PHYSICAL REVIEW B **71**, 045320 (2005)

Monte Carlo calculations for metal-semiconductor hot-electron injection via tunnel-junction emission

Ian Appelbaum*
Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts 02138, USA
and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

V. Narayanamurti
Division of Engineering and Applied Sciences and Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

(Received 8 July 2004; revised manuscript received 8 October 2004; published 18 January 2005)

We present a detailed description of a scheme to calculate the injection current for metal-semiconductor systems using tunnel-junction electron emission. We employ a Monte Carlo framework for integrating over initial free-electron states in a metallic emitter and use interfacial scattering at the metal-semiconductor interface as an independent parameter. These results have implications for modeling metal-base transistors and ballistic electron emission microscopy and spectroscopy.

DOI: 10.1103/PhysRevB.71.045320 PACS number(s): 73.23.−b, 72.10.Bg

I. INTRODUCTION

A tunnel junction consists of two conductors separated by an insulating potential barrier. When an electrical voltage bias is applied between the two conductors, electrons can convert their potential energy gained by the bias into kinetic energy by tunneling through the thin forbidden zone of the insulator.

Tunnel junctions were studied by Sommerfeld and Bethe[1] and since then have been used with much success for hot-electron spectroscopy. For instance, Giaever *et al.* studied metallic superconductor gap energies in the early 1960s.[2] More recently, magnetic tunnel junctions (formed between two metallic ferromagnets) have attracted much interest due to their potential as nonvolatile memory elements and hard-drive read heads.[3,4]

Tunnel junctions have been used for hot-electron injection from metals into semiconductors for decades. First, they were used in solid-state devices to make ultrafast transistors,[5,6] which mate a tunnel junction to a Schottky diode with a thin base. Later, they served as the basis for a three-terminal microscopy based on the scanning tunneling microscope, known as ballistic electron emission microscopy (BEEM) and its associated spectroscopy (BEES).[7,8] In this technique, the injected electron current is measured as a function of either the lateral probe position (BEEM) or the tip bias (BEES).

Most theoretical efforts describing metal-semiconductor hot-electron injection via tunnel-junction emission have focused on the latter application of tunnel-junction hot-electron injection (BEES). Many authors have presented theories to explain the features of BEES spectra on various metal-semiconductor systems, including those with heterostructure collectors.[9,10] In fact, although this paper concerns the Monte Carlo method for this purpose, this scheme has been used numerous times before to model various aspects of this problem: Schowalter and Lee used Monte Carlo to examine hot-electron transport in Au/Si,[11] Ke *et al.* used it in a simplistic analysis of Au/GaAs,[12] Bauer *et al.* used it to model Au/GaP and Mg/GaP,[13] and Lee has used a Monte Carlo method to simulate ballistic electron microscopy and attenuation below the Schottky interface.[14,15] de Andres *et al.* have also discussed Monte Carlo methods in their review on BEEM.[16] The present paper attempts to give a more detailed treatment of this theory and its finer points than has been previously described.

II. MONTE CARLO FOR BALLISTIC ELECTRONS

The entire process under investigation is essentially a series of two steps, as schematically shown in Fig. 1. The electrons from the emitter (in BEES, a scanning metal tip) first must couple to states in the metal-base layer via tunneling to form a base current. Then, they can couple to states in the semiconductor to form a collector current. The calculation of these currents involves the connection of one momentum state to another by using the principles of energy and parallel momentum conservation in the planar approximation. This is a good assumption for planar devices, such as the metal-base hot-electron transistor, and, in most cases, it is suitable for BEES.[17]

Traditional methods of modeling BEES spectra typically involve the evaluation of complex integrals over initial and final states. The Monte Carlo[18] method replaces numerical methods for solving continuous integrals with random sampling of the complicated integration region of phase space. Monte Carlo's power in this case is apparent in the fact that only the geometry of k space and knowledge of the connection coefficients are required to calculate values to arbitrary precision. This provides a means to alter the model simply for *ad hoc* additions, which would otherwise require substantial algorithm revision.

III. TUNNEL JUNCTIONS

For modeling tunnel-junction injection of hot electrons, we have only to sample the space of initial states (from the

 ©2005 The American Physical Society

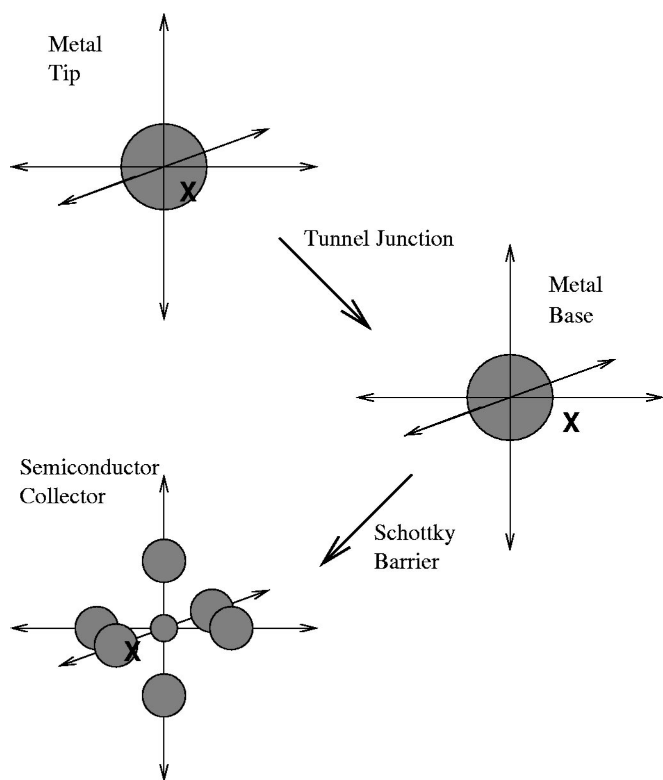

FIG. 1. Schematic diagram of the ballistic electron-injection process. The "X" shows the location of an example k-space state, which conserves both parallel momentum and energy throughout the process.

emitter), which contribute to the tunneling and transmitted hot-electron collector currents, summing the contributions from each sampling iteration.

The electrons involved in the calculation start their transport in the metal emitter (or STM tip). The simplest model for this metal is the free-electron gas, a model that takes into account only a quadratic dispersion relation and the Pauli exclusion principle. Since this means that

$$\frac{\hbar^2}{2m}(k_x^2+k_y^2+k_z^2) < E_F \tag{1}$$

for all occupied states at zero temperature, they form a sphere in k space whose radius is determined by the highest filled energy E_F (the Fermi energy), as shown in Fig. 1.

At nonzero temperature T, the edge of this sphere is smeared by thermal excitation according to the Fermi-Dirac distribution function,

$$P_{occupation}(E) = \frac{1}{e^{\beta(E-E_F)}+1}, \tag{2}$$

where $P_{occupation}$ is occupation probability and $\beta=1/(k_BT)$, where k_B is Boltzmann's constant.

The first step to the Monte Carlo method is to sample a state from among all that have non zero occupation probability. At zero temperature, this applies only to the states within the sphere, but at nonzero temperature, there are partially filled states at every point in k space. However, states far from the Fermi energy have very low occupation probability and so can, with very good approximation, be ignored. Subsequent calculations use a cutoff of $8k_BT$ past E_F.

In k space, this corresponds to a sphere with radius

$$k^2_{high\ cutoff} = \frac{2m}{\hbar^2}(E_F+8/\beta) = \frac{2mE_F}{\hbar^2}\left(1+\frac{8}{E_F\beta}\right), \tag{3}$$

$$k_{high\ cutoff} = \sqrt{\frac{2mE_F}{\hbar^2}}\sqrt{1+\frac{8}{E_F\beta}} \approx k_F(1+4/\beta E_F), \tag{4}$$

where k_F is the Fermi wave vector.

Our method of choosing an electron state is to randomly pick a point within a cube of side $2*k_{high\ cutoff}$ centered at $|\vec{k}|=0$, until it falls inside a circumscribed sphere of radius $k_{high\ cutoff}$. (This is known as the von Neumann rejection method.) Then, we first find the contribution of that state to the base tunneling current (the first step in Fig. 1) by calculating the products of the following:

(i) The thermal occupation probability of the initial (emitter) state.

(ii) The thermal vacancy probability of the final (base) state.

(iii) The incident charge flux on the metal-vacuum interface.

(iv) The tunneling probability.

Then we can calculate the probability that the tunneling electron makes the second step shown in Fig. 1 and contributes to the transmitted collector current.

We have already examined component (i). This probability is given by the Fermi-Dirac function. Component (ii) is the probability that the state in the base, which conserves energy and parallel momentum, is *not* occupied. Since the probability of occupation and vacancy must add to 1, the vacancy probability is given by

$$P_{vacancy}(E) = 1 - \frac{1}{e^{\beta(E-E_F+eV)}+1}, \tag{5}$$

where E is the kinetic energy in the emitter and eV is the potential energy gained from biasing the tunnel junction by a voltage V.

A. Electron flux

Component (iii), the incident flux, is proportional to the projection of the electron velocity in the direction normal to the interface ($\hbar k_\perp/m$); faster moving electrons will impinge on the interface more often and thus have a greater probability of contributing to the tunnel current. We immediately see that the original spherical phase space we want to sample can be reduced to a hemisphere because electrons with velocities having negative components in the emission direction (perpendicular to the interface) can be excluded; they are not incident on the metal-vacuum interface and therefore do not contribute to the base current.

The proportionality constant relating charge flux to perpendicular velocity is the phase-space volume of the sampled state times the electron charge e. To calculate this quantity $e\Delta^3k$, we must know how many states there are in the Fermi

sphere. This is determined by the normalization condition

$$n = \int_{|k|<k_F} \frac{d^3k}{(2\pi)^3} \tag{6}$$

(at zero temperature).

The real-space electron density n is an empirically measurable quantity. The Monte Carlo equivalent of this integral is

$$\sum \Delta^3 k = 2N\Delta^3 k, \tag{7}$$

where N is the sampling number. The factor of 2 is necessary because, although the integral is over the entire sampling sphere, our Monte Carlo sampling is restricted to only the forward-directed hemisphere. Therefore, $\Delta^3 k = n/2N$. For nonzero temperature, we replace N with the sum of the occupation probabilities,

$$N_{thermal} = \sum \frac{1}{e^{\beta(E-E_F)}+1} \tag{8}$$

because we are sampling over a volume that is less dense with electron states.

B. Tunneling probability

Component (iv), the tunneling probability, can be a nontrivial exercise in application of the Schrödinger equation if calculated exactly. However, we make use of an important simplification. The tunnel barrier height (for metal-base transistors: oxide conduction band offset; for BEEM/BEES: the work function of the metal tip) is many eV. The maximum tunneling electron energy (given by the applied bias voltage) is typically around 1 V. The Wentzel-Kramers-Brillouin (WKB) method, an approximation scheme, works well for tunneling processes such as this in which the tunneling energy is far below the barrier height.

The WKB tunneling probability is

$$P^{WKB} = e^{-2\gamma}, \tag{9}$$

where

$$\gamma = \frac{1}{\hbar}\int_a^b p(x)dx. \tag{10}$$

The integration bounds a and b are the bounds of the barrier.

For the planar approximation, the applied bias voltage drops evenly across the entire vacuum gap. Therefore, the barrier is a trapezoid, depicted in Fig. 2. We have, then,

$$\gamma = \frac{1}{\hbar}\int_0^{d_{gap}} \sqrt{2m\left(-\frac{eV}{d_{gap}}x + E_F + \phi - E\right)}dx, \tag{11}$$

where d_{gap} is the vacuum gap width, V is the applied bias voltage, and ϕ is the barrier height. In this planar approximation, E is the energy determined by the one-dimensional perpendicular component of the momentum $\hbar^2 k_\perp^2/2m$.

This is a trivial integral; we evaluate it to

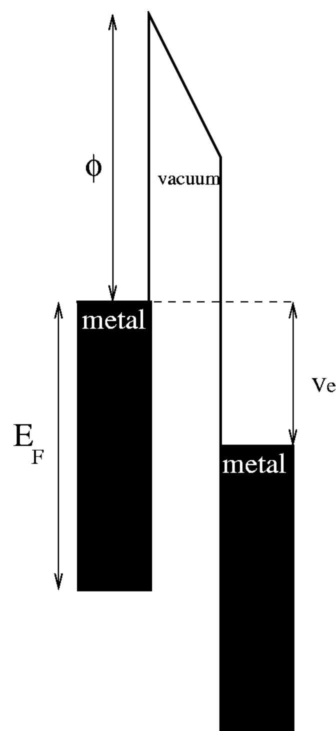

FIG. 2. Schematic tunnel-junction energy diagram.

$$\gamma = \frac{2^{3/2}m^{1/2}d_{gap}}{3\hbar eV}[(E_F+\phi-E)^{3/2} - (E_F+\phi-eV-E)^{3/2}]. \tag{12}$$

C. Base current summary

Now we have all the components of the base current contribution of each sampled electron. Since this is a planar theory, the current we calculate is actually a current *density*. We multiply by A, the effective junction area, to convert to a real current.

Our calculation is now to simply evaluate, for states within the sampling sphere, the sum

$$I_{\rm base} = eA\frac{\hbar}{m}\frac{n}{2N}\sum k_\perp P_{occupation}(E) \times P_{vacancy}(E+eV)P^{WKB}(k_\perp). \tag{13}$$

The dominating factor within this sum is the exponentially dependent WKB transmission probability, which is responsible for the sharply peaked energy probability shown in Fig. 3. The parallel momentum distribution is similarly peaked close to $k_\parallel = 0$ (Ref. 19).

D. A simplifying condition

We can now pause to notice a simplifying condition in our method. For emitter electron states with low energy (for instance, $E+eV \ll E_F - \beta$), the vacancy probability for a compatible state in the base is negligibly small. Therefore, we can restrict our sampling space further by ignoring initial

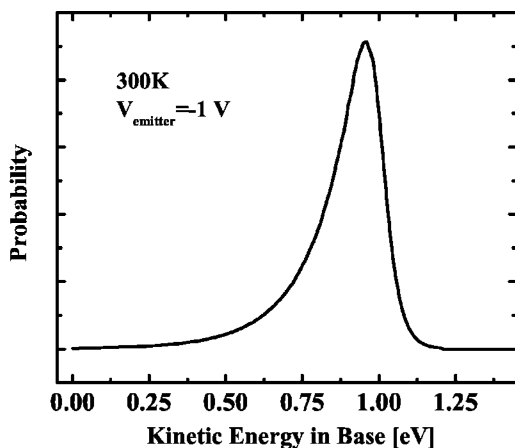

FIG. 3. Energy distribution of tunneling electrons from Au through a vacuum tunnel barrier with 1 V across it. The nonzero temperature (300 K) causes a smearing of the distribution near the cusp corresponding to the emitter Fermi energy.

emitter states with energy below a low-energy cutoff; subsequent calculations sample initial states with energy greater than

$$E_{low\ cutoff} = E_F - eV - 8/\beta, \tag{14}$$

resulting in a cutoff wave vector

$$k^2_{low\ cutoff} = \frac{2m}{\hbar^2}(E_F - eV - 8/\beta). \tag{15}$$

By decreasing the phase-space volume of sampled points, our sampling density increases. This greatly improves the accuracy of the calculation because the effect of discretizing phase space is minimized as the limit of continuous sampling is approached. However, our calculation must reflect this limitation of phase-space sampling; $\Delta^3 k$ must now be calculated with an effective sampling

$$N' = N_{thermal} + \frac{N}{V_{shell}} V_{sphere}. \tag{16}$$

The first term is the contribution from the electrons that form the base current, with $E_{low\ cutoff} < E < E_{high\ cutoff}$. These electrons are in states that form a spherical shell in k space. The second term accounts for the remaining electrons with $E < E_{low\ cutoff}$, which is a sphere. These states have occupation probability very close to unity, and, therefore, we need only count how many of them we would have sampled without the cutoff using the density determined by N, the sampling number.

E. Results of Monte Carlo tunnel current calculation

Using this Monte Carlo framework, we have calculated the tunnel-junction current-voltage characteristics using 10^6 samples in the emitter k space. This is shown in Fig. 4. We see at least two distinct regimes: at low bias, the current varies linearly with voltage and, at higher bias $\approx E_F$, the current increases exponentially.[20] We can understand this behavior with a simple analysis of our theory.

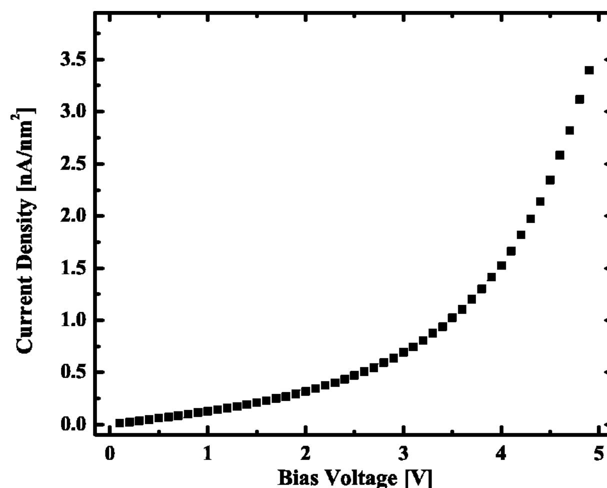

FIG. 4. Monte Carlo calculation of tunnel-junction current-voltage spectrum. The barrier width is 6 Å and the height is 5 eV above the Fermi energy (5 eV).

At low bias, the number of electrons contributing to the base current is small. These electrons originate in states close to the Fermi energy of the emitter, in a thin shell of outer radius k_F and inner radius of $\sqrt{2m(E_F - eV)}/\hbar$. Since $k_\perp$ varies little over the thickness of the shell, P^{WKB} and the incident flux can be treated to first order as a constant. The dominant effect determining the current change as the bias voltage changes is simply the change in the number of states contributing to the base current. This is determined by the volume of the spherical shell,

$$\text{Volume} = \frac{4}{3}\pi\left\{\left[\frac{\sqrt{2mE_F}}{\hbar}\right]^3 - \left[\frac{\sqrt{2m(E_F - eV)}}{\hbar}\right]^3\right\} \propto E_F^{3/2} - (E_F - eV)^{3/2}. \tag{17}$$

For small voltages, $eV \ll E_F$, we can make the approximation

$$\text{Volume} \propto E_F^{3/2} - E_F^{3/2}\left(1 - \frac{3eV}{2E_F}\right) \approx E_F^{3/2}\frac{3eV}{2E_F} = \left(\frac{3e\sqrt{E_F}}{2}\right)V. \tag{18}$$

Therefore, we see that at low bias voltage, the volume (and, hence, the tunnel current) varies, to first order, linearly with bias voltage.

For $eV > E_F$, every state contributes to the base current, so the above analysis does not apply. The current variation in this so-called Fowler-Nordheim regime is determined by the tunneling probability, which increases exponentially with increasing bias voltage. Even for $eV < E_F$, the volume of contributing states changes only a small amount as the voltage varies due to the small density of states at low energy, so the exponentially increasing regime begins below the Fermi energy.

IV. COLLECTOR CURRENT CALCULATION

Now we focus our attention on the collector current. This current is created by the tunneling electrons that ballistically

travel through the metal base and couple with available states in the semiconductor. Therefore, we must model the following:

(i) Ballistic elastic and inelastic attenuation in the metal base.

(ii) Interfacial elastic scattering at the metal-semiconductor Schottky interface.

(iii) Coupling to semiconductor states.

(iv) Quantum mechanical reflection at the Schottky interface.

A. Attenuation

Component (i) is modeled by using empirical elastic and inelastic mean-free paths for ballistic electrons.[10] The mean-free path is a length scale over which the electrons are attenuated by scattering in an average sense. The length scales are determined for different scattering processes; inelastic scattering is, in general, energy dependent because the phase space into which the electron can scatter changes with its initial kinetic energy if energy is not conserved in the scattering process. Higher-energy electrons are scattered more and therefore have shorter mean-free paths than low-energy electrons near the Fermi level.

Combining both elastic and inelastic processes yields

$$P_{attenuation} = \exp\left[-d_{metal}\left(\frac{1}{\lambda_{elastic}} + \frac{1}{\lambda_{inelastic}}\right)\right], \quad (19)$$

where d_{metal} is the thickness of the metal layer.

B. Interfacial scattering

Component (ii) is an addition to the original model of ballistic electron-emission spectroscopy incorporated to take into account empirical observations. It was found that the theoretical spectra fit the experiment observations far better if a portion of the transmitted current were elastically scattered, randomizing the parallel momentum. Therefore, we model this process in the Monte Carlo framework by randomly choosing, with a probability *SP*, electrons to be scattered via a random reorientation of the wave vector. To assure that this is an elastic process, the norm of the wave vector (proportional to the kinetic energy) is conserved. The algorithm to do this is very similar to our state-sampling algorithm: we choose a random point within a unit sphere and then divide by its norm to make it a unit vector. This unit vector is multiplied by the norm of the original wave vector.

C. Semiconductor conduction band model

We now discuss component (iii). A simple model of the semiconductor states in k space is the *spherical band* model. In this model, each of the conduction band valleys have spherical constant-energy surfaces with k-space radius

$$R = \frac{\sqrt{2m^*_{DOS}(E - E_{min})}}{\hbar}, \quad (20)$$

where m^*_{DOS} is the density of states effective mass and E_{min} is the conduction band minimum of the conduction valley, with respect to zero energy in the base metal.

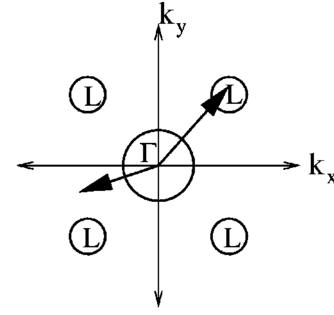

FIG. 5. Schematic Interface Brillouin Zone for GaAs(100). The arrow on the upper right indicates parallel momentum compatible with available semiconductor states, while the arrow in the lower left indicates parallel momentum at which there are no available states.

To find the semiconductor states compatible with our sampled electron state, we again apply conservation of energy and parallel momentum. The constant energy spheres are constructed and then projected onto the interface Brillouin zone (IBZ), or $k_\parallel$ plane. A schematic IBZ for GaAs(100), for energy greater than E_{min} for the two lowest conduction band valleys Γ and L, is shown in Fig. 5.

If we also project the electron wave vector on this IBZ, it may overlap a projected constant-energy sphere of one of the conduction band minima. If so, it contributes to the collector current because both energy and parallel momentum are conserved. However, its $k_\perp$ necessarily changes due to the band-structure effects. This discontinuity in wave vector causes quantum-mechanical reflection at the metal-semiconductor interface. The electron is transmitted with probability[9]

$$P_{QM} = \frac{4k_s k_m \frac{m_s}{m_m}}{\left(\frac{m_s}{m_m}k_m + k_s\right)^2}, \quad (21)$$

where m_m is the effective electron mass in the metal, m_s is the effective electron mass in the semiconductor, and k_m and k_s are the perpendicular components of the electron wave vectors in the metal and semiconductor, respectively. In a "free-electron" metal, $m_m=1$, a common approximation.

The collector current is then the evaluation of the sum

$$I_{collector} = eA\frac{\hbar}{m}\frac{n}{2N'}\sum k_\perp P_{occupation}(E) \times P_{vacancy}(E+V)P^{WKB}(k_\perp)P_{attenuation}(E) \times P_{QM}(k_\perp,E)\int \delta(k_s^\parallel - k_m^\parallel)d^2k, \quad (22)$$

where the integral over the δ function accounts for parallel momentum conservation at the metal-semiconductor interface.

V. BALLISTIC ELECTRON EMISSION SPECTROSCOPY

In contrast to solid-state tunel junctions where the geometry (and, hence, the tunnel-current–emitter-voltage relation-

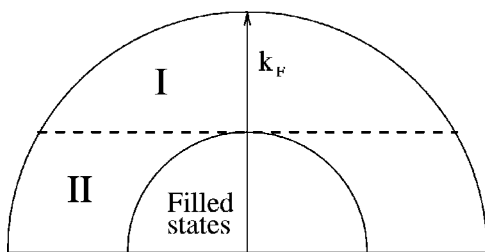

FIG. 6. A 2D projection of the shelllike integration region used in the calculation of the tunnel junction current.

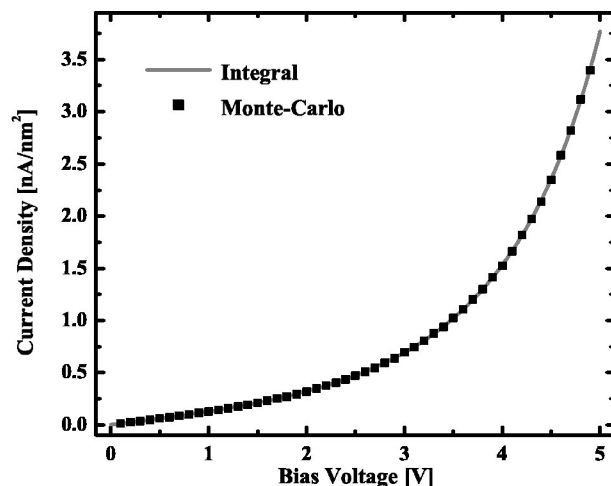

FIG. 7. Integral calculation of tunnel junction current-voltage spectrum. The barrier width is 6 Å and the height is 5 eV above the Fermi energy (5 eV). Also shown for comparison are the results of the previously presented Monte Carlo method.

ship) is fixed, the tunneling distance with scanning probe microscopy is variable. In the BEES experiment, as the voltage is varied, the tunnel current is held fixed by increasing the width of the vacuum barrier accordingly via dynamic feedback of tip height. To model this in our simulation, we need to calculate the base current before the collector current is calculated. However, implementing feedback directly in the Monte Carlo algorithm would be very inefficient. Therefore, we accelerate the process of finding the proper vacuum gap by using an integral expression for the tunnel current.

A. Integral expression for tunnel current

We know from the discussion of the zero-temperature tunnel-current theory that in the Monte Carlo formalism,

$$I_{base} = eA\frac{\hbar}{m}\sum k_{\perp}(E)P^{WKB}(E)\Delta^3 k. \tag{23}$$

We convert this sum into an integral by making the standard substitution

$$\Delta^3 k \rightarrow \frac{d^3k}{(2\pi)^3} \tag{24}$$

and specifying the bounds of integration.

In analogy to the arguments used before, our integration bounds enclose a half shell in k space with outer radius k_F and inner radius $\sqrt{2m(E_F-eV)}/\hbar$. Figure 6 schematically shows a projection of this region, broken up into two regions labeled I and II. Since the integrand is dependent only on $k_{\perp}$, we can reduce the three-dimensional integral to a one-dimensional integral by a suitable choice of infinitesimal volume element.

Region I is a spherical cap, with $\sqrt{2m(E_F-\mathrm{eV})}/\hbar < k_{\perp} < k_F$. To exploit the one-dimensional dependence, we use a thin disk as a differential volume element. The radius of this disk r is given by the equation of the circle projected by the Fermi sphere,

$$k_{\perp}^2 + r^2 = k_F^2, \tag{25}$$

and because the area of the circle is $\mathcal{A}=\pi r^2$, we have for the differential volume element

$$d\mathcal{V} = \mathcal{A}\cdot dk_{\perp} = \pi(k_F^2 - k_{\perp}^2)dk_{\perp}. \tag{26}$$

Region II is the portion of the spherical shell for which $0 < k_{\perp} < \sqrt{2m(E_F-\mathrm{eV})}/\hbar$. The volume element for this shell is then an annular disk. The area of the annulus is given by the difference between the areas of two concentric disks projected by the Fermi sphere and the inner sphere,

$$\mathcal{A} = \pi\left\{(k_F^2 - k_{\perp}^2) - \left[\frac{2m(E_F-eV)}{\hbar^2} - k_{\perp}^2\right]\right\} = \pi\frac{2meV}{\hbar^2} \tag{27}$$

with infinitesimal volume

$$d\mathcal{V} = \pi\frac{2meV}{\hbar^2}dk_{\perp}. \tag{28}$$

Therefore, we have

$$I_{base} = \frac{2\hbar}{m(2\pi)^3}\left(\int_0^{\sqrt{2m(E_F-eV)}/\hbar} k_{\perp}e^{-2\gamma}\pi\frac{2meV}{\hbar^2}dk_{\perp} + \int_{\sqrt{2m(E_F-eV)}/\hbar}^{k_F} k_{\perp}e^{-2\gamma}\pi(k_F^2-k_{\perp}^2)dk_{\perp}\right). \tag{29}$$

Since these two integrals cannot be evaluated analytically, we must use numerical quadrature. We now show the results of this calculation of tunnel current as a function of bias voltage in Fig. 7. Plotted on the same axis is the result of the Monte Carlo simulation previously presented for comparison. The remarkable agreement confirms the effectiveness of the Monte Carlo method.

To use these results for vacuum gap compensation to hold the tunnel current constant, we use a bisection algorithm. The effectiveness of this scheme allows the base current to be held to within 10^{-3} of the set point in the Monte Carlo simulation with negligible computational overhead.

B. Electron injection results

Due to its relative simplicity, we calculate the BEES for Au/GaAs for voltages where the only contribution is from the Γ valley (conduction band minimum). The Schottky barrier used in this calculation is 0.92 eV, and $m^*=0.067$. These results are shown in Fig. 8.

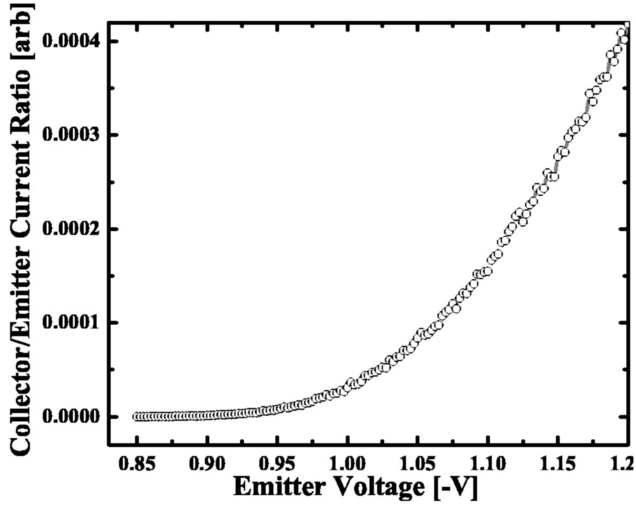

FIG. 8. Monte Carlo BEES calculation for Au/GaAs at 300 K. This simulation uses 10^8 samples of k space.

Notice that the collector current is zero for voltages below the Schottky barrier and increases near quadratically thereafter. It has been claimed that the leading order of the voltage dependence in BEES is actually a 5/2 power law.[21] From our development of the theory, we can see how this 5/2 power law arises.

C. Threshold behavior

The three major contributing factors to this power law are conservation of energy, conservation of momentum, and quantum-mechanical scattering. Here we determine the leading order contributions from each in the zero-temperature BEES theory.

1. Conservation of energy

Consider a system where the applied voltage (and thus the potential energy drop) between tip and base is just ϵ greater than the Schottky barrier height E_{SB}. Then, due to conservation of energy across the tunnel barrier,

$$\frac{\hbar^2 \vec{k}_{base}^{\,2}}{2m} = \frac{\hbar^2 \vec{k}_{tip}^{\,2}}{2m} + E_{SB} + \epsilon. \tag{30}$$

The condition for injection of electrons into the semiconductor is

$$\frac{\hbar^2 \vec{k}_{base}^{\,2}}{2m} \geqslant E_F + E_{SB}. \tag{31}$$

Combining Eqs.(30) and (31), we have

$$\frac{\hbar^2 \vec{k}_{tip}^{\,2}}{2m} \geqslant E_F - \epsilon. \tag{32}$$

Therefore, at just above threshold, only the electrons in the tip with more kinetic energy than $E_F - \epsilon$ will satisfy the total energy requirement. For these electrons, we have

$$|\vec{k}_{tip}| \geqslant \sqrt{\frac{2m}{\hbar^2}(E_F - \epsilon)}. \tag{33}$$

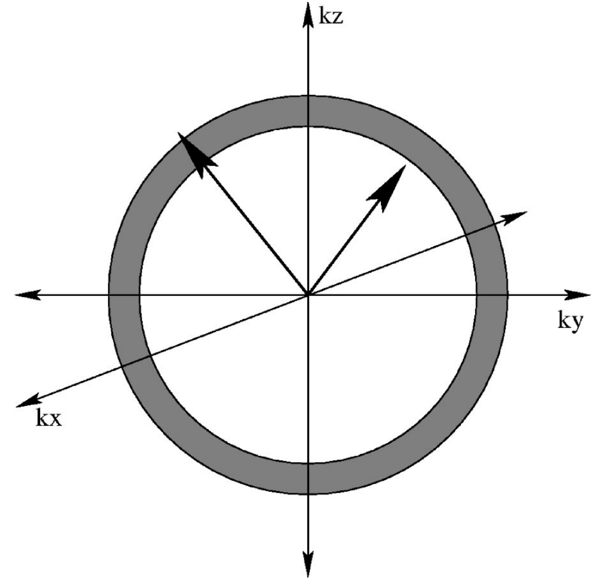

FIG. 9. A schematic of the "Fermi sea" in the tip metal. The arrow on the left represents the Fermi wave vector k_F and the arrow on the right represents the minimum electron wave vector magnitude [Eq. (33)]. At a tip bias above the Schottky barrier, only electrons within a shell (dark shaded region) between spheres of the two radii have enough energy to couple with states in the semiconductor.

The volume in k-space of those electrons that have enough energy to scatter into semiconductor states above the Schottky barrier is then the difference in volumes of two spheres with radii $|\vec{k}_{tip}| = k_F$ and $|\vec{k}_{tip}| = \sqrt{2m(E_F - \epsilon)/\hbar^2}$, as shown in Fig. 9:

$$\text{Volume} = \frac{4}{3}\pi\left(\sqrt{\frac{2m}{\hbar^2}}\right)^3 [(\sqrt{E_F})^3 - (\sqrt{E_F - \epsilon})^3]. \tag{34}$$

Expanding to first order gives

$$\text{Volume} \approx 2\pi\left(\sqrt{\frac{2m}{\hbar^2}}\right)^3 \sqrt{E_F}\epsilon. \tag{35}$$

Therefore, the volume of states in k space which contribute to the BEES collector current is proportional to ϵ.

2. Conservation of momentum

Not all electrons with enough energy will couple into semiconductor states. The component of k parallel to the interface plane must be compatible with available states in the semiconductor.

Consider the available states of the Γ valley, just above the Schottky barrier. These states form a surface, which projects a disk onto the interface Brillouin zone. Assuming quadratic $E(k)$ dependence, the radius of this disk is proportional to $\sqrt{\epsilon}$.

Since only the electrons with $\vec{k}_\parallel$ inside this disk will enter the semiconductor, the contribution to the BEES collector current above the Schottky barrier will be proportional to the ratio of the areas of the Γ valley disk and another disk with radius $k_\parallel$ characteristic of the tunneling process. This situa-

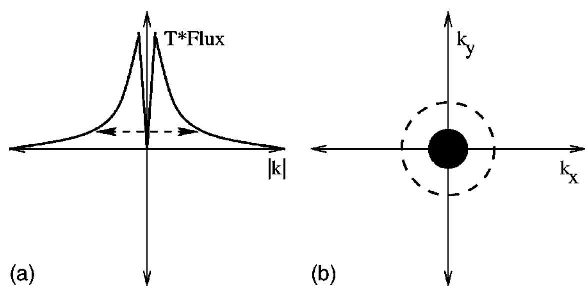

FIG. 10. (a) The tunnel process gives rise to a momentum distribution with a characteristic width (see Ref. 19, Fig. 3). (b) This distribution is compared to the projection of the available semiconductor states.

tion is depicted schematically in Fig. 10 (Ref. 19).

Therefore, momentum conservation contributes another factor proportional to the projected area $\pi(\sqrt{\epsilon})^2 \propto \epsilon$ to the behavior of the BEES collector current.

3. Quantum mechanical scattering

Due to quantum-mechanical scattering, the transmission coefficient for a state in the base metal with wave vector k_m coupling with a state in the semiconductor k_s is

$$P_{QM} = \frac{4k_s k_m \dfrac{m_s}{m_m}}{\left(\dfrac{m_s}{m_m}k_m + k_s\right)^2} = \frac{4\dfrac{m_s}{m_m}}{k_m\left(\dfrac{m_s}{m_m} + \dfrac{k_s}{k_m}\right)^2} k_s. \tag{36}$$

At just above threshold, k_s is very small and k_m is very large. Assuming again that the band structure has quadratic valleys, we see that

$$P_{QM} \propto k_s \propto \sqrt{\epsilon}. \tag{37}$$

These three threshold behaviors multiply to give the 5/2 power law.

VI. EFFECT OF SCATTERING PROBABILITY

We mentioned earlier the inclusion of the scattering probability (SP) to the theory so that it would more reflect empirical observations in certain systems. Here we explain this assertion further.

Since the probability that an electron tunnels across the insulating barrier increases with perpendicular momentum, the vacuum tunnel barrier acts as a filter that passes forward-directed electrons.[19] These electrons have relatively little parallel momentum and, therefore, cannot couple with conduction valleys that lie far away from the IBZ center.

Au/Si Schottky diodes provided an early experimental example.[11] Since the conduction band minimum lies near the X point in the (100) direction, the (100) crystal orientation has states that lie at the IBZ center. With the (111) orientation, all states require nonzero parallel momentum. Therefore, BEES on the (100)-crystal orientation should yield a larger I_c than (111). However, experiment has shown repeatedly that the two orientations yield virtually the same spectra.[22]

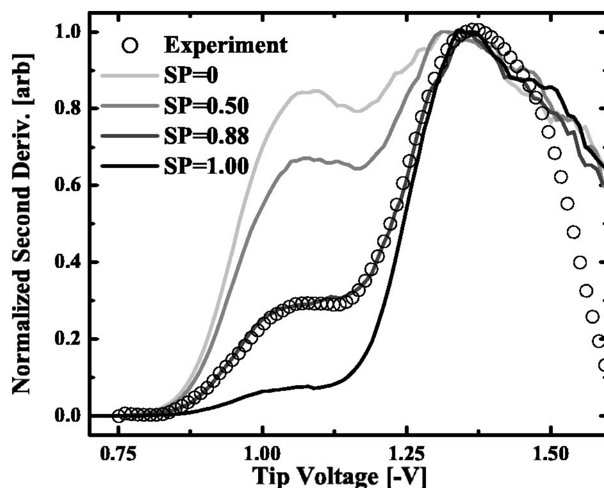

FIG. 11. Scattering probability (SP) changes the relative contributions of the Γ and L valleys in GaAs(100) to reflect experimental observations.

A similar difficulty arose in the interpretation of BEES spectra from Au/GaAs(100). Since the GaAs conduction band minimum (at Γ) lies at the zone center, one expects the contribution from this valley to dominate over any additional thresholds from higher conduction band minima, such as L, which in this crystal orientation lies near the perimeter of the IBZ. Contrary to this expectation, the contribution from L is typically 3–4 times stronger than Γ.

To explain this discrepancy, the standard planar tunneling model was modified to include s-wave scattering at the base metal-semiconductor interface.[23] Electrons initially highly forward-directed are scattered outside of the IBZ center where they can couple with states in the L valley. The scattering probability (SP), the probability that an individual electron is scattered out of the zone center, was determined to be approximately 0.9 by fitting the model to the data.

In Fig. 11 we show the effect of varying SP in our Monte Carlo algorithm. To model the scattering, SP×N electrons have their three-dimensional electron wave vectors randomly reoriented without altering the wave vector norm before coupling with semiconductor states. Instead of showing collector current spectra, we show the *second derivative* of the collector current (SD-BEES). This allows us to see the relative strengths of the thresholds in a clearer way. For comparison, experiment data are superimposed on the series of simulated spectra. We see that the simulation that most accurately reflects the relative strengths of the Γ and L valleys is the one for which SP is close to 0.9.

We also note the consistent discrepancy between the simulated and observed SD-BEES at high bias. While this is partly due to our neglect of electron-phonon scattering in the semiconductor at high electron energy, it is mostly due to the inadequacy of the spherical band model to accurately reflect the true band structure far from the conduction valley minimum.

VII. CONCLUSION

We have presented a detailed description of Monte Carlo calculations of hot-electron injection via tunnel-junction

emission for applications to modeling spectroscopies, such as metal-base transistor transfer characteristics and BEES. The method is straightforward but involves many subtleties that drastically increase the effectiveness and accuracy of the calculation. These improvements include reduction of phase space in the emitter Fermi sea to increase the effective sampling density and inclusion of integral methods and minimization to maintain constant tunnel current for BEEM spectroscopy simulation. After examination, this framework allows a simple analysis of the collector current dependence on emitter voltage bias near the Schottky threshold.

Although the computational difficulty of the Monte Carlo method increases when the path through phase space is extended and attaining high accuracy requires dense sampling, this technique is useful even with modest computational facilities. With greater computing power, *ad hoc* additions can be included to explore more sophisticated uses of hot-electron injection, such as quantum-mechanical transmission through buried heterostructures and interactions with phonons within the collector, leading to energy relaxation and momentum reorientation (reflection).

ACKNOWLEDGMENTS

The authors acknowledge support from the NSF under Contract No. ECS-9906047 and are grateful for help from N. Master, K.J. Russell, and R. Sheth. The work was supported in part by an ONR MURI subaward (S0149461) from the University of California at Santa Cruz.

*Present address: Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA. Electronic address: appelbaum@ee.udel.edu

[1] E. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford University Press, New York, 1989).
[2] I. Giaever, Phys. