The roles of apex dipoles and field penetration in the physics of charged, field emitting, single-walled carbon nanotubes

Jie Peng, Zhibing Li, Chunshan He, Guihua Chen, Weiliang Wang, Shaozhi Deng, Ningsheng Xu, Xiao Zheng, GuanHua Chen, Chris J. Edgcombe, and Richard G. Forbes

The roles of apex dipoles and field penetration in the physics of charged, field emitting, single-walled carbon nanotubes

Jie Peng,[1] Zhibing Li,[1,a)] Chunshan He,[1] Guihua Chen,[1] Weiliang Wang,[1] Shaozhi Deng,[1] Ningsheng Xu,[1,a)] Xiao Zheng,[2] GuanHua Chen,[2] Chris J. Edgcombe,[3] and Richard G. Forbes[4]

[1]*The State Key Laboratory of Optoelectronic Materials and Technologies Department of Physics, Sun-Yat-Sen University, Guangzhou, 510275, People's Republic of China*
[2]*Department of Chemistry, The University of Hong Kong, Hong Kong, China*
[3]*PCS Group, Department of Physics, Cambridge University, Cambridge CB3 OHE, United Kingdom*
[4]*Advanced Technology Institute, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom*

A 1 μm long, field emitting, (5, 5) single-walled carbon nanotube (SWCNT) closed with a fullerene cap, and a similar open nanotube with hydrogen-atom termination, have been simulated using the modified neglect of diatomic overlap quantum-mechanical method. Both contain about 80 000 atoms. It is found that field penetration and band bending, and various forms of chemically and electrically induced apex dipole play roles. Field penetration may help explain electroluminescence associated with field emitting CNTs. Charge-density oscillations, induced by the hydrogen adsorption, are also found. Many of the effects can be related to known effects that occur with metallic or semiconductor field emitters; this helps both to explain the effects and to unify our knowledge about FE emitters. However, it is currently unclear how best to treat correlation-and-exchange effects when defining the CNT emission barrier. A new form of definition for the field enhancement factor (FEF) is used. Predicted FEF values for these SWCNTs are significantly less than values predicted by simple classical formulae. The FEF for the closed SWCNT decreases with applied field; the FEF for the H-terminated open SWCNT is less than the FEF for the closed SWCNT but increases with applied field. Physical explanations for this behavior are proposed but the concept of FEF is clearly problematical for CNTs. Curved Fowler–Nordheim plots are predicted. Overall, the predicted field emission performance of the H-terminated open SWCNT is slightly better than that of the closed SWCNT, essentially because a C–H dipole is formed that reduces the height of the tunneling barrier. In general, the physics of a charged SWCNT seems much more complex than hitherto realized. © *2008 American Institute of Physics.*
[DOI: 10.1063/1.2946449]

I. INTRODUCTION

A. General background

The use of carbon nanotubes (CNTs) as field electron (FE) emitters has attracted many experimental and theoretical studies. This is mainly because CNTs have high aspect ratios, which allow operation at low applied voltage. For FE emission, the CNT is an interesting multiscale system, in that there is a strong interaction between its mesoscopic length and the nanoscale details of electronic structure at the CNT apex. The emission is sensitive to this electronic structure, which needs to be treated quantum mechanically.

A few years ago, it became possible to simulate long CNTs (up to several microns in length) by a hybrid approach involving quantum mechanics and classical analysis.[1,2] A (5, 5) single-walled CNT (SWCNT) closed with a fullerene hemisphere was simulated (we call this the "closed CNT"). This suggested that the superior emission behavior of a CNT

may involve not only a large field enhancement factor (FEF) but also barrier lowering due to field penetration (FP).

Recent improvements in our program allow a whole 1 μm long CNT, comprising about 80 000 atoms, to be simulated using only the quantum-mechanical procedures described in Refs. 1 and 2. Small inconsistencies related to the matching between the classical and quantum regions have thereby been removed. All results reported here have been regenerated in this way. We have also resimulated the closed SWCNT with similar computational effort. There were no qualitatively significant changes in outcomes.

The present paper originated as a study of the effects of hydrogen adsorption on FE from CNTs. There are contradictory conclusions in the literature. Zhou *et al.*[3] and Kim *et al.*[4,5] obtained the local density of states (LDOS) at the apex by *ab initio* methods; they found that the LDOS at the charge-neutrality level was suppressed by the hydrogen, and concluded that adsorption reduces emission. By contrast, Mayer *et al.*[6] calculated the tip barrier using a dipole and point charge model; they assumed that the tip barrier was reduced by the hydrogen, and concluded that adsorption enhances emission. Mayer[7,8] recently improved the model and

---
a)Authors to whom correspondence should be addressed. Electronic addresses: stslzb@mail.sysu.edu.cn and stsxns@mail.sysu.edu.cn.

illustrated the electrostatic potential around the CNT. There are also conflicting interpretations of related experiments.[9,10] Further study seemed useful, as our method could be extended to treat hydrogen adsorption.

To make theoretical comparisons, we first tried to simulate a non-H-terminated open SWCNT. This configuration is not expected to exist physically, because it has unsaturated dangling bonds. In practice, if it came into existence temporarily, we expect the bonds to rehybridize and the structure to reconstruct (most probably into a closed CNT). Our simulations produce results for the non-H-terminated open SWCNT for some applied fields. However, as expected, it is not a stable system and sometimes the results do not converge. We do not have confidence that these results have useful physical meaning, and do not present them. We think Mayer[7,8] was able to treat an open CNT successfully because his method was based on classical electrostatics rather than quantum mechanics. We considered it best to simulate an open-ended (5, 5) SWCNT terminated by hydrogen (H) atoms (we call this the "open CNT"), and compare this with the equivalent analysis of a closed SWCNT.

Interesting differences were found in the simulations. The least expected related to predicted FEFs (γ_s as defined here—see below). Application of the simple height/radius formula for a metal post with a hemispherical cap,[11] of apparently same dimensions as our closed CNT (length 1 μm and geometrical radius 0.344 nm) estimates the FEF as around 2910; a better classical estimate, derived from Eq. (20) in Ref. [11], is 1570. (For these estimates, the "geometrical radius" is defined by the positions of the carbon-atom nuclei.) Both our CNT configurations have γ_s significantly less than this. The closed CNT has γ_s in the range of 750–1000, with decreasing γ_s, with an increase in the applied macroscopic field F_M; the H-terminated open CNT has γ_s in the range of 370–450, with γ_s increasing with increase in F_M.

In looking for the physical origin of these results, we re-examined wider aspects of the physics of charged CNTs, including the role of FP and possible electric-dipole effects near/at the CNT apex. It soon became clear that the physics of a charged CNT is much more complex than hitherto realized. This paper presents our numerical results and sets them in the context of reasoned hypotheses about the physics of charged CNTs. As noted below, our theoretical method is not one of high numerical accuracy, but it should generate results that are qualitatively correct. We emphasize that our aim here is better qualitative understanding of CNT behavior.

B. Theoretical background

The normal FE convention is used that an electron extracting field is denoted by F and taken as positive. This quantity F is the negative of the quantity "electric field" defined in conventional electrostatics. Our model has a SWCNT standing upright on, and in good electrical contact with, one plate of two parallel plates. The applied macroscopic field F_M is the field between the plates, at large distance from the CNT. The name "barrier field" is used for the much higher local fields immediately above the CNT apex

that determine the tunneling barrier shape. Local field varies with position; more careful definitions are given later.

Conventional descriptions of FE and thermal electron emission, originally developed for metals, postulate the existence of an electron in the process of emission. This is often called the "external electron," but is called here the "departing electron." The total potential governing its motion is historically called the "motive."[12] The corresponding total potential energy (PE) is called here the "motive energy." The motive energy is what goes into the one-electron Schrodinger equation, and hence into approximate methods of solving it to obtain a tunneling probability. Conventional theory identifies two components.

