

# Fluctuation-induced tunneling conduction through nanoconstrictions

Hang Xie and Ping Sheng\*

*Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

(Received 18 November 2008; revised manuscript received 15 March 2009; published 14 April 2009)

We propose a generalization of the well-known fluctuation-induced tunneling conduction mechanism based on electron tunneling through finite segment(s) of nanoconstrictions whose transverse dimension is less than half the Fermi wavelength. By considering the effects of thermally induced voltage fluctuations across the nanoconstrictions, a temperature-dependent conductivity behavior is obtained which is in good agreement with the experiments, with reasonable parameter values. In the limit of high applied voltage the present model predicts interesting electronic Fabry-Perot behavior manifesting as peaks in differential conductance with linear variation in their voltage separations.

DOI: 10.1103/PhysRevB.79.165419

PACS number(s): 73.23.-b

## I. INTRODUCTION

In disordered materials characterized by large metallic regions separated by small insulating gaps, it has been widely observed that the conductance-temperature behavior follows a typical variation that can be well described by the fluctuation-induced tunneling conduction (FITC) mechanism.<sup>1</sup> In contrast to the usual expectation that tunneling conduction should exhibit small or no temperature variation, it was shown that, at finite temperatures, thermally induced fluctuation voltages across the insulating gap (which may be regarded as a small capacitor) can influence the tunneling probability and thereby impart a characteristic temperature dependence to the conductivity. The FITC can successfully explain the nonmetallic temperature dependence of conductivity in diverse materials such as carbon-polyvinylchloride films, tin-doped indium oxide (ITO) thin films, carbon nanotube (CNT) bundles, and conducting polymers.<sup>2-5</sup> These materials share the common trait that, other than the nonmetallic temperature variation in the conductivity, other properties, e.g., optical characteristics, all exhibit metallic behavior.

In spite of the success of the FITC, however, there are some cases in which the parameters extracted by applying the model are difficult to reconcile with known material properties. For example, the barrier height can be unphysically small (several meV) and the barrier width unreasonably large (50–100 nm). This is in contrast to the usual barrier height on the order of 1 eV and barrier width on the order of 1 nm. The existence of such cases imply that, in addition to the conventional picture of insulating potential barrier for the gap separating the metallic regions, there could be alternative physical scenario(s) that display similar characteristics.

In this paper we propose a model in which the insulating potential barrier is replaced by very narrow constriction(s), consisting of conducting elements (such as a chain of single metallic atoms) whose transverse width is less than half the electronic Fermi wavelength. As electrons behave like waves in the nanoscale, the narrow constriction acts like a waveguide with the transverse dimension below the cutoff. As the constriction is finite in length, tunneling transport can occur.

In what follows, we recapitulate the FITC mechanism in Sec. II, followed by the presentation of the model in Sec. III.

Outline of the mesoscopic calculations is given in Sec. IV. Results and discussion, together with comparisons to the experiments, are given in Sec. V. Good agreement is shown with the observed two-probe resistance measurements on metallic RuO<sub>2</sub> and IrO<sub>2</sub> nanowires.<sup>6</sup> We choose to compare with these experimental results because in these particular cases both four-probe and two-probe measurements were carried out, with opposite temperature variations (two-probe results showed nonmetallic behavior). Hence it is clear that the two-probe results are dominated by the contacts. Moreover, the parameters extracted by applying the FITC show unphysically small potential barrier height and large barrier width, exactly the deficiency we wish to correct with the model in Sec. III. This is in contrast to the results observed recently in CNT samples,<sup>5</sup> where the agreement between the two-probe and four-probe results indicates the measured characteristics to be intrinsic to the bulk sample, i.e., arising from CNT-CNT contacts. Details of some of the calculations are given in Appendixes A and B.

## II. FLUCTUATION-INDUCED TUNNELING CONDUCTION

In the model underlying the FITC, large metal grains are separated by insulating barriers and the contact gap between two metal grains can form a tunneling barrier with an effective capacitance as shown in Fig. 1(a). Figure 1(b) shows the potential profile of a junction barrier with an externally applied voltage  $V_A$  (or electric field  $\epsilon_A$ ) together with a thermally induced fluctuation voltage  $V_T$  (or electric field  $\epsilon_T$ ), which can be either along the direction of  $V_A$  or against it. The junction can experience two values of total voltage  $V_T + V_A$  or  $V_T - V_A$ . By assuming  $V_A \ll |V_T|$ , the resulting two tunneling currents  $I_+ = \frac{1}{2}I(V_T + V_A)$  and  $I_- = \frac{1}{2}I(V_T - V_A)$  (the factor of  $\frac{1}{2}$  arises since each has equal probability of occurrence) are opposite in directions. From Fig. 1(b), it is clear that the net current is given by

$$\Delta I = \frac{1}{2}[I(V_T + V_A) - I(V_T - V_A)]. \quad (1)$$

In the limit of  $V_A \rightarrow 0$ , the partial conductance with a fluctuation voltage  $V_T$  is defined as

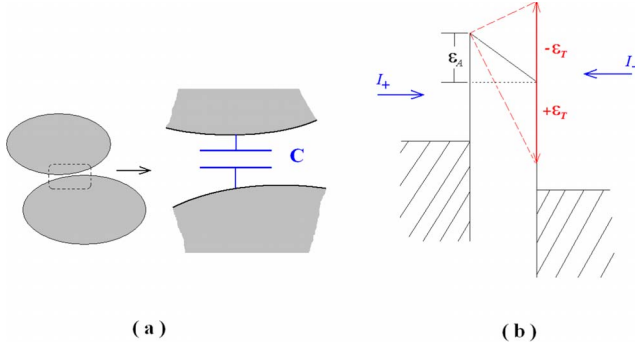


FIG. 1. (Color online) (a) Schematic illustration for the insulating junction between two large metallic granules, which may be regarded as a capacitor. (b) The tunneling potential barrier with the external electric field  $\varepsilon_A$  and the fluctuation electric field  $\varepsilon_T$  that can be along either the forward or the backward directions.  $I_+$  and  $I_-$  refer to the two tunneling currents, along opposite directions of  $\varepsilon_T + \varepsilon_A$  and  $\varepsilon_T - \varepsilon_A$ . Here we assume  $|\varepsilon_T| > \varepsilon_A$ . The dashed lines illustrate the effect of  $\varepsilon_T$  on the barrier.

$$\sum(V_T) = \lim_{V_A \rightarrow 0} \frac{\Delta I}{V_A} = \frac{dI(V_T)}{dV_T}. \quad (2)$$

To obtain the temperature dependence, it is observed that, since  $V_T$  is thermally induced, it obeys the Boltzmann distribution. This can be seen as follows.

