} n_{FD}(E_i), \tag{21}$$

where $n_{FD}(E)$ is the *Fermi-Dirac* function. Since we now partition the system in three regions, left and right infinite electrodes and a central (scattering) region, this is no longer possible.

Since $G_{\mu,\nu}^r(E)$ can be written in a *spectral* form[4, 15],

$$G_{\mu,\nu}^r(E) = \sum_i \frac{c_i^\mu c_i^{\nu*}}{E^+ - E_i}, \tag{22}$$

and using the relation (we consider that x_0 lies in the interval [a,b]), we have

$$\int_a^b \frac{f(x)}{x - x_0 \pm i\delta}\, dx = PV\left[\int_a^b \frac{f(x)}{x - x_0}\, dx\right] \mp i\pi f(x_0), \tag{23}$$

with

$$PV\left[\int_a^b \frac{f(x)}{x - x_0}\, dx\right] = \lim_{\delta \to 0}\left[\int_a^{x_0-\delta} \frac{f(x)}{x - x_0}\, dx + \int_{x_0+\delta}^b \frac{f(x)}{x - x_0}\, dx\right], \tag{24}$$

which is the *Cauchy principal value*. We see that

$$\int_{-\infty}^\infty \left(G_{\mu,\nu}^r(E) - G_{\mu,\nu}^a(E)\right) n_{FD}(E)dE = -2i\pi\rho_{\mu,\nu}, \tag{25}$$

where $G_{\mu,\nu}^a$ is the advanced (matrix) Green's function, defined in similar way as the retarded function, with E^+ being replaced by E^-,

$$E^- = \lim_{\delta \to 0^+} E - i\delta. \tag{26}$$

Considering that the system has time reversal symmetry, the retarded and advanced matrices are related,

$$\mathbf{G}^a(E) = \left(\mathbf{G}^r(E)\right)^\dagger = \left(\mathbf{G}^r(E)\right)^*, \tag{27}$$

and we obtain the (well-known) expression that is used to calculate the equilibrium density matrix $\rho_{C,C}$,

$$\rho_{C,C} = -\frac{1}{\pi}Im\left[\int_{-\infty}^\infty G_{C,C}^r(E)n_{FD}(E)dE\right]. \tag{28}$$

In this way, we see that this procedure is entirely equivalent to the usual scheme where one first obtains the Kohn-Sham orbitals, and from them the density matrix. However, we can now obtain the density matrix of an open system.

2. *Non-Equilibrium*

If the system is in non-equilibrium, the expression to obtain the density matrix is more involved than equation (28). We then use the NEGF technique, which is developed following the contour path formalism to calculate the expectation values of time dependent operators[17], in particular, Green's Functions, like the retarded and correlation (or lesser) Green's Function, which are defined, respectively, as

$$G^r(x,x') = -i\theta(t-t')\langle[\hat{\psi}_H(x),\hat{\psi}_H^\dagger(x')]_+\rangle, \tag{29}$$

$$G^<(x,x') = i\langle\hat{\psi}_H^\dagger(x')\hat{\psi}_H(x)\rangle, \tag{30}$$

where $x \equiv (\mathbf{r},t)$, and $\hat{O}_H(t)$ is a general operator $\hat{O}$ in the Heisenberg picture. More details can be found in references [18, 19]. For a derivation of the equations in the case of a representation in a non-orthogonal basis set, see the work of Thygesen [20].

It is established that, if the *central* (scattering) region is in contact with two reservoirs that do not interact directly with each other, but that are each one in thermodynamic equilibrium characterized by a different electrochemical potential, the density matrix of the central region can be obtained by

$$\rho_{C,C} = \frac{1}{2\pi}\int G_{C,C}^r(E)\big(n_{FD}^L(E)\Gamma_L(E) + n_{FD}^R(E)\Gamma_R(E)\big)G_{C,C}^a(E), \tag{31}$$

where all the matrices have the size of the number of basis elements of the central region. All the matrices above will be properly defined below.

C. Retarded Green's Function of the Central Region

Now we need a scheme to calculate the retarded Green's function of the central region, and the self-energies that represent the coupling to the leads. Due to the fact that a localized set of basis functions is used, the matrices $\mathbf{H}$ and $\mathbf{S}$ can be written in tri-diagonal form. This makes it possible to obtain an expression for the retarded matrix Green's function for the central region[4]. Note that the left and right electrodes do not have a direct interaction, that is $H_{L,R} = S_{R,L} = 0$ [21]. We begin defining the energy dependent matrix,

$$h_{A,B} = E^+ S_{A,B} - H_{A,B} \; (A \text{ and } B = L,C,R). \quad (32)$$

It is then possible to show that the retarded Green's function of the central region can be written as [4]

$$G^r_{C,C}(E) = \left(h_{C,C} - \Sigma^r_L(E) - \Sigma^r_R(E) \right)^{-1}, \quad (33)$$

where the self-energies are defined as

$$\Sigma^r_L(E) = h_{C,L} \, g^r_{L,L}(E) \, h_{L,C}, \quad (34)$$

$$\Sigma^r_R(E) = h_{C,R} \, g^r_{R,R}(E) \, h_{R,C}. \quad (35)$$

These self-energies depend on the leads Green's functions,

$$g^r_{A,A}(E) = \left(E^+ S_{A,A} - H_{A,A} \right)^{-1} \; (A = L,R). \quad (36)$$

This scheme would seem impractical, since equation (36) involves, in principle, the inversion of a *semi-infinite* matrix. However, the inter-block matrices $h_{C,L}$, $h_{L,C}$, $h_{C,R}$ and $h_{R,C}$, are non zero only for a finite number of elements. It is then necessary to find only a finite *portion* of $g_{L,L}$ and $g_{R,R}$, called the surface Green's function. In this implementation we have used the recursive algorithm described in Ref. [14], although other options can also be employed[12].

D. Density matrix of the scattering region

Either in the case of equilibrium or non-equilibrium, the density matrix of the central region is obtained via an integration over the energy variable.

The importance of the expression (28) is that, for a complex variable Z in place of the real E, $\mathbf{G}^r(Z)$ is analytic for $Im[Z] > 0$, and that this function is smoother for complex values of Z. This integral can then be evaluated using a complex contour path[5], considering that $n_{FD}(E)$ has poles at $Z = E_F + i(2n+1)\pi k_B T$, where k_B is the Boltzmann constant and T is the *electronic temperature* used in $n_{FD}(E)$ (formally, it should be a step function $\theta(E_F - E)$ instead of $n_{FD}(E)$, which is used to facilitate numerical convergence, and expected not to significantly alter the results). Using equations (28) and (33), we can determine $\rho_{C,C}$ self consistently, as described below.

When there is a voltage drop applied to the system, equation (31) involves both the retarded (analytic for $Im[Z] > 0$), as well as the advanced (analytic for $Im[Z] < 0$) Green's Function. This makes it impossible to use a complex contour. However, considering that

$$\Gamma(E) = \Gamma_L(E) + \Gamma_R(E) = i[\Sigma^r(E) - \Sigma^a(E)], \quad (37)$$

and adding and subtracting $G^r n^L_{FD} \Gamma_R G^a$ to (31), we have

$$\rho_{C,C} = -\frac{1}{\pi} Im \left[\int_{-\infty}^{\infty} G^r_{C,C}(E) n^L_{FD}(E) dE \right] + $$
$$\frac{1}{2\pi} \int_{-\infty}^{\infty} G^r_{C,C}(E) \Gamma_R(E) G^a_{C,C}(E) \quad (38)$$
$$\left(n^R_{FD}(E) - n^L_{FD}(E) \right) dE.$$

The first part of (38) can be done in a complex contour, just like in the equilibrium case, whereas the second part has to be done with real values of E, but for a limited range of energies because of the cancelation of the Fermi-Dirac distribution functions.

E. Calculation Procedure

The calculation procedure is divided in two steps: *i) Electrode (Bulk) calculation; ii) Scattering-region calculation.*

1. Electrode calculation

An electrode calculation consists of a *standard* super-cell (periodic boundary conditions) setup, where the cell is composed of an electrode principal layer, with all atoms in the "bulk" electrode geometry. Once self consistency is achieved, we can calculate the electrode self energies (equations (34) and (35)), which are stored in a file for further use. We also store the matrix elements of the Hamiltonian and density matrix (those that will be used in the second step, as explained below). The advantage of storing information in a file is that, for different scattering regions with the same electrodes, we do not have to recalculate the electrode part. The disadvantage is that these files can be quite large.

2. Scattering-Region Calculation

We then set up a geometry containing: i) one PL of each electrode on each side; ii) a middle region with the scattering region of interest plus enough extra layers of electrodes on each side in such a way that all the screening will occur within this region, i.e., the Hamiltonian elements and charge density will already have the bulk value in the PL regions defined in i). This middle region is sometimes referred to (like we do here) as the *extended molecule*[4]. The whole composition is what we call the central region (Fig. 3).

As a zero-order approximation, the density matrix is first obtained considering this setup to be periodic, that is, we perform a *standard* SIESTA calculation[7] within the supercell approximation. Once self consistency is attained, we turn to the steps of the Green's function methodology. Within this procedure, iterative cycles are performed where the density matrix is obtained via integration of the central region matrix Green's functions. Each one of these cycles consists in the following steps:
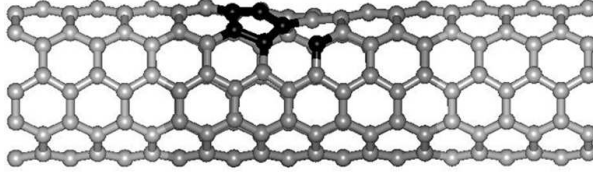
FIG. 3: Illustration of a central region geometry used in the calculations of the transmittance properties of a vacancy in a (5,5) carbon nanotube. The lighter gray atoms represent the two PLs (left and right) of the central region. The black spheres represent the atoms in the pentagon that is formed after the vacancy is created, which contains two nearest neighbor atoms to the vacancy (NN_V), plus the third NN_V.

1. Given $\rho_{C,C}$, the elements whose indices belong to the electrode regions (either left or right) are substituted by the corresponding electrode (bulk) values, including the inter-cell elements. This specifies a real-space electronic density $\rho(\mathbf{r})$, which is used to calculate a new $H_{C,C}$ matrix.

2. Given the new $H_{C,C}$, the *electrode matrix elements* (as $\rho_{C,C}$) are substituted by the (bulk) stored values. Using the previously calculated self energies, the matrix appearing in (28) is inverted, then $G_{C,C}$ is obtained giving a new $\rho_{C,C}$.

This procedure is repeated until a desired accuracy of self-consistency is achieved, defined by the relation

$$|(\rho_{\mu,\nu})_{i+1} - (\rho_{\mu,\nu})_i| < \Delta\rho_{max}, \qquad \mu, \nu \in C, \qquad (39)$$

where $(\rho_{C,C})_i$ is the *i-th* iteration density matrix.

Additional care is taken regarding the Hartree potential, which is obtained in SIESTA with a fast Fourier transform (FFT) algorithm[5, 7], with the choice of setting its average value to zero. This means that

$$\int_{supercell} V_H(\mathbf{r})\, d\mathbf{r} = 0. \qquad (40)$$

Since the forms of $V_H(\mathbf{r})$ are different within the supercells used in bulk and central region calculations, even with $\rho(\mathbf{r})$ *exactly* equal (at the PLs of the central region) to the bulk values, the solutions to the Hartree potential in these regions could be shifted in order to satisfy equation (40). We then compare the average value of $V_H(\mathbf{r})$ at regions close to the border of the bulk and at the central region of the supercells. The central region $V_H(\mathbf{r})$ is then shifted so that at the PL regions it matches the bulk values.

When a bias is applied, we consider that

$$V_H(\mathbf{r}) \rightarrow \tilde{V}_H(\mathbf{r}) + V_L, \qquad z < z_L; \qquad (41)$$

$$V_H(\mathbf{r}) \rightarrow \tilde{V}_H(\mathbf{r}) + V_R, \qquad z > z_R, \qquad (42)$$

where z_L and z_R define the borders between the extended molecule and the left and right PLs within the scattering region, respectively. We are considering that the transport occurs along the z direction. $\tilde{V}_H(\mathbf{r})$ is the Hartree potential solution obtained via a FFT [5]. All this means that (using atomic units)

$$H_{II} \rightarrow H_{II} + S_{II}V_I, \qquad I = L, R, \qquad (43)$$

and

$$\Sigma_I^r(E) \rightarrow \Sigma_I^r(E - V), \qquad I = L, R. \qquad (44)$$

In the extended molecule we have

$$V_H(\mathbf{r}) \rightarrow \tilde{V}_H(\mathbf{r}) + V_L + \left(\frac{V_R - V_L}{z_R - z_L}\right)(z - z_L), \qquad (45)$$

which satisfies the Poisson equation with the appropriate boundary conditions.

F. Spin Polarized Calculations

We perform spin polarized calculations as in a regular SIESTA calculation, namely, we obtain two different density matrices $\rho_{C,C}^\alpha$ and $\rho_{C,C}^\beta$, for the spin components α and β along the z-axis. These are used to obtain the DFT Kohn-Sham effective potentials[7] $v_{eff}^\alpha(\mathbf{r})$ and $v_{eff}^\beta(\mathbf{r})$.

G. General considerations

To end this section we briefly comment on a few relevant aspects.

i) Matrix inversion algorithm: Inversion of (large) matrices is a crucial part of the Green's functions techniques, and it is where numerical problems might occur. These matrices can be numerically close to being singular, and depending on the algorithm used, the process can easily fail. Different ways of handling this step might be quite important[12];

ii) Numerical value of the infinitesimal δ: In principle, it is an infinitesimal; in practice, it is a *small* number. We usually use a value of 1×10^{-5} Ry;

iii) Calculating the DOS and PDOS in the extended molecule region: For this calculation we use the relation

$$N(E) = -\frac{1}{\pi} Tr\big[Im[\mathbf{GS}]\big], \qquad (46)$$

where $N(E)$ is the density of states, and the trace is taken either over the whole set of orbitals (DOS) or over specific orbitals (PDOS) that belong to the extended molecule region;

iv) Forces, estimating and relaxing: We here relax the geometry considering the setup to be periodic, before performing the transport calculation, like in a regular supercell calculation. Since even for zero bias (as it is done in the present work later on for carbon nanotubes) it is important to

perform an iteration procedure to converge the density matrix for open systems, it might also be important to evaluate the forces and eventually relax again the atoms, within the open boundary condition. This step may be more critical when a (large) bias is present. So far this is not performed in our implementation, and we use the geometry mentioned above as an approximation;

v) Mixing the DM or H: In standard SIESTA calculations, the iterative process involves the mixing of the density matrix obtained from different iteration steps. Apparently, for open systems, mixing the Hamiltonian matrix works better [5, 12]. We are currently using this latter option, even though TRANSAMPA can handle both types of mixing.

III. RESULTS

The main goal of the following calculations is to compare the TRANSAMPA results with the ones obtained with a similar, well-established implementation, TranSIESTA [5, 22]. For this purpose, we briefly study the transmittance of a single vacancy in a (5,5) metallic carbon nanotube. Carbon nanotubes[23–25] have been attracting a great deal of attention as a promising material for nanoscale applications [26]. It has been recently shown that defects can alter in a dramatic way the transport properties of carbon nanotubes[27]. Even though it is believed that double vacancies are responsible for these large changes in the transport properties of metallic nanotubes, the influence of single vacancies in the transmittance has also been theoretically investigated[27, 28].
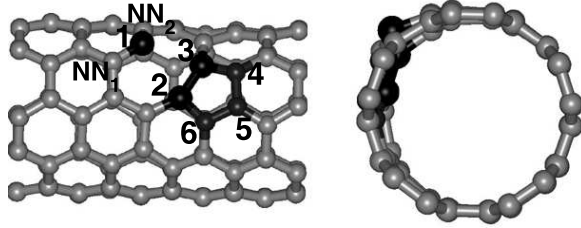

FIG. 4: Two different views of a detail of the (5,5) nanotube with a single vacancy, lateral (left) and frontal (right). The nearest-neighbor atoms to the vacancy are marked from 1 to 3, and are represented by larger spheres, and the atoms that reconstruct forming a pentagon are marked from 2 to 6. The relevant interatomic distances are given in Table I.

We begin by studying the single vacancy using a regular supercell approximation. The calculations are similar to what we have performed before for doped single wall carbon nanotubes (SWNT) [29–33]. They are all based on first-principles density-functional theory calculations using the SIESTA code[7], which performs a fully self-consistent calculation solving the standard Kohn-Sham equations using a numerical, strictly localized basis set to expand the Kohn-Sham orbitals. For the exchange and correlation energy and potential we use the PBE-GGA [34]. In the geometry relaxation calculations we have used a split-valence double-zeta

TABLE I: Interatomic distances (Å) between the marked atoms in Fig. 4. As a reference, the distance between two carbon atoms in a pristine (5,5) SWNT is also shown ($(CC)_{bulk}$). The number in parenthesis for the 2-3 distance is the distance between these atoms in a pristine nanotube, *i.e.*, before relaxation.

NN_1-1	NN_2-1	2-3	3-4	4-5	5-6	6-2	$(C-C)_{bulk}$
1.42	1.42	1.59 (2.47)	1.48	1.43	1.44	1.44	1.45

basis set with a confining energy shift of 0.1 eV[35]. Standard norm-conserving Troullier-Martins pseudopotentials[36] are used. A cutoff of 250 Ry for the grid integration was used to represent the charge density. We only considered a (5,5) SWNT (diameter of 6.92 Å). Our supercell has a lateral separation of 19 Å between image tube centers to make sure that they do not interact with each other. The supercell used in the geometry relaxations had 220-1=219 atoms (11 unit cells), which resulted in the value of 27.445 Å for the distance between a vacancy and its image in the next supercell. This is the structure shown in Fig. 3. We have used 5 Monkhorst-Pack **k**-points for the Brillouin zone integration along the tube axis, and the structural optimizations were performed using the conjugate gradient algorithm until the residual forces were smaller than 0.05 eV/Å. The atoms shown in lighter gray in Fig. 3 were held fixed at their bulk positions throughout the calculations. They will be the PL atoms within the central region.

A detail of the final geometry is shown in Fig. 4, both from a lateral as well as from a front view. The three atoms that are nearest-neighbors to the vacancy are shown as larger spheres, and are marked from 1 through 3. Atom 1 moves outward with respect to the original diameter, and atoms 2 and 3 get closer together and move inward. The distance between these atoms in a pristine (5,5) SWNT is 2.47 Å, since they are originally second nearest neighbors. After the relaxation, this same distance becomes 1.59 Å, and there is the formation of a pentagon, as indicated in Fig. 4. This pentagon suffers an inward relaxation, as can be seen in Fig. 3. This flexibility of SWNT that allows these inward relaxations with relatively low energy cost has as a consequence that defects can have a much lower formation energy as compared to graphene sheets[37]. All the other relevant distances are presented in Table I.

We now turn to the charge transport calculations. All the details of the calculation, such as basis set, energy shift and real space grid, are the same as in the geometry relaxation calculations. In fact, the size of the supercell was suggested by the requirements of the transport calculations, since due to the localization range of the used basis set, the PL region needs to have three bulk unit cells[38]. Therefore, the central region contains two PL regions, one to the left and one to the right side of the extended molecule, which has 5 bulk unit cells (this is the region where the vacancy is located). For the leads (bulk) calculations we thus have to also use a supercell with three bulk unit cells, since this is exactly one PL region.

In Fig. 5, we show the transmittance of a situation where the central region is identical to a single PL region, *i.e.*, for an infinitely periodic ideal system. Note that these results can be obtained directly from the electrode calculation. In this case,
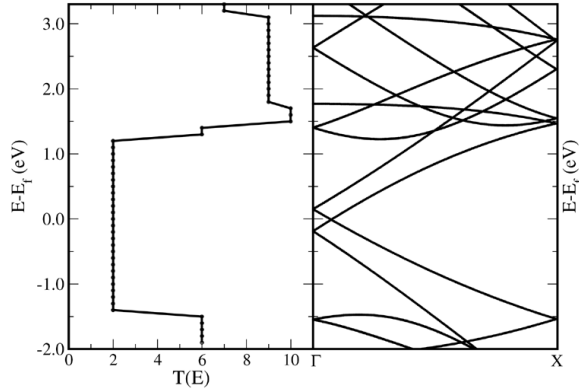
FIG. 5: Comparison between the transmittance and the band structure for the geometry used as leads in the transport calculation.

the transmittance at a certain energy is equal to the number of right (or left) propagating Bloch states with the same energy. These are equivalent to "positive" (or "negative") k_z, where **k** is the wave vector characterizing a Bloch state in the Brillouin zone (BZ) of this periodic system, and z is the *transport direction*. Thus, the transmittance in this case could also be directly extracted from the band structure, where $T(E)$ is equal to the number of bands that are crossed at this energy, increasing k_z from either the Γ point to BZ boundary (right propagating) or from the BZ boundary to the Γ point (left propagating)[39]. With this procedure, we can check that the surface Green's function is properly calculated, and know the upper bound for the transmittance obtained in the scattering calculation, specially at the Fermi energy.

Using the relaxed geometry described above, that already includes the electrode regions of the left and right leads that are needed, we did a *scattering region* calculation, and obtained the transmittance shown in Fig. 6. As can be seen, the agreement with TranSIESTA is very good.
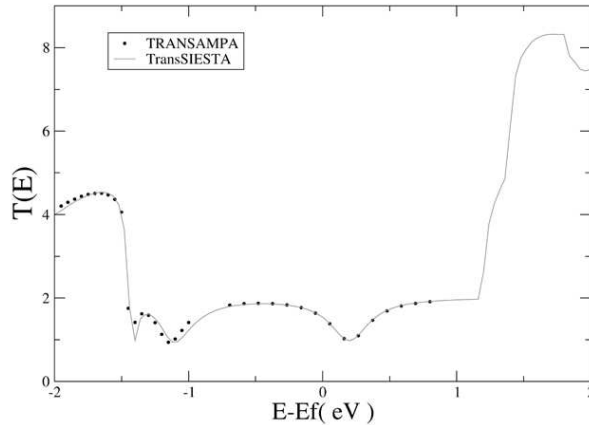

FIG. 6: Transmittance of a single vacancy in a (5,5) carbon nanotube. We show results obtained with TranSIESTA[5] and TRANSAMPA.

We can see that our calculation predicts a reduction of the

transmittance at the Fermi level by $\sim 25\%$, when a single vacancy is present ($T(E_F)$ drops from 2 G_0 to $T(E_F) \sim 1.5$ G_0, with $G_0 = 2e^2/h$). The important feature in the present calculation is not the particular value of $T(E_F)$, which may change with the use of a larger basis set, but the almost perfect agreement between TRANSAMPA and TranSIESTA.

IV. CONCLUSION

In summary, we presented the details of a numerical implementation of the DFT-NEGF method for charge transport calculations, which we have called TRANSAMPA. The DFT part of the method is based on the SIESTA code, and all the flexibilities of this program can be used in TRANSAMPA, such as multiple-ζ basis functions and spin polarized calculations. We illustrated the use of TRANSAMPA with the calculation of the transmittance of a single vacancy in a (5,5) SWNT. We compared our results with the TransSiesta code, and the agreement is excellent. Even though we only show results for a carbon nanotube, the method is very flexible, and can be used for a variety of situations of charge transport across nanostructures, such as, for example, molecules connecting two Au leads [40], or metallic nanowires with or without impurities. Of course the DFT-NEGF as implemented here is not the final answer to the challenging problem of quantitatively predicting the charge transport properties across nanosystems. However, it is an important first step that will allow us to learn a great deal about those physical problems, even though in some cases we can only learn why it does not work.

V. ACKNOWLEDGEMENTS

APPENDIX A: TIME REVERSAL SYMMETRY

If the system has time reversal symmetry, we have

$$\hat{T}^{-1}\hat{H}\hat{T} = \hat{H} \qquad (A1)$$

where $\hat{T}$ is the (antilinear) time reversal operator [41]. The operator definition of the Green's function,

$$(E^+\hat{1} - \hat{H})\hat{G}^r(E) = \hat{1}, \qquad (A2)$$

leads to

$$\left(E^+ \hat{1} - \hat{T}^{-1}\hat{H}\hat{T}\right)\hat{G}^r(E) = \hat{1},$$
$$\left(E^- \hat{T} - \hat{H}\hat{T}\right)\hat{G}^r(E) = \hat{T},$$
$$\left(E^- \hat{1} - \hat{H}\right)\hat{T}\hat{G}^r(E)\hat{T}^{-1} = \hat{1}, \quad \text{(A3)}$$
$$\left(\hat{G}^a(E)\right)^{-1}\hat{T}\hat{G}^r(E)\hat{T}^{-1} = \hat{1},$$

so that

$$\hat{T}\hat{G}^r(E)\hat{T}^{-1} = \left(\hat{G}^r(E)\right)^\dagger. \quad \text{(A4)}$$

Given an operator $\hat{Q}$[41] such that

$$\hat{T}\hat{Q}\hat{T}^{-1} = \begin{cases} \hat{Q}, & \hat{Q} \text{ real.} \\ -\hat{Q}, & \hat{Q} \text{ complex}, \end{cases} \quad \text{(A5)}$$

we can decompose $\hat{G}^r(E)$ into its real and imaginary parts,

$$\hat{T}\hat{G}^r(E)\hat{T}^{-1} = \left(\hat{G}^r(E)\right)^* = \left(\hat{G}^r(E)\right)^\dagger. \quad \text{(A6)}$$

[1] See articles and references in *I*ntroducing Molecular Electronics (Lecture Notes in Physics), Eds. G. Cuniberti, G. Fagas, and K. Richter, Springer-Verlag, Berlin, 2005.

[2] S. Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, Cambridge, 2001.

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

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

[5] M. Brandbyge, Jos-Luis Mozos, P. Ordejn, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).

[6] P. Hohenberg and W. Kohn, Phys. Rev **136**, 864B (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133A (1965); W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).

[7] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.-Cond. Matt. **14**, 2745 (2002).

[8] This acronym is related to two words, **TRAN**sport and **SAMPA**, which is by itself an acronym from the portuguese: **S**imulações **A**plicadas a **M**ateriais: **P**ropriedades **A**tomísticas. To obtain the code, please contact either fazzio@if.usp.br or ajrsilva@if.usp.br.

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

[10] L. V. Keldysh, Sov. Phys. JETP **20**, 1018 (1965); L. P. Kadanoff and G. Baym, *Q*uantum Statistical Mechanics, Benjamin, Menlo Park, 1962.

[11] R. Landauer, Philos. Mag. **21**, 863 (1970); M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985); M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).

[12] A. R. Rocha, V. M. Garcia-Surez, S. W. Bailey, C. J. Lambert, J. Ferrer, and S. Sanvito, Phys. Rev. B **73**, 085414 (2006).

[13] For a general discussion and references, see K. Capelle, cond-mat/0211443.

[14] M. P. L. Sancho, J. M. L. Sancho, and J. Rubio, J. Phys F:Met. Phys. **14**, 1205 (1984).

[15] E. N. Economou. *Green's Functions in Quantum Physics*. Springer-Verlag, Berlin, 1990

[16] A. R. Williams, Peter J. Feibelman, and N. D. Lang, Phys. Rev. B **26**, 5433 (1982).

[17] H. Haug, AP. Jauho.*Quantum Kinetics in Transport and Optics of Semiconductors*. Springer-Verlag, Berlin, 1996

[18] Mathias Wagner, Phys. Rev. B **44**, 6104 (1991)

[19] J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986)

[20] Kristian S. Thygesen, Phys. Rev. B **73**, 35309 (2006)

[21] This is not a restriction, since the *extended molecule* (as discussed further in the text) can always be made larger, including more layers of the bulk geometry, such that this condition is satisfied

[22] Since TranSIESTA is a commercial code, the results we present were obtained during its trial period.

[23] S. Iijima, Nature **354**, 56 (1991).

[24] S. Iijima e T. Ichihashi, Nature 363, 603 (1993).

[25] Bethune, D. S.; Klang, C. H.; de Vries, M. S.; Gorman, G.; Savoy, R.; Vazquez, J.; Beyers, R. Nature **363**, 605 (1993).

[26] See, for example, *Carbon Nanotubes: Synthesis, Structure, Properties and Applications*, Eds. M. S. Dresselhaus, G. Dresselhaus, and P. Avouris, (Topics in Applied Physics Series, Springer, Heidelberg, 2001).

[27] C. Gómez-Navarro, P. J. de Pablo, J. Gómez-Herrero, B. Biel, F. J. Garcia-Vidal, A. Rubio, and F. Flores, Nature Mat. **4**, 534 (2005).

[28] H. J. Choi, J. Ihm, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **84**, 2917 (2000).

[29] R. J. Baierle, S. B. Fagan, R. Mota, A. J. R. da Silva, and A. Fazzio Phys. Rev. B **64**, 085413 (2001).

[30] S. B. Fagan, A. J. R. da Silva, R. Mota, R. J. Baierle, and A. Fazzio Phys. Rev. B **67**, 033405 (2003).

[31] S. B. Fagan, R. Mota, A. J. R. da Silva, and A. Fazzio Phys. Rev. B **67**, 205414 (2003).

[32] S. B. Fagan, R. Mota, A. J. R. da Silva, and A. Fazzio J. Phys.: Condens. Matter. **16**, 3647 (2004).

[33] S. B. Fagan, R. Mota, A. J. R. da Silva, and A. Fazzio Nano Lett. **4**, 975 (2004).

[34] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[35] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. Garcia, and J. M. Soler, Phys. Status Solidi B **215**, 809 (1999).

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

[37] A. J. R. da Silva, A. Fazzio and A. Antonelli, Nano Lett. **5**, 1045 (2005).

[38] The size of the PL is determined in the electrode calculation, where it is chosen in such a way that, due to the strictly localized nature of the basis functions, one PL interacts only with its nearest neighbors. This is checked using the fact that the SIESTA code is designed to work with sparse matrices, and therefore easily provides which matrix elements are not strictly zero, allowing us to know if the supercell is interacting only with its two neighboring images.

[39] K. S. Thygesen, M. V. Bollinger, and K. W. Jacobsen, Phys. Rev. B **67**, 115404 (2003).

[40] R. B. Pontes, F. D. Novaes, A. Fazzio, and A. J. R. da Silva, J. Am. Chem. Soc. **128**, 8996 (2006).

[41] A. Messiah. *Quantum Mechanics*. North-Holland, Amsterdam, 1975