The first is the so-called electrostatic[12] component of electron PE (EEPE). The EEPE relates to the potential experienced by a (hypothetical) vanishingly small test charge moving in the field created by the *total* emitter charge distribution, including the departing electron. (This may look illogical but in real situations, there is a high-voltage generator connected to the emitter that pushes a replacement electron into the system as the departing electron leaves.) Thus, the EEPE is the PE that would apply to the motion of an *extra* electron if the rest of the charge distribution in the metal or CNT did not react in any way to the presence of this electron.

The second component in the motive energy represents the system response to the departing electron. This is sometimes called the "correlation-and-exchange" (C&E) component. For metals, the C&E PE is almost always approximated (in the region where the tunneling barrier exists) by a classical image PE. For a flat planar surface of a free-electron metal, this leads to the so-called standard (Murphy–Good) Fowler–Nordheim-type (FN-type) equation[13,14] for the FE current density. For a curved emitter surface, with a slightly different barrier shape, this leads to a FN-type equation with slightly different correction factors in the exponent and the pre-exponential.[15,16]

In principle, if the same general ideas apply to FE from CNTs as applied to metals, then the motive energy for a CNT should contain both "electrostatic" and "C&E" components. However, there is continuing unpublished debate about whether, and how, a C&E PE component should be included in the motive energy for FE from a CNT. No consensus exists on these issues, so in this paper we calculate only the electrostatic component (the EEPE). This is a useful first step, often used for FE from CNTs, and was the strategy adopted by Fowler and Nordheim in their original work on FE from free-electron metals.[14] Qualitative effects associated with the CNT charge distribution should emerge correctly.

A brief description of the main calculation method is given in Sec. II. Section III then discusses the zero-field barrier shapes, and Sec. IV discusses the results found when the applied macroscopic field is high enough to cause electron emission by tunneling. Section V presents a more detailed discussion of FEFs. Section VI considers estimates of emission current. Section VII discusses some specific issues, and Sec. VIII summarizes our conclusions.

The following phrases are used with the technical meanings defined in Ref. [17]. The *excess electron density* is the

difference between the (simulated) ground-state electron density and the electron density that would exist if neutral atoms were brought together (in the absence of any applied field), with no change in their electron distributions. The *induced electron density* is the difference between the excess electron densities in the presence, and absence, of the applied field. The excess electron density in zero applied field represents chemically induced changes. The induced electron density represents further changes induced by applying an electric field. The excess electron density in nonzero applied field includes both chemically induced and electrically induced effects. For diagrams, we use the convention that an increase in electron density is represented as a positive change in density (i.e., the convention is the reverse of that used in classical electrostatics).

The reason for presenting information in the form of excess electron densities is that, in the theory of FE, all long-range electrical effects are due to small changes in the total charge distribution: using excess electron densities picks out the changes in the CNT electron density (as compared to an assembly of neutral atoms) and helps identify the physical and chemical causes of electrical effects.

II. THE SIMULATION METHOD

To simulate the SWCNT in realistic FE experimental conditions, a linear-scale algorithm is needed. A key decision is to divide the long CNT into subregions.[1,2] Each consists of 40 atoms, except the apex which uses 80 atoms. A subregion and its adjacent subregion(s) form a subsystem (the dangling bonds are saturated by hydrogen atoms). Each subsystem was simulated by the "modified neglect of diatomic overlap (MNDO)" semiempirical quantum-mechanical method.[18] The charges in other parts of the nanotube contribute interactions to the subsystem as if they were point charges. The principles of our general approach have been described elsewhere.[1,2] In the present paper, a total of 1600–2000 subregions are used in the quantum-mechanical part of the calculation.

A macroscopic field applied to the CNT would, if no electron redistribution occurred, create an electrostatic potential variation along the length of the CNT. As in classical electrostatics, this induces electrons to move from the substrate into the CNT. We assume that quasithermodynamic equilibrium exists, with the electrochemical potential (Fermi level) constant along the nanotube and equal to that of the substrate. (Semiconductor FE literature, e.g., Ref. 19, calls this the "zero current approximation.") To guarantee the constant Fermi level, the number of electrons in each subregion must be adjusted. This is achieved numerically by an iteration method. The resulting density (and induced density) of electrons increases along the CNT, with the highest densities at its apex, as physically expected.

The EEPE is denoted by U, and is the superposition of the following: (a) the PE due to the applied macroscopic field; (b) the PE due to the excess-charge distribution in the entire nanotube; and (c) the PE due to the related substrate surface- charge distribution, the effect of which is represented by the images in the substrate of the excess CNT charges.

For positions close to atoms, in particular at the CNT apex, it is not sufficiently accurate to treat the electron distribution in terms of point charges. To obtain the apex surface PE barrier, the density matrices of the closest seven subregions of the CNT were used. The induced charges in other subregions were treated as point charges since they are less critical. To check the convergence, we used the density matrices of the closest 10 and closest 30 subregions in two test calculations, and found that results were not significantly different from those found using the closest 7 subregions.

III. EFFECTS WHEN NO APPLIED FIELD PRESENT

A. Charge distributions

We first look at the simulated CNT charge distributions in the absence of any applied field. Excess charge densities (as defined in Sec. I) have been obtained by calculating the difference between (a) the CNT ground-state electron densities and (b) the electron densities associated with noninteracting carbon and hydrogen atoms at the same positions.

Results are illustrated in Figs. 1(a) and 1(b) showing the positions of the atomic nuclei in the CNT apex, as seen from above, for the closed and open CNTs. Figures 1(c) and 1(d) show the excess electron densities in a plane that includes the CNT axis and at least one of the atoms in the top layer. The edges of these planes are shown in Figs. 1(a) and 1(b). Neither CNT is symmetrical about its central axis, so the planes pass through a carbon atom at the top pentagon ring of the closed CNT only in the top half of Fig. 1(c), and through a hydrogen atom only in the top half of Fig. 1(d).

Excess electron charge is shown as a positive excess density. Red regions have gained electron charge and thus carry conventional negative charge; blue regions have lost electrons and carry conventional positive charge. For both closed and open CNTs, charge rearrangements have taken place that have the nature of dipoles (or higher multipoles).