The charging energy of the capacitor for the barrier junction is

$$E_C = \frac{1}{2} C V_T^2, \quad (3)$$

where  $C$  is the capacitance, given by

$$C = \varepsilon_0 \frac{S}{H}, \quad (4)$$

with  $\varepsilon_0$  being the dielectric constant in vacuum,  $S$  as the area, and  $H$  as the thickness of the capacitor (tunnel junction). According to the Boltzmann distribution, the distribution of thermally activated  $E_C$  for the capacitor is given by

$$P(E_C) = N \exp\left(-\frac{E_C}{k_B T}\right), \quad (5)$$

where  $N$  is the normalization constant. From Eqs. (3) and (5) we obtain the average value of  $V_T^2$  as

$$\langle V_T^2 \rangle = \frac{k_B T}{C}. \quad (6)$$

This is the same result as obtained from the equipartition theorem. Substitution of Eq. (3) into Eq. (5) leads to

$$P(V_T) = \left(\frac{2C}{\pi k_B T}\right)^{1/2} \exp\left(-\frac{C}{2k_B T} V_T^2\right), \quad (7)$$

where the normalization constant has been explicitly evaluated. At a finite temperature, the total conductance is an integral over all the partial conductances at potentially possible

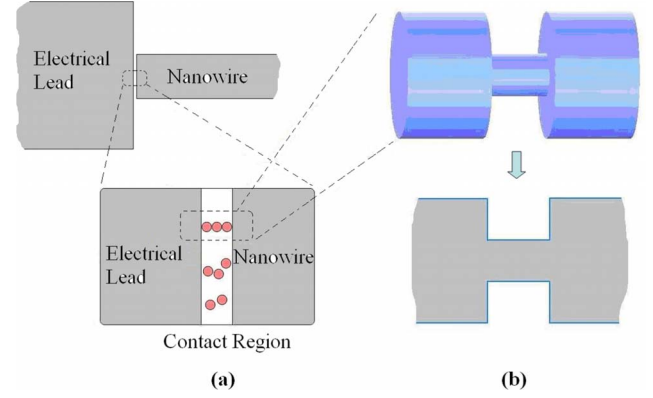


FIG. 2. (Color online) (a) Schematic illustration of the nanoconstrictions in the contact region between a nanowire and the metallic electrical lead. In the enlarged section (bottom panel) the red circles represent the metallic atoms. (b) The 3D (top) and 2D (bottom) schematic illustrations for a nanoconstriction connecting the conducting regions.

fluctuation voltages obeying the Boltzmann distribution. Hence

$$\sigma = \int_0^\infty P(V_T) \frac{dI(V_T)}{dV_T} dV_T, \quad (8)$$

with  $P(V_T)$  given by Eq. (7). Temperature dependence of the FITC directly follows. In fitting the experimental data, we shall treat the capacitance  $C$  to be slightly adjustable from its value as given by Eq. (4) so as to account for the possible deviations from the idealized geometry.

### III. MODEL PRESENTATION

We present our model by using as example the two-probe contacts with a metallic nanowire.<sup>6</sup> However, the arguments presented are applicable to other relevant scenarios as well.

Consider a *non-Ohmic* contact [the magnified figure in Fig. 2(a)]. There can be many possible contact microstructures. As there are no truly conducting contacts, the overall contact resistance is dominated by those structures that have the least resistance (as we assume the contact structures to be in parallel). The two conditions of being (1) nonconducting and (2) least resistance means that we can narrow down the possible structures to those constrictions of very small width made of (single) metallic atoms, which do not allow any conducting channel because the transverse dimension is below the waveguide cutoff threshold.

The realistic geometry of a nanoconstriction is naturally a three-dimensional (3D) structure connecting the two large metallic regions. The constriction can be regarded as a cylindrical segment as shown in the top panel of Fig. 2(b). If we consider only the azimuthally symmetric wave mode (reasonable for the lowest cutoff mode), then the problem can be simplified to two dimensions (2D). If we further approximate the cylindrical cross section by a square and require the mode number along the two transverse directions to be the same, then we reduce the geometry to that shown in the bottom panel of Fig. 2(b). For simplicity, we will treat the

problem in 2D and expect some differences with the 3D model. One is that the density of states (the number of modes) in the 2D case is less than that in the 3D case since the azimuthal degree of freedom is omitted. Such details of the 3D model will be pursued in the future. Also, for the 2D model the area ( $S$ ) of the capacitor [see Eq. (4)] will be set to be 0.02, 0.04, and 0.10 nm<sup>2</sup>, on the scale of single atom chain’s cross section. It is noted that the value of  $S$  does not significantly affect the predictions of the model.

IV. CALCULATIONAL APPROACH

At the mesoscopic scale, a transport “channel” can be viewed as an electronic waveguide mode. At zero temperature, the conductance of such a waveguide mode may be expressed by the Landauer formula<sup>7</sup>

$$G = \frac{2e^2}{h} t, \tag{9}$$

where  $e$  is the electron charge,  $h$  the Plank constant, and  $t$  the transmission coefficient for the electronic waveguide mode, which can be expressed as

$$t = \sum_{j=1}^M t_j. \tag{10}$$

Here  $t_j$  is the partial transmission coefficient for the mode “ $j$ ” in the waveguide,  $M$  being the total number of guided modes. The procedure for calculating  $t_j$  is detailed in Appendix A.

At nonzero temperatures, we must consider the smearing of the Fermi surface. In a two-terminal system with the chemical potential difference  $\Delta$  between the two electrical terminals, the net current may be written as<sup>7</sup>

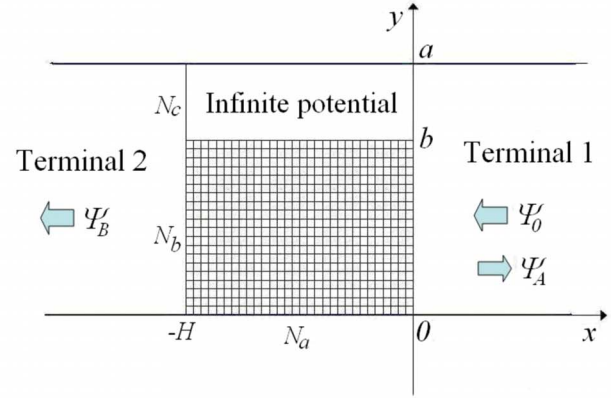
$$I(\Delta) = \frac{2e}{h} \int t(E, \Delta) [f(E) - f(E + \Delta)] dE, \tag{11}$$

where  $t(E, \Delta)$  is the transmission coefficient of an electron wave with incident energy  $E$  under a chemical-potential difference  $\Delta$  (or bias voltage  $V$ ,  $\Delta = eV$ ), and  $f(E)$  is the Fermi-Dirac distribution

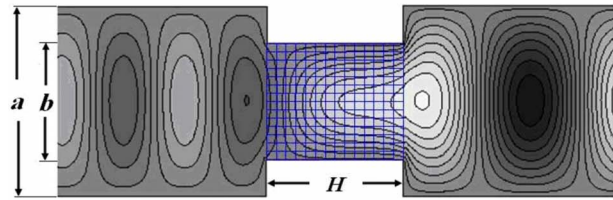
$$f(E) = \{1 + \exp[(E - \mu)/k_B T]\}^{-1}. \tag{12}$$

Given that the transmission coefficients can be evaluated, we obtain the  $I$ - $V$  relation from Eq. (11), from which  $dI/dV$  may be evaluated as a function of  $V$ . Then by using Eq. (8), the temperature dependence of conductance can be obtained through integration with the Boltzmann distribution.

