First-principles study of quantum tunneling from nanostructures: Current in a single-walled carbon nanotube electron source

Parham Yaghoobi, Konrad Walus, and Alireza Nojeh

Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

(Received 19 March 2009; revised manuscript received 14 July 2009; published 18 September 2009)

We present a first-principles calculation of the emission current in a single-walled carbon nanotube electron source. We have employed the nonequilibrium Green’s function and Fisher-Lee’s transmission formulation to describe electronic transport through the system. The simulation results reproduce the trends observed in experimental data closely and, in particular, the current saturation and deviation from the Fowler-Nordheim behavior. The proposed numerical approach is useful whenever a region of vacuum is present in the system Hamiltonian.

DOI: 10.1103/PhysRevB.80.115422

PACS number(s): 73.63.Fg, 73.40.Gk, 73.23.Ad

I. INTRODUCTION

Single-walled carbon nanotubes (SWNTs) are mechanically strong and physically stable structures that can be thought of as a rolled-up layer of graphene. They have diameters of about a nanometer and can have lengths of up to centimeters, resulting in very high aspect ratios. This leads to strong electric field enhancement at their tips. SWNTs are also capable of supporting current densities of up to \(10^9\) A cm\(^{-2}\), orders of magnitude higher than typical conductors such as copper and silver. These properties make them ideal candidates for electron emitters. Low turn-on voltages and brightness values of an order of magnitude higher than traditional cold field-emitters have already been demonstrated.

In 1928, Fowler and Nordheim proposed a model for field-electron emission (FEE) from surface emitters.\(^2\) The model assumes that the source has a metallic density of states and, upon the application of an electric field, electrons in the metal are accelerated toward a one-dimensional potential barrier. If the electric field is sufficiently high, some electrons will tunnel through the junction into the vacuum. The emitted current is exponentially dependent on the barrier width and, when graphically expressed on \(\ln[I(E^2)]\) vs \(E^{-1}\) scales, where \(I\) is current and \(E\) is electric field, will exhibit a linear relationship. This common way of expressing FEE is known as the Fowler-Nordheim (FN) plot. This model has been widely used by experimentalists to interpret the performance of electron sources, including nanotube-based emitters.\(^3,4\) However, the full applicability of this model to the emission process from SWNTs is not obvious and has been debated in the literature.\(^5\) The nanotube tip does not necessarily behave like a metal; electrons are not completely free and the effect of the tip’s atomic structure can be dominant. Moreover, a SWNT is a three-dimensional structure (3D) and not an infinitely wide surface; thus, a more detailed treatment than the FN model is required.

A SWNT emitter cannot be considered as a bulk and treated classically with macroscopic physical models. First-principles or mesoscopic models have to be used for accuracy. There has been significant work on first-principles modeling of SWNT emitters. These have shown the importance of tip orbitals in the emission process,\(^6\) the effective work function of SWNTs under a field,\(^7\) the behavior of the potential barrier,\(^8\) and the external field enhancement of SWNTs.\(^9\) There has also been attempts at estimating the emission current. In particular, in their valuable work, Han and co-workers calculated the many body wave function of the emitting nanotube using a pseudopotential-based, time-dependent density functional theory approach.\(^10,11\) However, in order to calculate transmission the authors used a one-dimensional potential barrier. Also, Buldum and Lu\(^12\) simulated the potential barrier of a (5,5) carbon nanotube under an electric field and calculated the emission current using a one-dimensional barrier.

To the best of our knowledge, an accurate calculation of the emission current using a first-principles method, which takes into account the full three-dimensional nature of the problem and incorporates vacuum in the Hamiltonian has not been reported previously. Here, for the first time we calculate the emission current by using a three-dimensional real-space first-principles Hamiltonian in the nonequilibrium Green’s function and Fisher-Lee formulation.\(^13\)

II. METHODOLOGY

In a typical FEE experiment using a SWNT as an emitter, a SWNT (a few micrometers long) is grown on a cathode electrode and placed in front of an anode, such as shown in Fig. 1(A). Our goal is to calculate the field-emission current from the SWNT using a fully first-principles approach, which is a very computationally intensive method and only allows a small section of the SWNT to be simulated. It is appropriate to choose an area close to the tip of the SWNT, as it has been shown that in an external field most of the nanotube remains an equipotential surface and the potential drop occurs mainly very close to the tip of the SWNT.\(^8\) As a result, the emission characteristics of the device are governed by the region that encapsulates the SWNT tip and vacuum, shown in Fig. 1(B). However, such a short section of the nanotube would not recreate the strong field enhancement that the entire nanotube would have due to its high aspect ratio. To compensate for this, we will use an electric field value that already contains the effect of the field enhancement (several hundred times stronger than the applied external field).\(^9\)

We assume that this section of the nanotube tip and vacuum [shown in Fig. 1(B)] and referred to as the channel
from now on] is connected to the continuation of the nanotube on the left side and vacuum on the right side. The current through such a channel can be calculated using the Landauer-Büttiker formula\(^1\)

\[
I = \frac{2e}{h} \int T(E) [f_1(E) - f_2(E)] dE,
\]

where \(f_1(E)\) and \(f_2(E)\) are the Fermi functions of the contacts, \(T\) is the transmission of the channel, \(q\) is the elementary charge, and \(h\) is Planck's constant, and \(E\) is energy.

The single particle nonequilibrium Green’s function (NEGF) has the following form in the spectral representation:\(^2\)

\[
G = [(E + i0^+) - H_{\text{channel}} - \Sigma_1 - \Sigma_2]^{-1},
\]

where \(H_{\text{channel}}\) is the channel Hamiltonian, \(i\) is the unit imaginary number, \(0^+\) is an infinitesimally small positive number, and \(\Sigma_1\) and \(\Sigma_2\) are the self-energy terms that account for the coupling of the semi-infinite contacts to the channel. The transmission of the channel can be computed using the Fisher-Lee relation

\[
T = T_1 [\Gamma_1 G \Gamma_2 G^\dagger],
\]

where \(\Gamma_1 = i[\Sigma_1 - \Sigma_1^\dagger]\) and \(\Gamma_2 = i[\Sigma_2 - \Sigma_2^\dagger]\).

This formalism provides a powerful tool for calculating the transport characteristics of nanoscale structures. The proper definition of the Hamiltonian operator, \(H_{\text{channel}}\), and the self-energy terms, \(\Sigma_1\) and \(\Sigma_2\), is essential in obtaining physically accurate results. Here, one challenge is how to deal with the vacuum region. In our approach, the three-dimensional Hamiltonian is built in a real-space basis and encapsulates both the SWNT and vacuum since it is straightforward to define vacuum in a real-space basis. This is done by discretizing the kinetic energy term using difference operators in Cartesian coordinates and projecting the potential term, obtained from some other first-principles packages that employ atomic basis sets. The potential in the vacuum region far away from the atoms is determined by the applied field. By encapsulating both the SWNT and vacuum as the channel, one can easily apply the Fisher-Lee formulation to calculate transmission.

In order to build this Hamiltonian in real space, we use the first-principles software package GAUSSIAN 03,\(^3\) which provides the advantage of different levels of theory and a built-in utility that can project the self-consistent field (SCF) potential onto a real-space basis. A good level of theory is the Hartree-Fock (HF) formalism, which has been shown to estimate the occupied levels accurately, which are the most relevant in our calculations.\(^4\) Density functional theory (DFT) would also be a natural choice, however, the most commonly used exchange-correlation functionals, namely the local density approximation (LDA) and the generalized gradient approximation (GGA), have been shown to overestimate the potential term under an electric field.\(^5,6\) TRANSIESTA\(^7\), an existing transport solver, is based on SIESTA,\(^8\) which lacks the flexibility of different levels of theory (is limited to DFT only), even for DFT it is limited to LDA and GGA. However, SIESTA offers the advantage of better scaling of computation time with the size of the system. In this work we used GAUSSIAN 03.

The self-energy terms in the real-space basis can be calculated by using the formulation developed by Appelbaum et al.\(^9\) This method assumes that the contacts are equipotential bodies, with potential values equal to the potential at the corresponding ends of the channel. Each self-energy term takes the form

\[
\Sigma = -t^2 g, \quad (4)
\]

where \(t = \frac{\hbar^2}{2ma}\), in which \(a\) is the lattice spacing in the real-space basis, \(m\) is the electron mass, and \(g = \frac{\Delta S - \Delta_S^0 - 4\sqrt{\Delta S}}{2\epsilon}\), where \(\Delta = E - H_{\text{D}}\) (\(E\) is energy and \(I\) is the identity matrix) and \(\Delta_S^0 = S^{-1} AS\). Here, \(H_{\text{D}}\) is the surface Hamiltonian for the three-dimensional contacts and \(S\) is the diagonalizing matrix of \(A\). More detail on the formulation of the Hamiltonian and the nonequilibrium Green’s function is provided in the Appendix.

Once a system is coupled to contacts, it experiences a shift in energy levels, as well as level broadening,\(^10\) and these effects can have an impact on the transport characteristics of the device. However, this effect will be minimal if a sufficiently long section of the nanotube is included in the device Hamiltonian (that is, if the contact is far enough from the tip). Computational limitations prevent a long nanotube to be included. Whether the length used in simulations is adequate can be verified by comparing the results to experiment or by performing the simulation with various nanotube lengths and determining the length at which the emission current no longer changes. Since the vacuum region is included in the Hamiltonian, the right contact should have little effect on the SWNT if there is sufficient distance between the tip and the right contact. This can be easily checked by examining the gradient of the potential in vacuum. If the local field is higher than the applied electric field then the vacuum level is perturbed by the molecule, but if the local field is equal to the applied field we know that we are suffi-
The simulated device, which encapsulates both the SWNT and vacuum. It consists of eight unit cells of a (5,5) SWNT capped with half a C_{60} molecule. The other end of the nanotube is terminated with hydrogen to avoid dangling bonds. Σ_1 is calculated 5 Å away from the hydrogen termination because the H-C bond produces an undesired dipole at the end as a result of the difference in electronegativity between carbon and hydrogen. Σ_2 is calculated 10 Å away from the tip of the structure, where the vacuum level has been found to be unperturbed by the molecule.

III. RESULTS AND DISCUSSION

As shown in Fig. 2, we have simulated 8-unit cells of a (5,5) SWNT capped with half of a C_{60} molecule. In order to speed up the geometrical optimization process in first-principles calculations in GAUSSIAN 03, we first relaxed this structure in NANOHIVE-1,25 a molecular dynamics software package, with adaptive intermolecular reactive empirical bond order (AIREBO) potentials.25 These potentials are widely used in the study of carbon and hydrogen structures.26,27 The system was relaxed again and the electronic structure was calculated under different electric fields using the restricted HF method and the 6–31g(d) basis set in GAUSSIAN 03. To compensate for the lack of field enhancement in our short nanotube, a scaled external electric field (value that already contains the effect of field enhancement) was used. However, note that the actual field distribution was calculated self-consistently in the simulation and the corresponding charge accumulation is illustrated in Fig. 3 (figure shows the Mulliken charge distribution). The hydrogen termination creates an undesired local dipole because carbon is more electronegative than hydrogen. As illustrated in Fig. 2 that section (about two and a half unit cells) of the device was excluded in constructing the Hamiltonian.

For each applied electric field and after the electronic structure calculations had converged, the SCF potential was projected onto a real-space basis with a grid spacing of 0.7 Å in x and y directions and 0.35 Å in z direction (the axis along the SWNT). At each applied electric field the left contact Fermi level was fixed at 4.5 eV below the vacuum level. This value was taken from the work of Zheng et al.,8 which presented an accurate calculation of the potential distribution along a micrometer-long capped (5,5) SWNT using a hybrid quantum/molecular mechanics method. The right contact Fermi level was fixed at the right boundary’s potential level. The energy integration range was from 10 eV above the left contact’s Fermi level to 10 eV below the right contact Fermi level. The Fermi functions at either contacts were evaluated at room temperature. Using the aforementioned formulation, the I-V characteristics of the structure was calculated (Fig. 4). The corresponding plot on FN scales is shown in Fig. 5, obviously deviating from a straight line and exhibiting a non-FN behavior. The current saturation behavior, which has been previously observed experimentally in SWNT emitters, is quite evident.28,29

Insight can be gained into this saturation behavior by examining the transmission spectrum (Figs. 6 and 7). At low
applied fields, as the relevant energy range in the Landauer-Büttiker integral increases (corresponding to an increase in the field), more “conduction channels” enter that range. This is reflected on the transmission spectrum (Fig. 6), where the shaded area shows the portion of the spectrum below the left contact Fermi level. In this circumstance, one would not expect current saturation, similar to a conventional electron emitter following the FN model. However, at high fields, the area under the transmission curve in the integration range remains relatively constant (Fig. 7), causing current saturation.

In order to better understand this saturation mechanism, we look at the behavior of the potential profile around the tip region of the SWNT. Figures 8 and 9 illustrate the potential profile along the SWNT axis for various electric fields in the pseudoexponential turn-on region of the I-V curve (0.1–0.4 V Å⁻¹), and the saturation region of the I-V curve (0.6–1.2 V Å⁻¹), respectively. The dotted line represents the highest occupied molecular orbital (HOMO) of the SWNT in each case. It can be seen that at low fields, the barrier width decreases with increasing the field. At high fields, however, the barrier width remains relatively constant with increasing the field, hinting at current saturation. This can be clearly seen on the plot of the barrier width vs field (Fig. 11).

For comparison the barrier width (Fig. 10) outside an infinite, planar electron emitter under electric field is $F \Phi - \varepsilon$, where $F$ is the work function of the material and $\varepsilon$ is the applied field. The work function of our carbon nanotube is 5 eV from our simulation. In Fig. 11, the solid line shows the behavior of the potential barrier width for the planar case and the dots are the simulated results at the tip of the carbon nanotube. What is most interesting is how the carbon

---

**FIG. 5.** The plot of the I-V characteristics of the simulated results (black squares) on FN scales. The dashed line shows a visual FN fit for comparison. The deviation from FN behavior at high field is apparent.

**FIG. 6.** Transmission spectrum at various low electric fields. The spectrum is shifted as necessary so that the left contact Fermi level is kept at 0 for all cases.

**FIG. 7.** Transmission spectrum at various high electric fields. The spectrum is shifted as necessary so that the left contact Fermi level is kept at 0 for all cases.

**FIG. 8.** Potential profile along the center of the SWNT in the turn-on regime. Dotted line is the HOMO level of the structure at: (A) 0.2, (B) 0.25, (C) 0.3, (D) 0.35, and (E) 0.4 V Å⁻¹.
nanotube reaches a relatively constant barrier width very rapidly. This is due to the aspect ratio of the nanotube, which greatly enhances the applied electric field at the tip. In other words, in contrast with the planar case, even at relatively low applied fields, the total field at the nanotube tip is already quite strong (due to enhancement by the nanotube), and strong emission occurs. Further increasing the external field does not change the barrier and emission current significantly.

A more complete picture can be obtained by also considering the potential barrier height. In fact, the tunneling current is directly dependent on both potential barrier height and width. Therefore, investigating the product of the two may be more illuminating in terms of seeing the overall behavior of emission current vs applied field. Figure 12 shows this product at the center of the carbon nanotube cap surface. It can be seen why saturation occurs at high electric fields. The product is large at low electric fields (both barrier height and width are large), but rapidly decreases with field and eventually, at high fields, remains relatively constant, consistent with the saturation of current. Also, a deep potential well exists at the tip, indicating that only a limited number of electrons can be accommodated in this well and made available for emission at any given time. These would explain the current saturation and the corresponding non-FN behavior shown in Fig. 5, which resemble the results obtained in our previous experiments and those performed by Kwo et al.33

So far, these deviations from the FN behavior have not been taken into serious consideration and have usually been overlooked by fitting a straight line to experimental results. This study could also explain the current saturation behavior observed in other sharp emitters like nanowires.30,31 Finally, we would like to emphasize that electronic transport is a nonequilibrium problem. Therefore, the Hamiltonian has to be constructed under nonequilibrium conditions. In the present work, although the Hamiltonian was constructed using the SCF method and coupled to contacts to account for the shift in energy and level broadening, the transport NEGF part was not solved self-consistently. Also, due to first-principles computational limitations, only a short section of the SWNT was simulated although the effect of field enhancement was considered by scaling the applied field and the actual potential distribution was calculated. Future work will include solving the NEGF self-consistently.
on a longer SWNT. We will also study the effect of various tip structures and adsorbates on emission current.

IV. CONCLUSION

A first-principles approach was applied to calculating current in a SWNT field emitter by encapsulating vacuum as part of the Hamiltonian. The transmission was calculated using a real-space basis and taking into account the full 3D nature of the problem. The results clearly demonstrate the commonly observed current saturation behavior in electron field-emission experiments. It was suggested that this saturation mechanism arises because of two main reasons. First, there exists a deep potential well that limits the number of electrons that can occupy the tip, and second, and more importantly, the dependence of the emission potential barrier on the number of electrons that can occupy the tip, and second, and more importantly, the dependence of the emission potential barrier on applied field is significantly less in the high-field saturation regime as compared to the pseudoexponential regime, due to the dominant local electric field distribution.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Canada Foundation for Innovation (CFI), and the British Columbia Knowledge Development Fund (BCKDF). The authors also thank Mona Berciu for insightful discussions on mesoscopic transport.

APPENDIX: REAL-SPACE FORMULATION AND NONEQUILIBRIUM GREEN’S FUNCTION

The problem of interest is a molecule-vacuum junction and how they can both be encapsulated in a single Hamiltonian. With existing first-principles software packages (i.e., GAUSSIAN, SIESTA, etc.), it is easy to define a molecular structure since they use an atomic orbital basis set. By moving the obtained Hamiltonian for this problem into the real-space basis set (delta functions on a discretized 3D space), one can easily add in the vacuum part. The Hamiltonian is defined as

\[ H = -\frac{\hbar^2}{2m} \nabla^2 + U, \]  

(A1)

where \( \hbar \) is the reduced Planck constant, \( m \) is the mass of an electron, and \( U \) is potential energy. In Cartesian coordinates the Hamiltonian can be defined as

\[ H = -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + U. \]  

(A2)

Using the central difference approximation of the second derivative,

\[ f''(x) \approx \frac{\delta^2 f}{\Delta^2} = \frac{f(x + \Delta) - 2f(x) + f(x - \Delta)}{\Delta^2}, \]  

(A3)

where \( \Delta \) is the grid spacing, one can define the Hamiltonian matrix (similar expressions to Eq. (A3) apply to the \( y \) and \( z \) direction). For example, for a three dimensional space with two discretized points in each dimension, the matrix would look like

\[
\begin{bmatrix}
* & t_x & t_y & t_z & 0 & 0 & 0 & 0 & 0 \\
0 & * & t_x & t_y & t_z & 0 & 0 & 0 & 0 \\
t_y & 0 & * & t_x & t_y & t_z & 0 & 0 & 0 \\
0 & t_y & 0 & * & t_x & t_y & t_z & 0 & 0 \\
0 & t_z & 0 & 0 & * & t_x & t_y & t_z & 0 \\
0 & 0 & t_z & 0 & 0 & * & t_x & t_y & t_z \\
0 & 0 & 0 & t_z & 0 & 0 & t_y & 0 & * \\
0 & 0 & 0 & 0 & t_z & 0 & 0 & t_y & 0 \\
0 & 0 & 0 & 0 & 0 & t_z & 0 & 0 & t_y \\
0 & 0 & 0 & 0 & 0 & 0 & t_z & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & t_z & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_z \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

where \( t_x = \frac{-\hbar^2}{2m_2 \Delta_x^2}, \) \( t_y = \frac{-\hbar^2}{2m_2 \Delta_y^2}, \) and \( t_z = \frac{-\hbar^2}{2m_2 \Delta_z^2}; \) \( \Delta x^2, \Delta y^2, \) and \( \Delta z^2 \) are the grid spacing in \( x, y, \) and \( z, \) respectively, and \( * = -2(t_x + t_y + t_z) + U. \) \( U \) is calculated using a first-principles program and projected onto this discretized 3D grid. For example, in GAUSSIAN 03, \( \Sigma \) is obtained using the Cubegen utility once the self-consistent solution is obtained.

Once the Hamiltonian is defined, the nonequilibrium Green’s function can be calculated using

\[ G = [(E + i0^+) - H - \Sigma_{\text{CNT}} - \Sigma_{\text{Vacuum}}]^{-1}, \]  

(A4)

where \( \Sigma_{\text{CNT}} \) is the self-energy of the carbon nanotube contact on the left side of the channel and \( \Sigma_{\text{Vacuum}} \) is the self-energy of the vacuum contact on the right side of the channel.

In the nonequilibrium Green’s function formulation, the self-energy terms are defined as \( \Sigma = \tau G_R \tau^*, \) where \( \tau \) is the coupling term between the channel and the contact and \( G_R \) is the Green’s function of the contact or the reservoir. In the real-space basis the coupling term \( \tau \) would only involve the nearest neighboring points. Therefore, we can redefine the self-energy term as \( \Sigma = \tau G_R \tau \), with \( \tau = -\frac{\hbar^2}{2m_2 \Delta^2} \) with \( a \) being the grid spacing. What Appelbaum et al.\textsuperscript{12} noticed was that since \( G_R = [(E + i0^+) - H_R]^{-1} \) and \( H_R = [H_{\text{surface}} - \Sigma] \) in the real-space basis, one can write an equation under the assumption that the contacts are at equipotential, with values equal to the potential at the corresponding ends of the channel. This turns out to give a simple quadratic equation of the form

\[ -\tau^2 G_R^2 + (E - H_{\text{surface}}) G_R - I = 0, \]  

(A5)

where \( I \) is the identity matrix. This equation can be solved for \( G_R \), and then the self-energy terms can be calculated.
FIRST-PRINCIPLES STUDY OF QUANTUM TUNNELING...

PHYSICAL REVIEW B 80, 115422 (2009)

*anojeh@ece.ubc.ca

24. B. Helfrich, NANOHIVE-1, Revision 1.2.0-b1, Nanorex, Inc.