With the H-terminated open CNT, dipole formation with the positive end outward could be expected on chemical grounds, from the Pauling electronegativities, and can be seen in Fig. 1(d). A dipole of this sense tends to reduce the surface electron barrier, as found in some experiments.[20] However, in addition, there seems to be a decaying charge-density oscillation that extends a small distance down the CNT wall. We believe that a decaying charge-density oscillation is to be expected physically, for much the same reason that Friedel oscillations exist. Friedel oscillations in CNTs were discussed by other authors, e.g., Ref. 21, although we are not aware of any previous suggestion that they would be caused by adsorbed hydrogen. However, our oscillations appear not to have the right wavelength for genuine Friedel oscillations. It is possible that the detail in our results is a consequence of using the MNDO model. Nevertheless, we do believe that adsorbed hydrogen could induce charge-density oscillations, and suggest that the effect should be looked for in more sophisticated CNT models.
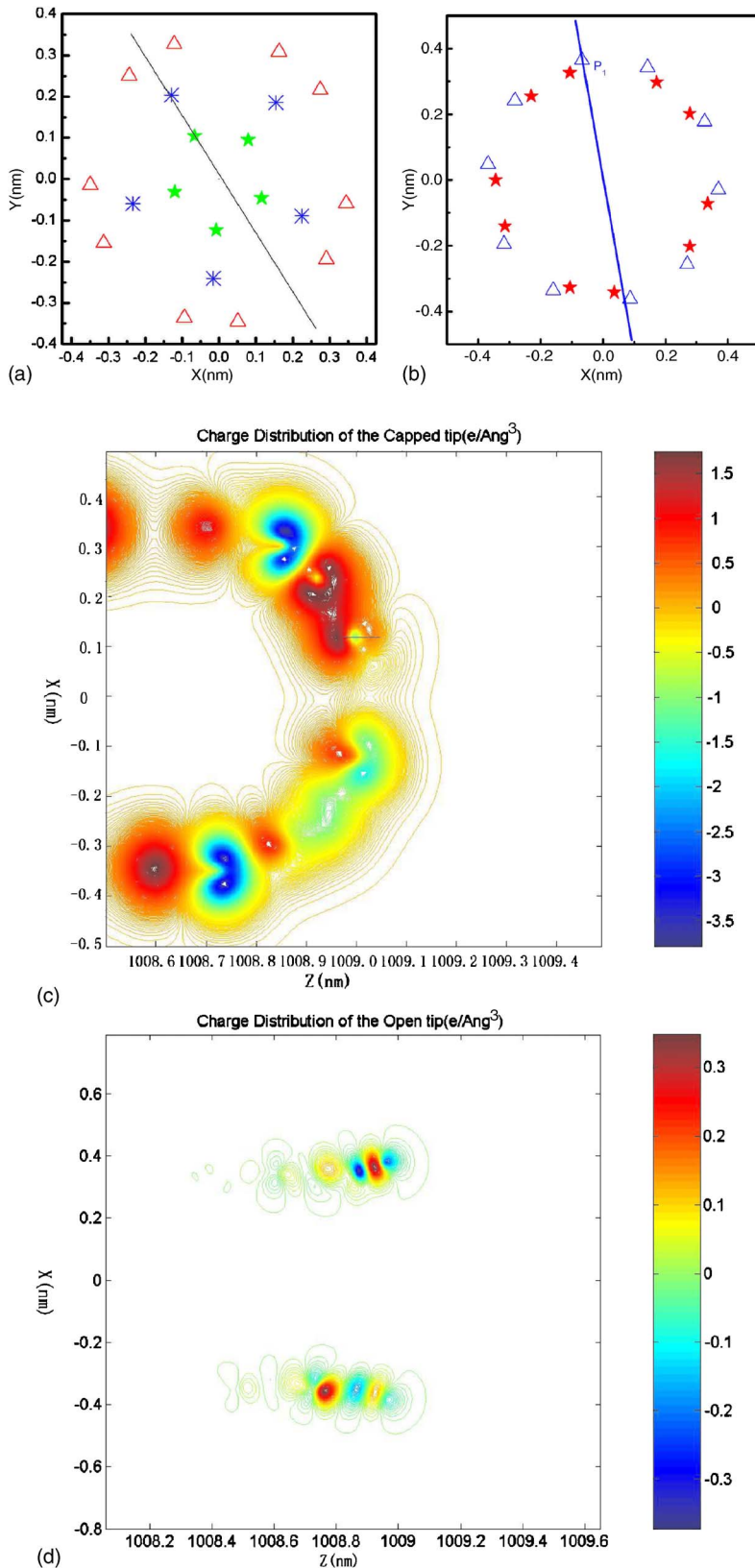
FIG. 1. (Color online) Excess electron densities, as defined in Sec. I, for closed and open CNTs in the absence of any applied field. (a) Projection onto a plane transverse to the axis, showing the positions of carbon atoms in the top layer (green ⋆), second layer (blue ∗) and third layer (red Δ), for closed CNT. (b) Projection onto a plane transverse to the axis, showing the positions of the hydrogen atoms (blue Δ) and the first ring of carbon atoms (red ⋆), for open CNT. [(c) and (d)] Excess electron densities on planes containing the CNT axis and passing through one of the first layer atoms [the edges of the planes are shown in (a) and (b) as straight lines]. Distances parallel to the CNT axis are measured from the substrate surface. Positive density indicates excess of electrons, so red regions have conventional negative charge while blue regions have conventional positive charge. Note the existence of charge re-arrangements having the general nature of dipoles.

Less expected was the charge rearrangement, with the closed CNT, that is seen in Fig. 1(c). [This effect can also be seen, but less clearly, in Fig. 5(b) in Ref. 2.] The effect is presumably due to a difference between the electronic structures of the CNT cylindrical wall and the CNT quasihemispherical cap, and/or electronic effects occurring at the join of these structures. There is both experimental and theoretical evidence for the existence of localized states at the end of closed CNTs.[22–25] Presumably, differences in local electronic structure are such that it is energetically favorable for electrons to move from the top of the cylindrical CNT wall into a localized state associated with the top pentagonal ring (or some other state associated with the cap), thus creating a dipole with its negative end outward and raising the apex
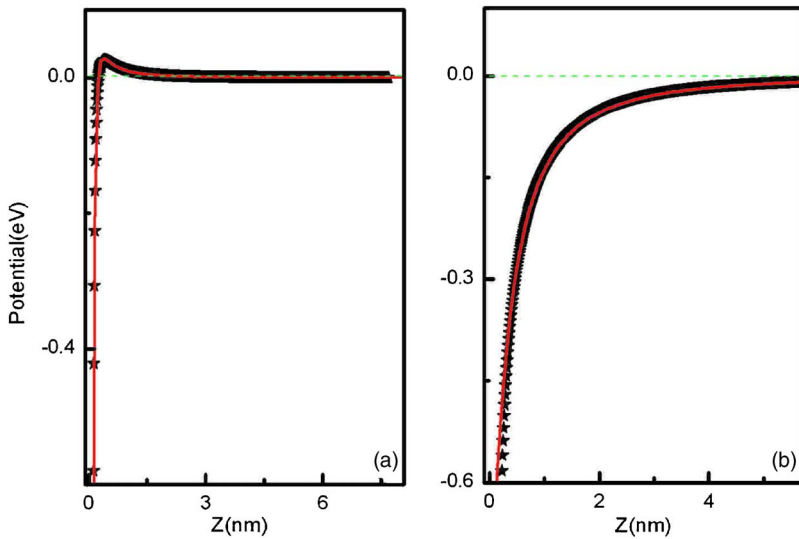
FIG. 2. (Color online) Total EEPEs in the absence of applied field: (a) above the nucleus of a pentagonal-ring atom in the closed CNT; (b) above a hydrogen-atom nucleus in the H-terminated open CNT. The starred points and the solid curve show the calculated values and a fitted curve, respectively.

band structure. This effect would resemble the upward band bending that occurs when electrons transfer from the surface region of a semiconductor to surface states in the band gap and below the bulk Fermi level.[19] However, there looks to be some evidence that the total effect is more complicated than this, and also involves charge redistribution stretching slightly down the CNT wall.

B. Intrinsic barriers

The motive energy near the CNT apex is important for FE. When no applied field is present, this PE confines the electrons inside the CNT and provides the so-called intrinsic tip barrier. We denote its electrostatic component (the EEPE) by U_0. The zero of U_0 is taken as the EEPE of an electron at a point in remote space, in the absence of any applied field; it can also be identified as the local vacuum level just outside the metal substrate.

For the closed CNT, for $F_M=0$, Fig. 2(a) shows how U_0 varies along a line, parallel to the central axis of the tube that passes through the nucleus of an atom in the topmost pentagonal ring of carbon atoms. Distance z is measured from the atomic nucleus and the atoms in the top pentagonal ring are centered at 1009.00 nm from the substrate. Similar PE variation was found along the axis itself, and along a line midway between atoms in the pentagonal ring.

For the open CNT, Fig. 2(b) shows how U_0 varies along the wall of the nanotube and forward into vacuum, along a line that runs parallel to the central axis of the tube, passing through one of the adsorbed hydrogen atoms. Here, distance z is measured from the hydrogen-atom nucleus, and the hydrogen atoms are centered at 1008.69 nm from the substrate.

In Fig. 2(a), U_0 for the closed CNT has a small peak at a value of z greater than zero. This form of EEPE variation is a characteristic of that produced by a dipole with its negative end outward. One possibility is that this is the result of local asymmetry in the electron distribution in the fullerene hemisphere. Neutral flat graphene has no dipole or quadrupole moment normal to the plane of the graphene sheet. The CNT cap is highly curved, so it is conceivable that a small spontaneous polarization might arise from the mixing of the p and

s atomic states, thereby giving rise to a dipole and/or quadrupole moment. However, it seems more likely that this small peak is a consequence of the dipolelike charge rearrangement seen in Fig. 1(c).

C. A simple model

We now present a simple model that brings out more clearly the differences between the closed and open CNTs seen in Fig. 2. Both diagrams show the deep well associated with the atomic core of the underlying atom [carbon in Fig. 2(a), hydrogen in Fig. 2(b)]. For simplicity, we model this part of the EEPE variation with a quadrupole PE term. We then add a dipole term to represent effects discussed above. We take both the dipole and the quadrupole as located at position $z=z_0$, on the line along which the barrier is considered. Thus we write

$$U_0(z) = -\frac{e}{4\pi\epsilon_0}\frac{P_0}{(z-z_0)^2} - \frac{e}{4\pi\epsilon_0}\frac{D_0}{(z-z_0)^3}, \quad (1)$$

where P_0 and D_0 are the strengths of the dipole and the quadrupole, respectively, in the absence of the applied field. The convention is used that dipole strength is positive when its positive end is outward.

As shown, the two plots in Figs. 2(a) and 2(b) are well fitted by the dipole-plus-quadrupole expression, and yield the following values: for the closed CNT: $z_0=-0.035$ nm, $P_0=-0.01$ e nm, and $D_0=0.0028$ e nm^2 and for the H-terminated open CNT: $z_0=-0.7$ nm, $P_0=0.27$ e nm, and $D_0=0.014$ e nm^2. The dipoles are of opposite polarities for the closed and open CNTs, as expected. This is the main result we wish to draw from this simple modeling; it confirms the impression given by the plots in Fig. 2. A similar conclusion is drawn from a more complicated fitting model that uses a better mathematical representation of the screened Coulomb potential associated with the atomic core.

Thus, from the full simulations and from this model, in the absence of any applied field, we conclude the following. For the closed CNT the apex dipole creates a field $F_{0,\text{closed}}$ in the same sense as the macroscopic field that will be applied;
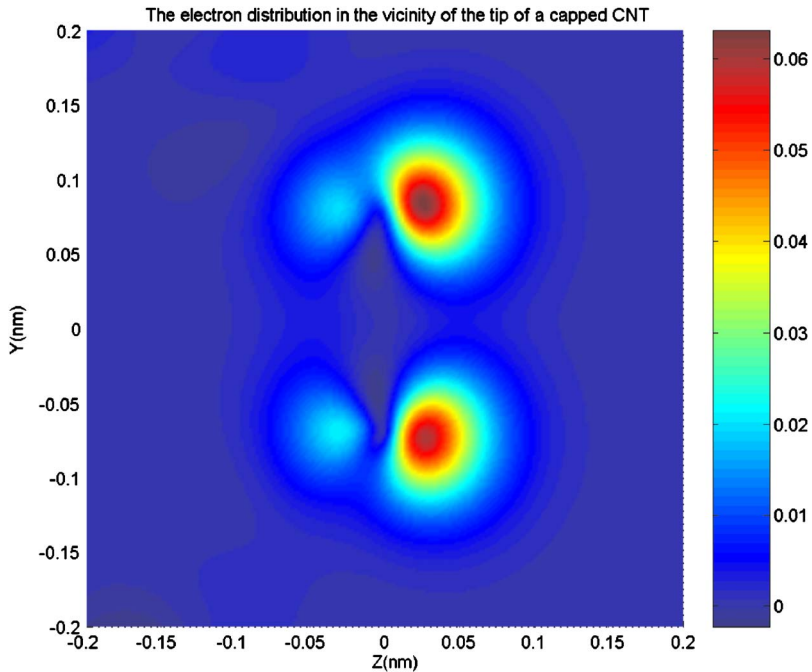
FIG. 3. (Color online) To illustrate the excess electron densities associated with two of the atoms at the top pentagonal ring of the closed CNT. Field-induced surface dipoles are prominently visible. The plane of the view passes through two of the carbon-atom nuclei and is parallel to (but does not include) the tip axis. The horizontal axis represents distance parallel to the CNT axis.

for the open CNT the apex dipole creates a field $F_{0,open}$ in the opposite sense. We return to this result in Sec. III.

IV. EFFECTS OF THE APPLIED FIELD

A. Charge distributions

With the present applied macroscopic field, the overall electron density distribution for the closed CNT is similar to Fig. 5 of Ref. 2. Local-field-induced dipoles (with their negative ends outwards) are clearly visible on the carbon atoms at the top pentagonal ring. The associated excess electron densities are shown in more detail in Fig. 3.

A field-induced electric-dipole layer always exists at the surface atoms of an electrically charged material (including metals). A simple model based on a parallel-plate capacitor shows why. The positive charge on one electrode acts on the electrons of the surface atoms in the opposing plate. These electrons move away from their local atomic nuclei until the forces on their electrical center due to these nuclei and to the distant charge are equal and opposite. It follows that a dipole layer must exist.[26] However, the small closed CNT seems particularly favorable for strong dipoles, as depolarization effects are smaller here than in many other contexts.

Surface dipoles with their negative ends outward were discussed in the context of FE by Drechsler[27] and Becker,[28] and were used by Mayer[7] in his analysis of open SWCNTs (without hydrogen termination), but we think ours was the first detailed quantum-mechanical calculation[2] to exhibit them clearly. The equivalent positive-end-outward dipoles have long been part of the theory of the field ion emission techniques,[26] and appear in advanced quantum-mechanical simulations[29] of positively charged metal surfaces. The existence of these surface electric dipoles has significant implications in the physics of charged surfaces.[26]

B. Tunneling barriers

The overall EEPE distribution associated with the closed CNT is essentially similar to Fig. 3 in Ref. 9. The corresponding distribution associated with the open CNT is shown in Fig. 4.

Figure 5 compares the tip barriers for the closed and open CNTs in an applied macroscopic field of 11 V/μm. The upper curve is for the closed CNT, the lower one is for the open CNT. For the closed CNT, the horizontal axis has the same meaning as in Fig. 1(a), for the open-ended structure, it is the same as in Fig. 1(b). The dashed line shows the Fermi level; for reasons of theoretical self-consistency associated with the code used in the calculations, the work function of the SWCNT cylindrical wall has been taken as 5.08 eV.

In making these comparisons we are assuming that, for the open CNT, the emission comes out of a covalent state associated with the C–H bonding, and that the transport characteristics of the CNTs are good enough to provide plenty of electrons into this state. In these circumstances, the barrier between the H atom and the vacuum is the main constraint on the electron emission.

Because the dipole field of the closed CNT is in the same direction as the applied macroscopic field, whereas the dipole field of the open CNT is in the opposite direction, the total calculated field for positive z values will be greater for the closed CNT than for the open one. On the other hand, the dipole associated with the closed CNT will raise the apex surface potential barrier, whereas the C–H dipole of the open CNT will lower this barrier. Therefore, in applied fields, the tip barrier of the closed SWCNT is higher and sharper, while that for the open one is lower but smoother.

The barrier maximum is lowered by the applied field more for the open CNT than for the closed one. This is
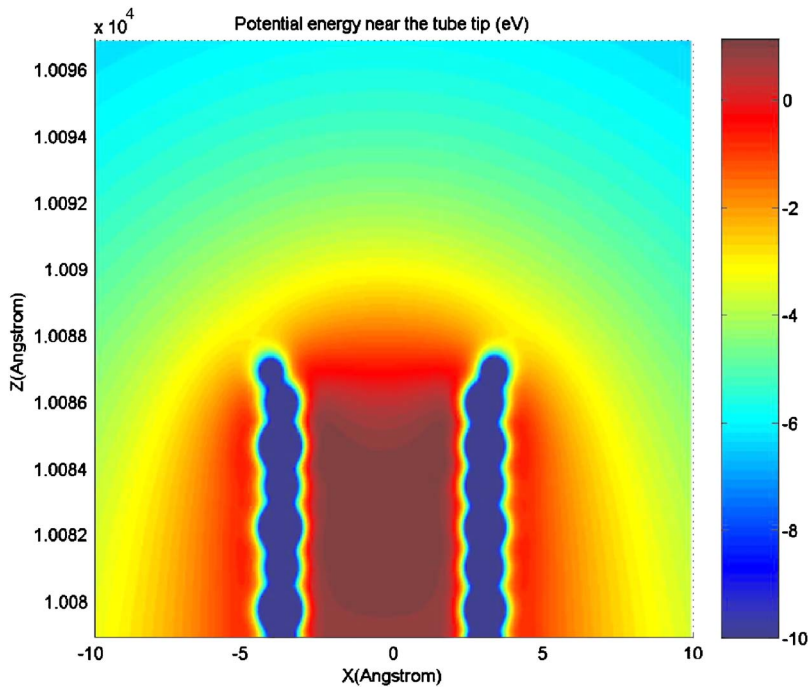
FIG. 4. (Color online) The EEPE (in eV) in an intersection plane, parallel to the tube axis, through one of the terminated hydrogen atoms of the open CNT, in an applied macroscopic field of 10.0 V/μm.

expected if the induced electron charge accumulates partly on the last carbon atom, rather than totally on the hydrogen atom.

We note that the effects broadly similar to those reported here might be expected to occur with any adsorbate onto a CNT where the chemistry of the situation is such that (1) a chemically induced transfer of electron-type charge takes place inward toward the carbon substrate, thereby creating an electric dipole with its positive end outward and (2) a covalent-type state is formed that spans both the adsorbate and some parts of the carbon substrate, are at or closely below the CNT Fermi level, and are easily populated by charge transport from the main part of the CNT.
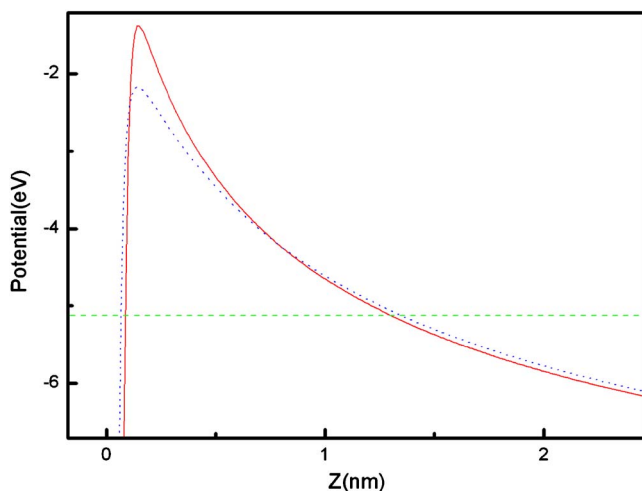

FIG. 5. (Color online) Variation of electrostatic electron PE for closed (red, solid curve) and open (blue, dotted curve) CNTs, along the lines specified in the text, in an applied macroscopic field of 11.0 V/μm. The dashed green line represents the Fermi level, which is at −5.08 eV relative to the zero-field vacuum level.

V. FIELD ENHANCEMENT FACTORS

A. Introduction and numerical results

In the literature of field emission from planar films, the concept of a FEF plays an important role. This is because it is often assumed that a FN-type equation can be used to calculate/analyze the emission current, provided that the barrier field F in the FN-type equation is replaced by

$$F = \gamma F_M, \tag{2}$$

where F_M is the macroscopic field as defined above and γ is defined as the FEF.

When this approach is applied to CNTs, as it often is, it is being implicitly assumed that the CNT is behaving as a classical conductor, that the barrier field F is well defined, and that the FEF value depends only on geometry and is independent of macroscopic field. The results to be reported here show, certainly for CNTs with small apex radii, that none of these assumptions is justified, that the concept of "FEF" is very problematic, and that the issues involved deserve deeper investigation.

In the circumstances of these calculations it is difficult to define a quantity that corresponds well to the enhancement factor as defined in classical-conductor calculations. Part of the reason is that the EEPE, as defined here, "sees" the screened Coulomb potential of the atomic core, so if we want to define the FEF in terms of the local field at a point in space, then this point has to be chosen an adequate distance away from the atomic core. However this inevitably means that this point has to be chosen some distance away from where we might place the equivalent surface of a classical conductor. In making comparisons, one then has to recognize that a CNT apex is so small that the field for a classical conductor falls off significantly with distance.

There is no easy way out of this dilemma. In Mayer's latest monopole-dipole treatment,[8] point charges are modeled
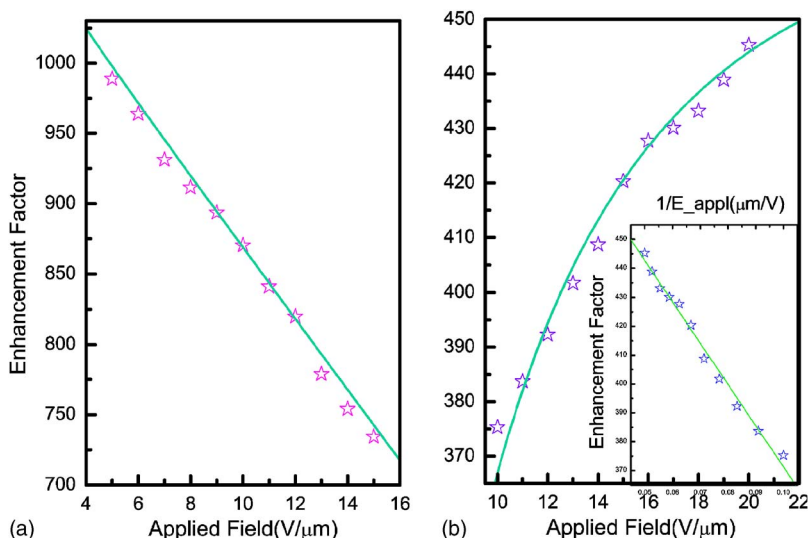
FIG. 6. (Color online) To show how the FEF γ_s varies with applied macroscopic field F_M for (a) the closed CNT and (b) the H-terminated open CNT. The solid lines are fitted curves. The inset in (b) shows the linear relationship between γ_s and $1/F_M$.

by Gaussian distributions. For his definition of FEF, Mayer uses a calculated local field at the position of a relevant atomic nucleus. This leads to a well-defined FEF. However we can deduce, from older work[30] on depolarization effects in planar monopole-dipole arrays, that a local field of this type is expected to be significantly lower than the local external field in space above the array. So the derived FEF would be correspondingly lower. In the hemisphere-on-a-post formula[11] "FEF$=\alpha\times$(height/apex-radius)," Mayer finds proportionality-coefficient (α) values much lower than those found by researchers who used a local field defined at a position (near the outer edge of the electron charge cloud[17]) that corresponds to the classical electrical surface of the CNT (see discussion after Fig. 8 in Ref. [8]). The difference in definition of "local field" may be part of the reason why.

The rapid field falloff above a CNT apex of very small radius complicates matters further. Neither Mayer's FEF nor the classically defined FEF corresponds well to the FEF that would be derived from conventional FN plot analysis because neither uses a definition of local field that corresponds to the field in the region of space where the tunneling barrier is actually located.

What we do here is to adopt a new, slightly arbitrary, definition of FEF (denoted here by γ_s) and explore how the resulting FEF value changes with macroscopic field. On the relevant line as specified earlier, we choose a fixed z value, z_s, in space outside the peak in the barrier but close to the position where the field has its maximum absolute value (in the range of F_M of interest). We then evaluate the field F_s at position z_s and define a FEF γ_s as F_s/F_M. In practice we take $z_s=0.20$ nm for the closed CNT and $z_s=0.24$ nm for the open CNT. The philosophy of this approach is that we are trying to choose a point such that the field looks as if it would be roughly the right value for an "equivalent triangular barrier."

The variation of this γ_s with applied macroscopic field F_M is shown in Figs. 6(a) and 6(b). The values of γ_s are significantly lower than simple classical-conductor calculations would suggest, and in neither case is the enhancement factor constant. For the closed CNT, γ_s decreases linearly

with F_M; for the open CNT, γ_s increases with F_M, but the form of the variation is better described as a linear decrease with $1/F_M$.

B. Theoretical discussion

For the closed CNT, the simple classical-conductor calculations noted earlier took the emitter radius as defined by the carbon-atom nuclei positions. In reality, the constant-potential surface assumed in these calculations needs to be taken at the outer edge of the electron charge cloud.[17] This makes the CNT electrical radius about 0.415 nm and decreases the classical prediction of γ_s to about 2410 for the simple formula and about 1330 for Eq. (20) in Ref. [10]. Further, the position z_s is outside this new classical-conductor surface by about 0.13 nm. If field falloff outside this surface is taken into account, then an even lower classical γ_s value (perhaps as low as 750) would be calculated. This is similar to the values predicted by the MNDO calculations. So it is unsurprising that Fig. 6 values of γ_s are significantly less than those derived from naive classical calculations. These lower values are also broadly consistent with the finding, by Bonard,[31] that the simple "height/radius" formula often over-predicted experimental γ_s values by a factor of roughly two.

The γ_s values calculated for a classical conductor would be constant, but ours are not. In Mayer's theory,[8] also, the FEF does not depend on the applied macroscopic field F_M. We think this is because his theory is primarily electrostatic in nature and does not deal with chemically induced charge transfers in the same way that ours does. Our approach is more fundamental and should be better at uncovering chemically induced effects.

In fact, the FEF as conventionally defined for an old-style metal field emitter is not strictly independent of field, either. Real field emitters, that expose facets with different values of local work function, are surrounded by a system of so-called patch fields.[12,32] Consider such an emitter mounted on one face of a parallel-plate arrangement. In this case, a simple theory of field enhancement exists. At a given part of the emitter surface, let the local field component induced by the applied voltage be F_V and the component of patch field

parallel to F_V be F_P, and suppose that these are independent. Let the total local field parallel to F_V be F_B and the macroscopic field be F_M. Then the total FEF γ_B is given by

$$\gamma_B = (F_B/F_M) = (F_V + F_P)/F_M = \gamma_{class} + F_P/F_M, \qquad (3)$$

where γ_{class} is the FEF you would get from an appropriate calculation involving a classical conductor of uniform work function.

Since F_P can be positive or negative, what one expects is that γ_B varies with $1/F_M$, either positively or negatively. This is how the H-terminated open CNT behaves in our calculations. For the open CNT the dipole field $F_{0,open}$ discussed earlier is negative: thus a linear decrease in γ_s as $1/F_M$ increases is expected (and observed).

However the closed CNT is behaving unexpectedly. Assume Eq. (2) is applicable to a CNT, with F_s taking the place of F_B, and γ_s that of γ_B. For γ_s to decrease linearly as F_M increases, there needs to be a source of local field F_P that itself has a strength that varies as $[-F_M^2]$, at least over the range of fields in question. We have examined several alternatives in search of a plausible physical explanation. So far, FP seems the only effect likely to generate this form of dependence.

Physically, the FP argument for field reduction at point z_s outside the closed CNT is as follows. If the nanotube was truly a metallic conductor, then the applied macroscopic field would generate an induced charge that is strongly localized near its apex. However, if the density of states of the carbon sheet is not high enough to allow sufficient charge to accumulate near the apex, then this induced charge will be spread over both the fullerene-type SWCNT cap and the near part of the cylindrical SWCNT column, and FP occurs. Due to the postlike geometry, and the small size of the SWCNT under investigation, movement of induced charge away from the CNT apex will clearly have the effect of reducing the local field acting on the top atoms in the pentagonal ring, and this in turn would reduce the strength of the local field just outside the CNT apex.

Figure 4 of Ref. 2 provides numerical evidence that FP occurs in CNT simulations. Alternatively, a basic theoretical argument is as follows. If S denotes the density of states per unit area, near the charge-neutrality level of a graphene or fullerene sheet, then the electron-type charge per unit area σ associated with downward band bending by an amount $|\Delta|$ is $\sigma = Se|\Delta|$. (The FE sign convention takes σ as positive.) From Gauss's theorem the local field close above this surface-charge density is $F = \sigma/\epsilon_0$. A value $S = 1$ eV^{-1} nm^{-2} may be derived from a formula given by Saito *et al.*[33] Using this and the value $F = 10$ V/nm yields $|\Delta| = 0.5$ eV. This gives some idea of the possible magnitude of such shifts. The effect is clearly sufficiently large that it needs to be taken into consideration. This is also the view of Chen *et al.*[34]

It is already known experimentally that FP can reduce the field above an emitter apex. From field ion microscope investigations long ago,[35] on a uranium dioxide emitter (which undergoes a metal to semiconductor type transition as temperature is reduced from about 110 to about 80 K), it is known that FP can cause reduction in the apex field by a factor of order 3. (The actual observation is a threefold in-

crease in the so-called best image voltage as temperature is lowered: this implies that near 80 K you need an applied voltage three times higher than that near 110 K, to get the same apex field as before.) Apex-field reductions in the size shown in Fig. 6(a) are thus not physically surprising.

FP is known to involve mathematical nonlinearity.[36] So it seems not implausible that, in the lowest useful approximation, the effective field F_a acting on a pentagonal-ring atom should go approximately as

$$F_a = c_1 F_M - c_2 F_M^2, \qquad (4)$$

where c_1 is a constant and c_2 is a slowly varying (with F_M). The field F_s (at position z_s, 0.20 nm away from a pentagonal-ring atom) should have the same dependence on F_M. In the expression for the FEF factor γ_s, this would generate a component proportional to $[-F_M]$, as observed in the MNDO calculations.

A variation in γ_s that goes as F_M would also occur if there were contributions to the effective polarizability of a pentagonal-ring atom from a first-hyperpolarizability term that gave a contribution to the dipole strength of the form $(1/2)\beta F_a^2$. The first-hyperpolarizability β would be zero for a flat graphene sheet (or a spherically symmetric atom), but we think it possible that it could be nonzero for a carbon atom located in the sharply curved apex of a CNT. We find it difficult, however, to convince ourselves that this term would be negative rather than positive and of any significant size. If it exists, we think its effect would be to reduce the consequences of FP slightly.

It might be asked why FP effects are, apparently, not found with the open CNT. We think the answer is that, in this case, the chemically induced C–H dipole is closer to the point at which F_s is defined and plays the dominant role in determining how F_s varies with F_M. However there is a need for more detailed investigation.

In summary, neither classical-conductor theory nor patch-field type effects can predict a linear decrease of the FEF with F_M. If, however, part of the reason for the low γ_s values is FP, then this dependence can be explained, at least in principle.

An increase in enhancement factor with applied macroscopic field was reported by Buldum and Lu,[37] in their classical-conductor-type calculations made as part of a simulation of emission from closed (5, 5) SWCNTs very much shorter than ours (their lengths were 3.8–9 nm). This variation appears to be the result of the way that they have defined their local field, E_{loc}, as the field at the point at which the electron leaves the barrier. This point moves inward, toward the classical conductor surface, as the applied macroscopic field increases, so an increase in FEF would be expected. This effect is different from those considered above. We have deliberately avoided the need to consider it, by choosing to define our local field F_s at a fixed point in space, relative to the relevant atomic nucleus.

More generally, there seems a strong need for more detailed investigation of the theory of FEFs for CNTs. This needs to embrace both more sophisticated quantum-
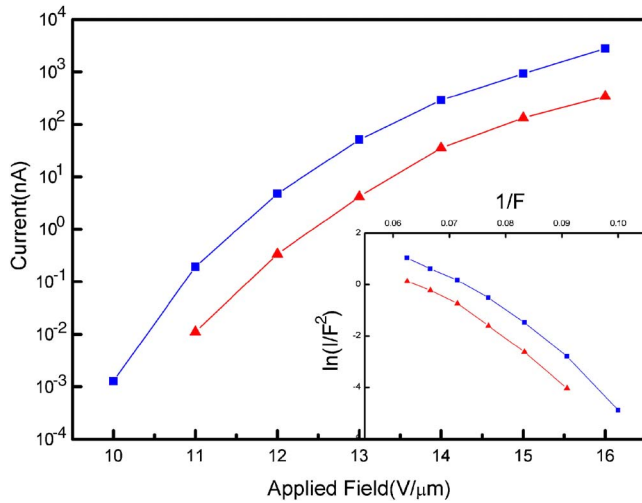
FIG. 7. (Color online) Emission current as a function of applied macroscopic field. The triangular points show the current from the closed CNT while the square points show the current from the open CNT. The inset shows the relationship in the form of FN plots.

mechanical modeling and deeper analysis of precisely what physical parameter the FEF derived from a FN plot corresponds to.

VI. EMISSION CURRENT

At present, it is far from clear what the correct detailed procedure should be for calculating the emission-current I from a CNT. Obviously, the density of states relevant to the CNT emitting region differs significantly from a free-electron-metal density of states, so it is not strictly appropriate to use conventional FN-type formulae derived using free-electron theory. Thus, we use here the simple approximate formula used in Ref. 2, namely,

$$I = \nu q_{ind} D. \tag{5}$$

Here, D is the transmission coefficient for a "typical barrier" at the CNT tip, q_{ind} is the total magnitude of the induced electron charge in the emitting region at the tip (as determined from our calculations), and ν is an assumed frequency of electron collision with the barrier. Values of ν are estimated by using the uncertainty principle.[2] D is obtained by carrying out a numerical integration on the relevant barrier, as described above, using the usual simple-JWKB (Jeffreys–Wentzel–Kramers–Brillouin) formula.

The values used for q_{ind} and ν vary slowly with applied macroscopic field F_M. Values for $F_M = 12$ V/μm are typical and are as follows. For the closed CNT $q_{ind} = 0.16e$, $\nu = 1.2 \times 10^{15}$ Hz; for the open CNT $q_{ind} = 0.45e$, $\nu = 1.7 \times 10^{15}$ Hz. This leads to the "typical ratio"

$$(\nu q_{ind})_{open}/(\nu q_{ind})_{closed} = 4. \tag{6}$$

Figure 7 shows semilogarithmic plots of current versus applied field, with the corresponding FN plots shown as an inset. The emission current from the open CNT is an order of magnitude greater than that from the closed one, for the same applied macroscopic field. This is partly due to differences in D and partly due to the differences in the pre-exponential (with the calculated difference in q_{ind} much more significant

than the difference in ν). The overall result goes in the same qualitative direction as the predictions of Mayer *et al.*[38] for a very much shorter SWCNT.

As noted earlier, C&E effects are neglected in this treatment. If the behavior of conventional FN-type equations is any guide,[39] these plots will underpredict true currents by a factor of around 100. Thus, the absolute emission-current values shown should be treated with great caution.

A feature of note is that the predicted FN plots for both closed and open CNTs are curved downward, rather than linear. The FN plots calculated by Buldum and Lu[37] for their short closed (5, 5) CNTs show much the same kind of curvature. Curvature has also been found in FN plots derived from model calculations on classical spherical emitters, both in Ref. 15 and in unpublished work by one of us (Forbes[40]). Thus, we think that the smooth curvature seen in our FN plots and in some other FN plots could be associated, at least partly, with rapid field falloff with distance. This makes the barrier shape different from that assumed in theory applicable to planar emitters, and leads to FN-plot nonlinearity.[15,41]

Experimental FN plots from CNTs do often exhibit convex (i.e., downward) curvature. However, there are further possible explanations, such as "saturation of the electron supply" (i.e., progressive breakdown of the assumption that the electron distribution inside the emitter, at its apex, is in thermodynamic equilibrium with electron pool in the substrate). This type of saturation effect is not included in the present model, so is not the cause of the curvature shown in Fig. 7.

VII. DISCUSSION

A. General issues

These calculations, together with previous works,[1,2] identified various physical effects that may influence FE from SWCNTs. They suggest that the physics of a charged nanotube may be significantly more complex than hitherto realized. In particular, we have established that FP should occur, and have identified four effects that might give rise to apex dipoles, namely, for the closed CNT (1) mixing of orbitals at the carbon atoms at the highly curved fullerene cap; (2) transfer of electron charge between the near part of the CNT column and the cap, and (3) field-induced polarization of the apex carbon atoms; and for the open CNT (4) electron transfer associated with hydrogen adsorption. It has also been suggested, for the open CNT, that hydrogen termination may lead to charge-density oscillations.

The existence of FP might help to explain the electroluminescence reported by Bonard *et al.*[42] FP would allow both the states in their two-state mechanism to be "further up the band structure" (and hence closer to the vacuum level) than they assumed. In consequence, the lower state need not be quite such a deep level, and this would make it easier for a hole to be created in such a state by tunneling.

The CNT end structure has significant influence on the tip barrier. For the open CNT, the apex dipole resulting from hydrogen termination plays a dominant role, in that it reduces the height of the tunneling barrier. In zero applied field

the EEPE variation is relatively sharp for the closed CNT, but smoother and longer range for the open CNT, due to this large C–H dipole.

However, Fig. 2 shows that in both cases, as z gets larger, the EEPE climbs back to the level of the reference zero within about 5 nm of the last CNT carbon atom. This behavior is markedly different from that associated with hydrogen adsorption on a flat metal plane (assuming the rest of the emitter is clean): in the metal-plane case there is a PE plateau just outside the surface, and the EEPE climbs back to the reference level over a distance comparable with the diameter of the atom. The absence of a PE plateau in the case of a H-terminated CNT makes it difficult to create a meaningful definition of local work function that can be used in the presence of an applied field, and this in turn makes it difficult to estimate tunneling probability by thinking of the hydrogen adsorption as leading to a reduction in work function and then applying some simple formula.

Nevertheless, barrier reduction due to the C–H dipole does produce an effect. Figure 5 confirms that, with the applied field present, the barrier height is lower for the open CNT. For a given applied field, the emission current is higher for the open CNT than for the closed CNT, even though the FEF (as calculated here) for the open CNT is significantly smaller (as is evident from slopes of the curves in Fig. 5, near $z=0.2$ nm).

On the issue of the role of hydrogen adsorption, we have found that the H-terminated open CNT should have slightly the better field emission performance. This finding does not necessarily contradict the previous *ab initio* calculations.[3,4] This is because, in our picture, FP means that electrons are emitted from energy states above the charge-neutrality level; so the density of states in the immediate vicinity of the charge-neutrality level should be of diminished relevance.

However, the performance difference between open (H-terminated) and closed CNTs is small. So, in the choice between SWCNT forms (assuming both could be manufactured reliably), other technological factors such as system stability would probably be of more significance.

B. The correlation-and-exchange problem

As already noted, a possible deficiency is the neglect of C&E effects associated with the departing electron. We thought it useful to first look at the electrostatic effects by themselves. We also wished to draw wider attention to apparent difficulties in deciding whether, and how, to take C&E effects into account.

For FE from metals, the C&E component in the motive energy is the metal response to a departing *point* electron: in the first approximation this is the well-known image PE, but there are more detailed models involving interaction with a C&E hole,[43] or involving some form of density functional theory. On the other hand, in the calculation of electronic ground states, every electron is represented by a delocalized wave function, and the final result is a self-consistent solution in which exchange effects have been taken into account. A central question is as follows. In determining the C&E response of the CNT to the presence of an electron in the

process of emission (which we need to do, in order to create the motive-energy expression that goes into the one-electron Schrodinger equation), what shall we assume about the spatial distribution of the electron being emitted?

The conventional approach is to treat the departing electron as a point for the purposes of determining the system response, but treat it as delocalized and wavelike for the purpose of determining tunneling probability). It would be helpful to have access to a clear explicit proof that this is the correct fundamental way to apply quantum mechanics. We offer this issue for wider discussion.

C. Use of the MNDO method

Exact quantum-mechanical analysis of the surfaces of charged systems is notoriously difficult, especially in situations of limited symmetry. This is because (as the surface dipoles discussed earlier show) the charge is associated with a strong modification of the system wave functions in the vicinity of the topmost layer of atoms (as compared with the corresponding uncharged surface). The MNDO method used here is a long standing but relatively unsophisticated method, of limited numerical accuracy. Its advantage, in the analysis of charged surfaces, is that it is likely to bring out any special local atomic-level phenomena that may physically exist. It is not surprising that our results exhibit several such phenomena. At qualitative and semiquantitative levels, most of these are clearly consistent with experimental evidence and/or more general theoretical arguments. The main value of the present work is that it suggests and provides support to hypotheses that these effects do exist physically. It also provides some provisional numerical estimates. These effects should be looked for in more sophisticated treatments capable of greater numerical accuracy.

VIII. CONCLUSIONS

In summary, this work (and our previous papers) confirmed that, when calculating surface PE barriers and FEFs, it is not satisfactory to treat SWCNTs as if they were solid metal objects. Essential differences are that (1) the density of states near and above the charge-neutrality level is much lower for a SWCNT than it is for a metal, so FP and band-bending effects may occur, especially near the emitter apex; (2) various local-dipole effects may occur near/at the emitter apex; and (3) it is difficult to formulate a definition of FEF that is both useful empirically and can also be calculated from first principles. Another difference, not brought out in this paper, is that electron energy is effectively quantized in the direction normal to the fullerene-type carbon sheet that forms the emitting cap; this will affect observed energy distributions.[42]

The hydrogen-terminated open SWCNT behaves differently from the closed CNT, the main differences being the variation of FEF with applied macroscopic field (increases with F_M rather than decreases), the apparent presence of hydrogen-induced charge-density oscillations, and the slightly better emission performance.

In some respects, FE from a closed CNT (even from a so-called metallic CNT) is more analogous to FE from a

semiconductor than to FE from a metal, with the best simple analogy perhaps being to a situation where the conduction band is significantly degenerate and the resulting triangular well has quantized levels in it. However CNT really needs to be treated as a field emission situation in its own right. There is, of course, no reason to expect that its behavior will be well described by the standard Fowler-Nordheim-type equation or by the simpler FN-type equations sometimes used in literature.

There are three important general messages from this work. First, the physics of charged nanotubes seems, even qualitatively, to be much more complex than has hitherto been realized. Second, the concept of FEF can be very problematic in the context of CNTs. Third, several effects uncovered here have equivalents in conventional FE from larger metal and/or semiconductor emitters: so studies of this kind may help unify knowledge concerning FE emission.

ACKNOWLEDGMENTS

[1] X. Zheng, G. H. Chen, Z. B. Li, S. Deng, and N. Xu, Phys. Rev. Lett. **92**, 106803 (2004).

[2] J. Peng, Z. B. Li, C. S. He, X. Z. Deng, N. S. Xu, X. Zheng, and G. H. Chen, Phys. Rev. B **72**, 235106 (2005).

[3] G. Zhou, W. H. Duan, and B. G. Gu, Phys. Rev. Lett. **87**, 095504 (2001).

[4] C. Kim, Y. S. Choi, S. M. Lee, J. T. Park, B. Kim, and Y. H. Lee, J. Am. Chem. Soc. **124**, 9906 (2002).

[5] C. Kim, K. Y. Seo, B. Kim, N. Park, Y. S. Choi, K. A. Park, and Y. H. Lee, Phys. Rev. B **68**, 115403 (2003).

[6] A. Mayer, N. M. Miskovsky, P. H. Cutler, and Ph. Lambin, Phys. Rev. B **68**, 235401 (2003).

[7] A. Mayer, Phys. Rev. B **71**, 235333 (2005).

[8] A. Mayer, Phys. Rev. B **75**, 045407 (2007).

[9] C. Y. Zhi, X. D. Bai, and E. G. Wang, Appl. Phys. Lett. **81**, 1690 (2002).

[10] A. Wadhawan, R. E. Stallcup II, K. F. Stephens II, J. M. Perez, and I. A. Akwani, Appl. Phys. Lett. **79**, 1867 (2001).

[11] R. G. Forbes, C. J. Edgcombe, and U. Valdrè, Ultramicroscopy **95**, 57 (2003).

[12] C. Herring and M. H. Nichols, Rev. Mod. Phys. **21**, 185 (1949).

[13] E. L. Murphy and R. H. Good, Phys. Rev. **102**, 1464 (1956).

[14] R. H. Fowler and L. W. Nordheim, Proc. R. Soc. London, Ser. A **119**, 173 (1928).

[15] C. J. Edgcombe and A. M. Johansen, J. Vac. Sci. Technol. B **21**, 1519 (2003).

[16] R. G. Forbes, Surf. Interface Anal. **36**, 395 (2004).

[17] R. G. Forbes, Ultramicroscopy **79**, 25 (1999).

[18] M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. **99**, 4899 (1977).

[19] A. Modinos, *Field, Thermionic, and Secondary Electron Emission Spectroscopy* (Plenum, New York, 1984).

[20] P. Ruffieux, O. Groening, M. Bielmann, P. Mauron, L. Schlapbach, and P. Groening, Phys. Rev. B **66**, 245416 (2002).

[21] M. F. Lin and D. S. Chuu, Phys. Rev. B **56**, 4996 (1997).

[22] R. Tamura and M. Tsukada, Phys. Rev. B **52**, 6015 (1995).

[23] D. L. Carroll, P. Redlich, P. M. Ajayan, J. C. Charlier, X. Blase, A. De Vita, and R. Car, Phys. Rev. Lett. **78**, 2811 (1997).

[24] P. Kim, T. W. Odom, J.-L. Huang, and C. M. Lieber, Phys. Rev. Lett. **82**, 1225 (1999).

[25] S. Han and J. Ihm, Phys. Rev. B **61**, 9986 (2000).

[26] R. G. Forbes, Ultramicroscopy **95**, 1 (2003).

[27] M. Drechsler, Z. Elektrochem. **61**, 48 (1957).

[28] J. A. Becker, in *Solid State Physics: Advances in Research and Applications*, edited by F. Seitz and D. Turnbull (Academic, New York, 1958), Vol. 7.

[29] G. C. Aers and J. E. Inglesfield, Surf. Sci. **217**, 367 (1989).

[30] R. G. Forbes and M. K. Wafi, Surf. Sci. **93**, 192 (1980).

[31] J.-M. Bonard (private communication).

[32] L. K. Hansen, J. Appl. Phys. **37**, 4498 (1966).

[33] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College, London, 1998).

[34] C. W. Chen, M. H. Lee, and S. J. Clark, Appl. Surf. Sci. **228**, 143 (2004).

[35] R. Morgan, J. Mater. Sci. **5**, 445 (1970).

[36] W. Moench, *Semiconductor Surfaces and Interfaces* (Springer, Berlin, 1995).

[37] A. Buldum and J. P. Lu, Phys. Rev. Lett. **91**, 236801 (2003).

[38] A. Mayer, N. M. Miskovsky, and P. H. Cutler, Appl. Phys. Lett. **79**, 3338 (2001).

[39] R. G. Forbes and J. H. B. Deane, Proc. R. Soc. London **463**, 2907 (2007).

[40] R. G. Forbes (unpublished).

[41] C. J. Edgcombe and N. de Jonge, J. Vac. Sci. Technol. B **24**, 869 (2006).

[42] J.-M. Bonard, J.-P. Salvetat, T. Stoeckli, L. Forro, and A. Chatelain, Appl. Phys. A: Mater. Sci. Process. **A69**, 245 (1999).

[43] T. L. Loucks and P. H. Cutler, J. Phys. Chem. Solids **25**, 105 (1964).