To calculate the transmission coefficient  $t$ , the mixed boundary matching approach was used to solve the Schrödinger equation (see Appendix A for details) in which the wave functions in the electrical terminal regions are expressed as the superposition of (waveguide) eigenmodes, and in the sample region by the discrete basis (defined at each grid node). The values at the grid nodes and the expansion coefficients are the unknowns. The relevant equations are of two types: one is from the discretized Schrödinger equations for each grid node, and the other is from the boundary con-



(a)



(b)

FIG. 3. (Color online) (a) Geometry and the calculation coordinate for the two-terminal system. The lower sample region is discretized by a  $N_a$  by  $N_b$  mesh. For the upper infinite potential region, the grid point number along the  $y$  direction is  $N_c$  (see Appendix A for details). (b) A calculated example of the wave-function contours in a nanoconstriction. Here the terminal width is  $a=0.30$  nm; width of the nanoconstriction is  $b=0.19$  nm and its length is  $H=0.40$  nm. Fermi wavelength is 0.421 nm. The transverse dimension of the constriction is noted to be slightly less than half the Fermi wavelength.

ditions at the interfaces between the sample and the two terminals. Figure 3(a) shows the geometry and the relevant coordinate system for the two-terminal system.

Figure 3(b) shows a calculated example for electron wave transport through a nanoconstriction. An incident eigenmode from the right terminal area passes through the narrow constriction in the middle onto the left terminal area. The real part of the electron wave function is drawn as contours. The blue mesh in the sample region stands for the discrete grid upon which the Schrödinger equation was solved numerically by the finite difference method. The empty region, above the mesh in Fig. 3(a), refers to the area with an infinite potential barrier, in which electron wave is excluded. Here  $H$  is the length of the narrow constriction. The width of the terminal is taken to be  $a$  and the transverse dimension of the narrow constriction is denoted as  $b$ . Their values are given in the figure captions.

It should be noted that, in the case of tunneling transport, very fine grid must be used in order to ensure convergence. That fine grid necessarily increases the computational load.

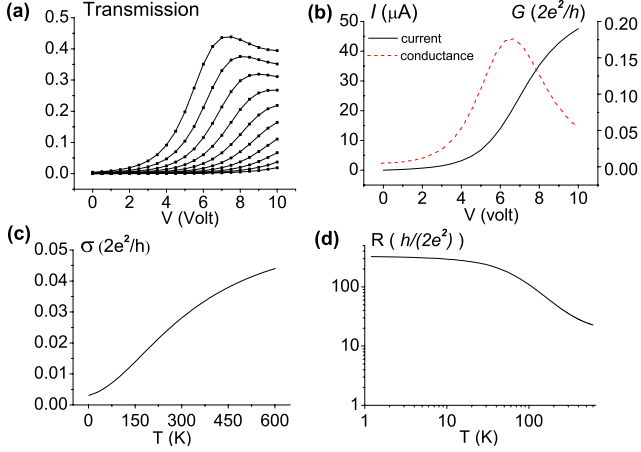


FIG. 4. (Color online) Simulation results for a nanoconstriction with  $a=0.30$  nm,  $b=0.19$  nm,  $H=0.40$  nm,  $\lambda_F=0.421$  nm, and  $C=4.43 \times 10^{-22}$  F. (a) Transmission plotted as a function of bias voltage. Each curve is for one incident energy, ranging from  $E_F$  (top) to  $(E_F-4.0$  eV) (bottom). (b)  $I$  vs  $V$  (solid curve) and  $G(dI/dV)$  vs  $V$  (dashed curve). (c) Conductance plotted as a function of temperature. (d) Resistance plotted as a function of temperature.

Some numerical technique, such as the sparse matrix approach, has been used to achieve efficiency.<sup>8</sup>

## V. RESULTS

The approach described above has been used to calculate the temperature variation in the fluctuation-induced tunneling conductance for the contact resistance to the metallic RuO<sub>2</sub> nanowire.<sup>6</sup> The Fermi energy of RuO<sub>2</sub> is  $\sim 8.5$  eV, obtained from the density-functional calculations.<sup>9,10</sup> The corresponding Fermi wavelength  $\lambda_F$  is 0.421 nm.

### A. Temperature dependence of the conductance

The calculated transmission curves with different bias voltage ( $V$ ) and incident energies ( $E_i$ ) are shown in Fig. 4(a) [ $V_{\max}=10$  (V)]. Here nanoconstriction's geometric parameters are  $a=0.30$  nm,  $b=0.19$  nm, and  $H=0.40$  nm. The cutoff wavelength is 0.38 nm ( $2b$ ), less than the Fermi wavelength in the RuO<sub>2</sub> nanowire.

From Eq. (11), the integral should be done over the range from  $E_F - eV_{\max}$  to  $E_F$  {the Fermi distribution in Eq. (12) is nearly a step function for up to  $T=500$  K since the width of  $f(E)$ 's sharp drop is  $k_B T$ , much less than the maximum chemical-potential difference ( $eV_{\max}$ ) [ $k_B T(500$  K)  $=0.043$  eV  $\ll eV_{\max}$ ]; thus the range  $[E_F - eV_{\max}, E_F]$  includes almost all the relevant electron energies}. If the incident energy is lower than  $(E_F - eV_{\max}/2)$ ,  $t(E_i, \Delta)$  is almost zero. Hence we only need to calculate the incident energy from  $E_F$  to  $(E_F - eV_{\max}/2.5)$ .

By using Eq. (11) to integrate the current at different incident energies, we obtain the bias voltage dependence of the total current [solid curve in Fig. 4(b)]. The differential conductance [dashed curve in Fig. 4(b)] is obtained through numerical differentiation.

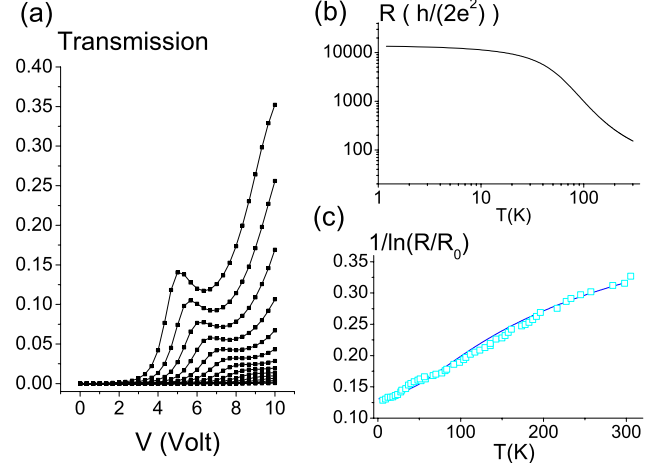


FIG. 5. (Color online) (a) Transmission plotted as a function of bias voltage. Here  $a=0.30$  nm,  $b=0.19$  nm,  $H=0.65$  nm,  $\lambda_F=0.421$  nm, and  $C=5.45 \times 10^{-22}$  F, and each curve is for one incident energy, ranging from  $E_F$  (top) to  $(E_F-4.0$  eV) (bottom). It is noted that the top curve exhibits a peak with a section to its right that might be interpreted as negative differential conductance for that energy. This behavior arises from quantum interference of the electronic wave and is the harbinger of the electronic Fabry-Perot effect to be detailed in Sec. VB. (b) Resistance dependence on temperature, for the same nanoconstriction as in (a). (c)  $1/\ln(R/R_0)$  vs  $T$  relation from the calculated 2D model (blue curve) and from the experimental data (colored open squares). The fitted FITC parameter values for the theory curve are  $T_1=1491$  K,  $T_0=190$  K, and  $R_0=6.8 \cdot (h/2e^2)$ .

Since the fluctuation voltage across the constriction is thermally induced, the above conductance results [shown in Fig. 4(b)] are integrated with the (thermal fluctuation) voltage distribution to obtain the temperature dependence of the conductance [see Eq. (8)]. The result is shown in Fig. 4(c). By inverting the conductance, the resistance-temperature ( $R$ - $T$ ) relation is obtained, shown in Fig. 4(d).

For a constriction with a length of 6.5 Å, the calculated transmission is plotted as a function of bias voltage in Fig. 5(a). Here each curve corresponds to one incident energy, ranging from  $E_F$  to  $E_F-4.0$  eV. With the same approach as described above, we obtain the  $R$ - $T$  relationship as shown in Fig. 5(b).

In the original FITC model,<sup>1,5,6</sup> the resistance-temperature relation has the following characteristic expression

$$R(T) = R_0 \exp\left[\frac{T_1}{T_0 + T}\right], \quad (13)$$

where  $T_1$  can be regarded as a measure of the energy required for an electron to overcome the energy barrier and  $T_0$  the temperature below which elastic tunneling conduction dominates. To compare our 2D tunneling model to the behavior given by Eq. (13), we rewrite Eq. (13) as a linear function of  $T$ :

$$\frac{1}{\ln(R/R_0)} = \frac{T}{T_1} + \frac{T_0}{T_1}, \quad (14)$$

where  $R_0$ ,  $T_0$ , and  $T_1$  are the fitting parameters. With reasonable parameter values, we find our simulation results to have



TABLE I. The fitting parameters obtained from the 2D tunneling model. The effective barrier height  $\varepsilon^*$  is evaluated in accordance with Eq. (15). Here the FITC parameter values  $T_0$  and  $T_1$  are obtained as the best fit of Eq. (14) to the theory curves.

Sample	$H$ (nm)	$b$ (nm)	$T_1$ (K)	$T_0$ (K)	$\varepsilon^*$ (eV)
H <sub>3</sub>	0.40	0.19	317	75	1.93
H <sub>1</sub>	0.65	0.19	1491	190	1.93
H <sub>6</sub>	1.10	0.21	19.3	1.2	0.04

a near-linear behavior [solid curve in Fig. 5(c)]. The theory is noted to agree very well with the experimental data from Lin *et al.* (from the  $H_1$  curve of Fig. 3 in Ref. 6), shown as colored squares in Fig. 5(c). Here the FITC fitting parameter values to our theoretical curve are  $T_0=190$  K,  $T_1=1491$  K, and  $R_0=6.8(h/2e^2)$ . We also fit our simulation results to the other data from Lin *et al.*<sup>6</sup> (see Table I), with good agreements as well.

Table I lists all the size parameters, adjusted to fit the experiment data displayed in Table II, and the fitted  $T_0$  and  $T_1$  values for the theory curves in the three cases. Here  $\varepsilon^*$  is the effective barrier height, evaluated from Eq. (15). We compare these parameters with the fitting parameters from Lin *et al.*'s<sup>6</sup> nanowire data using the original FITC model, listed in Table II. It is easy to see that our constriction lengths are much shorter and the barrier heights much larger, which are more physical. Our model also shows that, for the narrower tunneling constrictions, the effective barrier height can be much larger than that for the wider constrictions (but still below the cutoff).

For a nanoconstriction that is much longer—1.1 nm in length (0.21 nm in width), the transmission curves for the different incident energies and the corresponding  $I$ - $V$ ,  $G$ - $V$ , and  $R$ - $T$  relations are shown in Fig. 6. The parameters are fit to the data curve H<sub>6</sub> in Ref. 6. Their values are shown in third row of Table I.

For the temperature dependencies of resistance shown in Figs. 4–6, we can see that, at very low temperatures (less than 20K), the resistance is almost independent of the temperature, indicating elastic tunneling as the conduction mechanism. At higher temperatures, the resistance begins to decrease, owing to the thermal fluctuation mechanism.

Furthermore, comparison between the different  $R$ - $T$  relations for the three nanoconstrictions shows that in the third

TABLE II. The fitting parameters from Lin *et al.*'s experimental data (Ref. 6), using the form of Eqs. (13) and (14). The values of  $S$  are those estimated in Ref. 6 for the actual electrode contact areas. They can be significantly different from the values in the estimation of  $C$  [Eq. (4)] for the fluctuation voltages.

Sample	$H$ (nm)	$S$ (nm <sup>2</sup> )	$T_1$ (K)	$T_0$ (K)	$\varepsilon$ (meV)
H <sub>3</sub>	6.4	60 × 370	363	104	4.6
H <sub>1</sub>	15.7	400 × 650	1396	173	4.1
H <sub>6</sub>	8.7	70 × 400	20	8.6	1.1

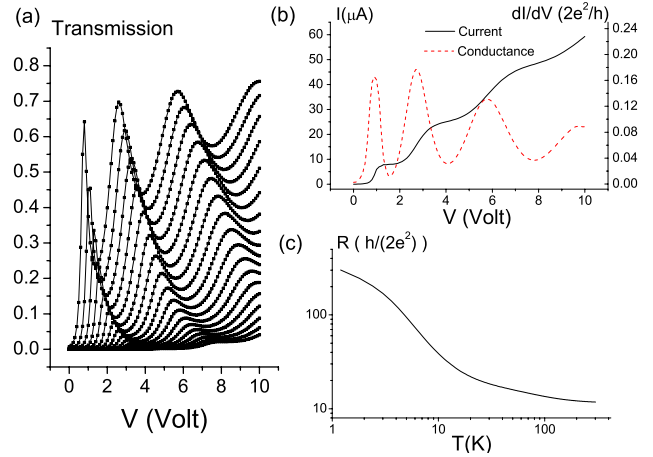


FIG. 6. (Color online) Simulation results for a nanoconstriction with  $a=0.30$  nm,  $b=0.21$  nm,  $H=1.1$  nm,  $\lambda_F=0.421$  nm, and  $C=8.05 \times 10^{-22}$  F. (a) Transmission dependence plotted as a function of bias voltage. The curves from top to bottom are for different incident energies, ranging from  $E_F$  to  $(E_F-2.6$  eV), respectively. (b)  $I$ - $V$  (solid curve) and  $G$ - $V$  curves (dashed curve) at  $T=0$  K. (c) Resistance plotted as a function of temperature.

case the resistance tends to decrease very quickly when temperature increases from zero, and at high temperatures the resistance tends to a constant. This can be explained as follows. The third case (shown in Fig. 6) involves a constriction which is wider (closer to the cutoff threshold) and hence easier for the electron wave to tunnel than the other two cases [see the transmission curves in Figs. 4(a), 5(a), and 6(a)]. As a result, at low temperatures the third case has a lower resistance [see Fig. 6(c)]. Also, because the relevant capacitance is smaller for the third case, the average fluctuation voltage tends to be larger than the other two cases, and the temperature-independent elastic tunneling process therefore tends to disappear more quickly (this is also evident from the small fitting parameter  $T_0$ ). This also explains the fact that at higher temperatures the  $R$ - $T$  curve tends to a constant, owing to the thermal activation effect.

## B. High voltage behavior

For a tunneling constriction that is relatively long (e.g., 1.1 nm) and only slightly less than the cutoff width (e.g., 0.21 nm, less than  $\lambda_F/2=0.2105$  nm), there can be an interesting prediction of the model at high applied voltage. Figure 6(a) shows the transmission curves as a function of the bias voltage and incident energy. The incident energy ranges from  $E_F$  to  $E_F-2.6$  eV.

After integrating all the incident current for the different incident energies, an  $I$ - $V$  curve is obtained (at 0 K). The differential conductance ( $dI/dV$  or  $G$ ) can be obtained from the  $I$ - $V$  curve through numerical differentiation, and both are shown in Fig. 6(b). Here we do not restrict ourselves to the low-field (voltage) regime but the applied voltage is still within a reasonable range, i.e., a few volts per nm. It is seen that there exists steplike  $I$ - $V$  behavior with corresponding peak-and-valley oscillations in the  $G$ - $V$  curve. This interesting behavior comes from the fact that, since the electronic

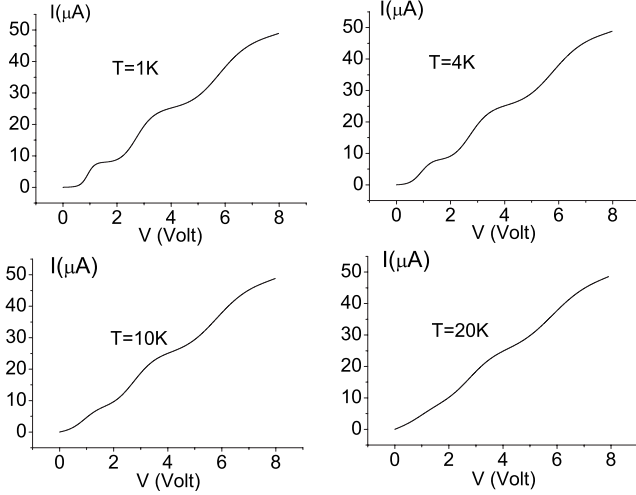


FIG. 7.  $I$ - $V$  curves of a nanoconstriction at different temperatures. Here  $a=0.30$  nm,  $b=0.21$  nm,  $H=1.1$  nm,  $\lambda_F=0.421$  nm, and  $C=8.05 \times 10^{-22}$  F. With increasing temperature, the steplike behavior is smeared.

energy is only slightly less than the cutoff threshold, once the electron gains energy from the applied electric field so that the cutoff threshold is exceeded, it can become a propagating mode, leading to interference effect owing to the finite length of the constriction. This phenomenon can serve as a predictive validation for our model, in conjunction with the temperature dependence of the resistance.

The  $I$ - $V$  curves at finite temperatures were also calculated (by including the influence of the voltage fluctuations as well as the smearing of the Fermi levels); the results are shown in Fig. 7. It is seen that, when the temperature increases, the fluctuation voltage becomes larger and hence the steplike behavior in the  $I$ - $V$  curve is less pronounced.

As mentioned before, the peaks and valleys in Figs. 6(a) and 6(b) come from interference of the electronic waves. Such interference is in exact analogy to the Fabry-Perot interference of light as it passes through a thin film. However, in contrast to the classical Fabry-Perot effect where the peaks are equally spaced as a function of frequency, here the positions of the peaks in the  $dI/dV$  vs  $V$  curve follow a *quadratic variation*. This is due to the different dispersion relations between the quantum Schrödinger wave (in which the energy is a quadratic function of wave vector) and the electromagnetic wave (in which the energy is a linear function of wave vector), shown in the following section.

It should also be noted that, under a high electric field, the electronic density might be altered, and the problem would have to be treated in a self-consistent manner. In our treatment it is assumed that the electronic density change is not so large as to cause significant distortion in our results.

### C. Equivalent one-dimensional tunneling

In this section we show that, even though the transport calculations are 2D, some relevant parameter (such as the effective potential barrier height  $\varepsilon^*$ ) and the steps in the  $I$ - $V$  behavior can be obtained from one-dimensional (1D) model considerations.

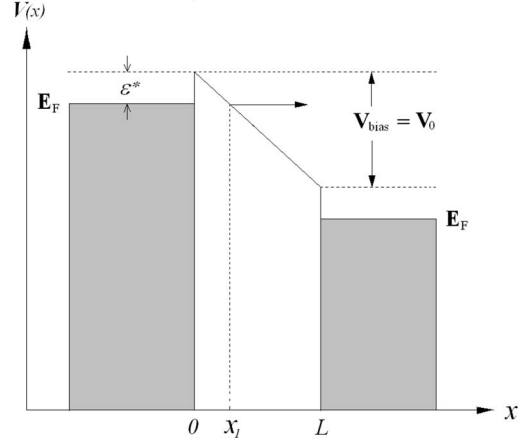


FIG. 8. A simplified 1D tunneling model to explain the  $\varepsilon^*$  evaluation and the appearance of peaks in the differential conductance vs applied voltage.

For the effective barrier height, it is simple to deduce it from the kinetic energy of the electron along the tunneling direction. That is, since the kinetic energy is given by  $\hbar^2(k_{\parallel}^2 + k_{\perp}^2)/2m$ , where  $k_{\parallel(\perp)}$  is the transverse (longitudinal) wave vector, the cutoff energy is given by  $k_{\perp}=0$  and  $k_{\parallel}=2\pi/\lambda_{\text{cut}}$ , where  $\lambda_{\text{cut}}=2b$  denotes the transverse dimension of the constriction. Hence  $\varepsilon_{\text{cutoff}}=(\hbar^2/2m)(2\pi/\lambda_{\text{cut}})^2$ . The effective barrier height is defined as  $\varepsilon^*=\varepsilon_{\text{cutoff}}-E_F$ , or

$$\varepsilon^* = \frac{\hbar^2}{2m} \left[ \left( \frac{2\pi}{\lambda_{\text{cut}}} \right)^2 - \left( \frac{2\pi}{\lambda_F} \right)^2 \right], \quad (15)$$

where  $\lambda_F$  is the Fermi wavelength. Numerical values of  $\varepsilon^*$  are listed in the last column in Table I for several 2D tunneling constrictions. For the first two cases,  $\varepsilon^*$  is of the order of 1.0 eV, which differs significantly from the unphysically small barrier height (several meV) obtained from the previous fits to the traditional FITC model (Table II).

To understand the quadratic separation of the peaks in the  $G$ - $V$  relation, consider a 1D tunneling model as shown in Fig. 8. On the left side the potential barrier is higher than the Fermi level by  $\varepsilon^*$ , and on the right side the Fermi level is lowered by a bias voltage  $V_0$ . A triangular tunneling barrier results. The range  $[0, x_1]$  is the tunneling range. In the other range  $([x_1, L])$  the electron has higher energy than the potential and the wave function is complex (and hence propagates like a wave) [see Eq. (16)]. This illustrates the fact that, once the electron gains energy from the electric field so as to exceed the cutoff threshold, it can convert itself from an exponentially decaying wave function to a wavelike propagating wave function.

For this 1D model, we use the WKB approximation<sup>11</sup> to write the wave function in the range  $[x_1, L]$  as

$$\psi(x) = \frac{C_1 e^{i\varphi(x)} + C_2 e^{-i\varphi(x)}}{4 \sqrt{\frac{2m}{\hbar^2} [E - V(x)]}}, \quad (16)$$

where  $\varphi(x) = \int_{x_1}^x \sqrt{\frac{2m}{\hbar^2} [E - V(x')]} dx'$ , and  $V(x)$  is the potential in the 1D case, depending on the bias voltage  $V_0$  and the potential barrier  $\varepsilon^*$ .

For the simulation case under consideration, the cutoff wavelength ( $0.21 \times 2 = 0.42$  nm) is just a bit smaller than the Fermi wavelength (0.421 nm) so  $\varepsilon^* \ll V_0$  (see Table I). We can omit the tunneling part (in Fig. 8,  $x_1 \approx 0$ ) and carry out the phase integration to obtain

$$\begin{aligned}\varphi(L) &= \int_0^L \sqrt{\frac{2m}{\hbar^2} [E - V(x)]} dx \\ &= \int_0^L \sqrt{\frac{2m}{\hbar^2} \frac{V_0}{L}} x dx \\ &= \frac{2}{3} \sqrt{\frac{2m}{\hbar^2}} V_0 L,\end{aligned}\quad (17)$$

where  $L$  is the constriction length and  $V_0$  is the bias voltage at  $x=L$ .

A simple derivation from Eq. (16) shows that the peak positions in  $|\psi(L, V_0)|^2$  (also for transmission- $V$  curve) obeys the condition

$$2\varphi(V_0) + \delta = 2\pi n, \quad (18)$$

where  $\delta$  is a constant phase factor and  $n$  is an integer. If we denote  $V_n$  as the peak values in the transmission vs bias voltage curve, we have

$$2\varphi_n + \delta = \frac{4}{3} \sqrt{\frac{2m}{\hbar^2}} V_n L + \delta = 2\pi n, \quad (19)$$

and it is easy to deduce from Eq. (19) that

$$V_n = K_1 n^2 + K_2 n + K_3, \quad (20)$$

where  $K_1$ ,  $K_2$ , and  $K_3$  are constants depending on  $L$ . As the separation between the peaks increases linearly, it follows that the peak positions in transmission-bias ( $t$ - $V$ ) curve follow a quadratic variation.

Since the conductance  $G(V)$  is the differential of  $I$ , i.e.,  $dI(V)/dV$ , and from Eq. (11)  $I(V)$  is the integral of  $t(E, V)$  over all energies,  $G$ - $V$  curve has a very similar shape with the  $t(E_F, V)$ - $V$  curve (Appendix B gives a detailed derivation to prove this point). Hence the peak position in  $G$ - $V$  curve must obey a quadratic variation.

## VI. CONCLUDING REMARKS

We present a model of FITC involving electron tunneling through nanoconstrictions. Landauer formula is used to calculate mesoscopic tunneling in finite-length nanoconstrictions with transverse dimensions less than half the Fermi wavelength. Through the consideration of thermally induced fluctuation voltages, we obtained a temperature dependence of conductivity very similar to the original FITC model but with more realistic parameter values. At finite voltages, step-like  $I$ - $V$  behavior is predicted which is a manifestation of electronic Fabry-Perot interference effect.

## ACKNOWLEDGMENTS

We would like to thank J. J. Lin and Y. H. Lin for discussions on the experimental results. Partial financial support from RGC Grant No. CA04/05.SC02 is gratefully acknowledged.

## APPENDIX A: TRANSMISSION CALCULATION IN A TWO-TERMINAL SYSTEM

Our model is based on the 2D geometry in which a sample is connected with two semi-infinite electrical terminals on the two sides. At the mesoscopic scale, the electrical terminals may be viewed as electronic waveguides and the sample may be viewed as the wave scattering area. In the tunneling problem, the terminals are wider than the nanoconstriction, which is our sample area. We set an infinite high potential wall on one side of the sample to narrow down the nanoconstriction, shown in Fig. 3(a). Since the electron wave function is excluded from the infinite potential area, we only discretize the remaining sample area.

With the coordinate system shown in Fig. 3(a), we can see that the electrical terminals are defined in the areas of  $x > 0$ ,  $x < -H$ , and  $0 \leq y \leq a$ . Assuming the potential in the terminal areas to be uniform along the  $x$  direction and substituting the basic waveguide solution  $\psi = e^{ik_x x} \varphi(y)$  into the Schrödinger equation

$$-\frac{\hbar^2}{2m} \vec{\nabla}^2 \psi + V(\vec{r}) \psi = E \psi, \quad (A1)$$

we obtain the differential equation for  $\varphi_n(y)$ ,

$$\frac{d^2 \varphi(y)}{dy^2} + \left[ \frac{2m}{\hbar^2} [E - V(r)] - k^2 \right] \varphi(y) = 0, \quad (A2)$$

through separation of variables. For simplicity, we define the upper and lower boundaries of the terminals by a hard wall potential:  $V(y)=0$  ( $0 < y < a$ ) and  $V(y)=\infty$  ( $y=0$  or  $a$ ). With the boundary conditions  $\varphi_n(0)=\varphi_n(a)=0$  the eigenfunctions can be simply expressed as

$$\varphi_n(y) = \sin\left(\frac{n\pi}{a} y\right), \quad (A3)$$

together with the eigenvalue  $k(n) = \sqrt{\frac{2m}{\hbar^2} E - \left(\frac{n\pi}{a}\right)^2}$ . Since the potential may be different on two sides of sample ( $E_1$  for terminal 1,  $x > 0$  and  $E_2$  for terminal 2,  $x < -H$ ), we define

$$k_r(n) = \sqrt{\frac{2m}{\hbar^2} E_1 - \left(\frac{n\pi}{a}\right)^2}, \quad (A4a)$$

$$k_l(n) = \sqrt{\frac{2m}{\hbar^2} E_2 - \left(\frac{n\pi}{a}\right)^2}. \quad (A4b)$$

The wave functions in the two electrical terminals may be expressed as

$$\psi^{(1)} = \psi_0 + \psi_A = e^{-ik_r(n_0)x} \varphi_{n_0}(y) + \sum_{n=1}^{N_1} A_n e^{ik_r(n)x} \varphi_n(y), \quad (A5)$$

$$\psi^{(2)} = \psi_B = \sum_{n=1}^{N_1} B_n e^{-ik_f(n)(x+H)} \varphi_n(y), \quad (\text{A6})$$

where  $\psi_0$  and  $\psi_A$  are the incident and reflected wave functions in terminal 1, respectively.  $\psi_0$  consists only one mode with index  $n_0$ .  $\psi_B$  is the transmitted wave function in terminal 2. All modes including evanescent waves are considered for accurate results. From the discrete sine transformation, we set  $N_1 = N_m - 2$  ( $N_m = N_b + N_c$ ) in Eqs. (A5) and (A6), where  $N_1$  is the number of free grid points at the sample/terminal interface.

The sample area ( $-H < x < 0$ ) consists of a low potential area ( $0 < y < b$ ) and a high potential (infinite) area ( $b \leq y \leq a$ ). The Schrödinger equation [Eq. (A1)] can be written as

$$\nabla^2 \psi_S + k_0^2 \varepsilon(\vec{r}) \psi_S = 0, \quad (\text{A7})$$

where  $\varepsilon(\vec{r}) = 1 - V(\vec{r})/E$  and  $k_0^2 = \frac{2mE}{\hbar^2}$ .

We use center difference discretization to write

$$\nabla^2 \psi_{i,j}^S \approx (\psi_{i+1,j}^S + \psi_{i-1,j}^S + \psi_{i,j+1}^S + \psi_{i,j-1}^S - 4\psi_{i,j}^S)/d^2,$$

where  $\psi_{i,j}^S$  is the discrete wave function at the grid node with coordinate  $(i, j)$ . Each unit cell is a square with size  $\Delta_x = \Delta_y = d$ , and  $d = \frac{H}{N_a - 1} = \frac{a}{N_m - 1}$ , where  $a = b + c$ ,  $N_m = N_b + N_c$ . The discretized Schrödinger equation is (we use  $P_{ij}$  to stand for  $\psi_{i,j}^S$ ) given by

$$(P_{i+1,j} + P_{i-1,j} + P_{i,j+1} + P_{i,j-1} - 4P_{i,j}) + (k_0 d)^2 \varepsilon(i, j) P_{i,j} = 0. \quad (\text{A8})$$

For the computational grid, there are  $N_a \times N_b$  nodes as shown in Fig. 3(a). Along the  $x$  direction the node index ranges from one to  $N_a$ , corresponding to the points from  $(0, 0)$  to  $(-H, 0)$ ; along the  $y$  direction the node index ranges from one to  $N_b$ , corresponding to the points from  $(0, 0)$  to  $(0, b)$ . There are a total of  $(N_a - 2)(N_b - 2)$  independent equations [Eq. (A8)] in the grid area.

On the sample-electrical terminal boundaries, we set up equations from the boundary conditions, which state that the wave function and its derivative should be continuous at the interfaces. At  $x=0$ , we have

$$\psi^{(1)}|_{x=0} = \psi_{1,j}^S, \quad (\text{A9a})$$

$$(d\psi^{(1)}/dx)|_{x=0} = (\psi_{1,j}^S - \psi_{2,j}^S)/d. \quad (\text{A9b})$$

By substituting Eq. (A5) into Eq. (A9), we get

$$1.0 \sin\left(\frac{n_0 \pi}{a} y_j\right) + \sum_{n=1}^{N_1} A_n \sin\left(\frac{n \pi}{a} y_j\right) = P_{1,j}, \quad (\text{A10a})$$

$$\begin{aligned} & -ik_r(n_0) 1.0 \sin\left(\frac{n_0 \pi}{a} y_j\right) + \sum_{n=1}^{N_1} A_n [ik_r(n)] \sin\left(\frac{n \pi}{a} y_j\right) \\ & = \frac{P_{1,j} - P_{2,j}}{d}. \end{aligned} \quad (\text{A10b})$$

These equations are used at every node at the  $x=0$  boundary,  $(1, y_j)$ , where  $j$  ranges from 2 to  $N_b - 1$ .

For a more precise calculation, we find that  $(P_{2,j} - P_{1,j})/d$  is the approximation of  $d\psi/dx$  at  $x=0$  with first-order accuracy but it is also the approximation of  $d\psi/dx$  at  $x=-d/2$  with second-order accuracy. So when the boundary is chosen at  $x=-d/2$ , it would lead to better precision. Thus Eq. (A10b) can be modified as

$$\begin{aligned} & -ik_r(n_0) e^{ik_r(n_0)d/2} \sin\left(\frac{n_0 \pi}{a} y_j\right) \\ & + \sum_{n=1}^{N_1} A_n [ik_r(n)] e^{-ik_r(n)d/2} \sin\left(\frac{n \pi}{a} y_j\right) \\ & = \frac{P_{1,j} - P_{2,j}}{d}. \end{aligned} \quad (\text{A10c})$$

Similarly, we have the boundary-condition equations at the boundary  $x=-H$

$$\sum_{n=1}^{N_1} B_n \sin\left(\frac{n \pi}{a} y_j\right) = P_{N_a, j}, \quad (\text{A11a})$$

$$\sum_{n=1}^{N_1} B_n [-ik_f(n)] \sin\left(\frac{n \pi}{a} y_j\right) = (P_{N_a-1, j} - P_{N_a, j})/d. \quad (\text{A11b})$$

The modified boundary-condition equation of Eq. (A11b) is (at  $x=-H+d/2$ ) given by

$$\sum_{n=1}^{N_1} B_n [-ik_f(n)] e^{-ik_f(n)d/2} \sin\left(\frac{n \pi}{a} y_j\right) = (P_{N_a-1, j} - P_{N_a, j})/d. \quad (\text{A11c})$$

In Eq. (A11), it should be noted that  $2 \leq j \leq N_b - 1$ .

It should be noted that, in “ $N_c$  range” (the infinite potential part,  $b < y < a$ ), only the wave-function continuity condition applies:

$$1.0 \sin\left(\frac{n_0 \pi}{a} y_j\right) + \sum_{n=1}^{N_1} A_n \sin\left(\frac{n \pi}{a} y_j\right) = 0, \quad (\text{A12})$$

$$\sum_{n=1}^{N_1} B_n \sin\left(\frac{n \pi}{a} y_j\right) = 0. \quad (\text{A13})$$

In Eqs. (A12) and (A13),  $N_b \leq j \leq N_m - 1$ . The total number of unknowns is noted to be  $N_x = 2N_1 + N_a(N_b - 2) = 2(N_b - 2) + 2N_c + N_a(N_b - 2)$ , which equals the total number of equations  $N_f = 4(N_b - 2) + 2N_c + (N_a - 2)(N_b - 2)$ . We assemble all these equations into a large matrix equation. Since there are many zero elements in the matrix, the sparse matrix technique can be used.<sup>8</sup>

After solving the equations, the transmission coefficient of the system can be evaluated. It starts from the flux expression



$$\vec{j} = \frac{\hbar}{m} \text{Im}(\psi^* \vec{\nabla} \psi). \quad (\text{A14})$$

By substituting the forms of the wave functions given above and integrating along the  $y$  direction (and taking into account the orthogonality of the eigenfunctions  $\{\varphi_n(y)\}$ ), total current  $I_x$  in electrical terminal 2 may be written as

$$I_x^{(2)} = \frac{\hbar}{m} \frac{a}{2} \sum_{n=1}^M |B_n|^2 k_f(n). \quad (\text{A15})$$

Here  $M$  is the number of guided modes in the terminal (evanescent modes have no contribution to the flux). The incident current in terminal 1 can also be obtained as

$$I_x^{(1)} = \frac{\hbar}{m} \frac{a}{2} k_r(n_0). \quad (\text{A16})$$

It follows that the transmission coefficient ( $I_x^{(2)}/I_x^{(1)}$ ) for the  $n_0$ th incident mode is given by

$$t(n_0) = \sum_{n=1}^M |B_n|^2 k_f(n) / k_r(n_0). \quad (\text{A17})$$

Under the *incoherent assumption* in mesoscopic physics, which states that in electrical terminals all the guided modes should be considered as incoherent with respect to each other,<sup>7</sup> the total transmission coefficient is the summation of those for all the allowed guided modes

$$t = \sum_{m=1}^M t_m = \sum_{m=1}^M \sum_{n=1}^M |B_n^m|^2 k_f(n) / k_r(m). \quad (\text{A18})$$

Here the expansion coefficients ( $B_n^m$ ) correspond to the transmission coefficient for the incident  $m$ th mode scattered into the outgoing  $n$ th mode.

## APPENDIX B: RELATION BETWEEN $t$ - $V$ AND $G$ - $V$ CURVES IN A TUNNELING CHANNEL

In this appendix we show that the transmission curve in Fig. 6(a) should follow the behavior of the  $dI/dV$ - $V$  ( $G$ - $V$ ) curve in Fig. 6(b). In other words, the similarity is not an accident.

We recall that the conductance  $G(V)$  is the differential of  $I(V)$ , and from Eq. (11),  $I(V)$  is the integral of transmission spectrum  $t(E, V)$  over all energy  $E$  [here we use  $t(E, V)$  instead of  $t(E, \Delta)$ , with  $V = \Delta/e$ ]. It may seem that, since the differentiation and integration are two inverse operations,  $G(V)$  should be the same as  $t(E, V)$ . But the variables in the differentiation (by  $V$ ) and in the integration (by  $E$ ) are not the same. Hence this equivalence cannot be directly deduced. However, it can be shown below that indeed  $G(V) \approx Ct(E_F, V)$ . The key is to use the approximation of Eq. (B2)

to combine the two variables ( $E, V$ ) into one function  $K(x)$ , then the differential and integral operations indeed cancel, leading to the desired result.

In low temperatures, Eq. (11) can be rewritten as

$$I(V) = \frac{2e}{h} \int_{E_F - Ve}^{E_F} t(E, V) dE, \quad (\text{B1})$$

where  $E$  is the incident energy and  $V$  the bias voltage.

In Fig. 6(a), from the transmission curves' dependencies on  $E$  and  $V_0$ , we can construct an expression for  $t(E, V)$  as

$$t(E, V) = K[Ve - (E_F - E)\delta] A[(E_F - E)\delta], \quad (\text{B2})$$

where  $K(Ve) = t(E_F, V)$ , which is the first transmission curve at the top of Fig. 6(a) ( $E = E_F$ ), and here  $\delta$  is a proportionality constant.

Equation (B2) means that all the transmission curves with different incident energy  $E$  is just like the first curve  $t(E_F, V)$ , except for some *displacement*  $[(E_F - E)\delta]$  along the  $x$  axis, with some amplitude modulation factor  $A(x)$  that has a smooth variation with  $E$ .

From the integral of Eq. (B2), the differential conductance may be written as

$$\begin{aligned} G(V) &= \frac{dI(V)}{dV} \\ &= \frac{2e}{h} \int_{E_F - Ve}^{E_F} \frac{dK[Ve - (E_F - E)\delta]}{dV} A[(E_F - E)\delta] dE \\ &= \frac{2e}{h} \int_{-\delta Ve}^0 \frac{dK[Ve + y]}{dV} A[-y] \frac{dy}{\delta}, \end{aligned}$$

where  $y = -(E_F - E)\delta$ . With partial integration, we obtain

$$\begin{aligned} G(V) &= \frac{2e^2}{h} \frac{1}{\delta} K(Ve + y) A(-y) \Big|_{-\delta Ve}^0 \\ &\quad - \frac{2e}{h} \frac{1}{\delta} \int_{-\delta Ve}^0 \frac{dA(-y)}{dy} K(Ve + y) dy. \end{aligned}$$

Since  $A(x)$  is a smooth function of  $x$ , we drop the second term above, then

$$G(V) \approx \frac{2e^2}{h} \frac{1}{\delta} [K(Ve)A(0) - K(Ve - \delta Ve)A(\delta Ve)]. \quad (\text{B3})$$

As the transmission evaluated from  $K(Ve - \delta Ve)A(\delta Ve) = t(E_F - Ve, V)$  ( $E = E_F - Ve$ ) is small [see Fig. 6(a)], the second term of Eq. (B3) can be omitted. As a result we have

$$G(V) \approx \frac{2e^2}{h} \frac{1}{\delta} K(Ve)A(0) = \frac{2e^2}{h} \frac{1}{\delta} t(E_F, V). \quad (\text{B4})$$

The similarity between  $G(V)$  and  $t(E_F, V)$  is thus established.

\*sheng@ust.hk

- <sup>1</sup>P. Sheng, Phys. Rev. B **21**, 2180 (1980); P. Sheng, E. K. Sichel, and J. I. Gittleman, Phys. Rev. Lett. **40**, 1197 (1978).
- <sup>2</sup>E. K. Sichel, P. Sheng, J. I. Gittleman, and S. Bozowski, Phys. Rev. B **24**, 6131 (1981).
- <sup>3</sup>J. Ederth, P. Johansson, G. A. Niklasson, A. Hoel, A. Hultaker, P. Heszler, C. G. Graqvist, A. R. van Doorn, M. J. Jongerius, and D. Burgard, Phys. Rev. B **68**, 155410 (2003).
- <sup>4</sup>S. Paschen, M. N. Bussac, L. Zuppiroli, E. Minder, and B. Hilti, J. Appl. Phys. **78**, 3230 (1995).
- <sup>5</sup>M. Salvato, M. Cirillo, M. Lucci, S. Orlanducci, I. Ottaviani, M. L. Terranova, and F. Toschi, Phys. Rev. Lett. **101**, 246804 (2008).
- <sup>6</sup>Y. H. Lin, S. P. Chiu, and J. J. Lin, Nanotechnology **19**, 365201 (2008).
- <sup>7</sup>S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).
- <sup>8</sup>A free software package UMFPACK is used in solving the sparse matrix equation. Its website is <http://www.cise.ufl.edu/research/sparse/umfpack/>
- <sup>9</sup>K. M. Glassford and J. R. Chelikowsky, Phys. Rev. B **47**, 1732 (1993).
- <sup>10</sup>L. F. Mattheiss, Phys. Rev. B **13**, 2433 (1976).
- <sup>11</sup>J. J. Sakurai, *Quantum Mechanics* (Addison-Wesley, Massachusetts, 1994).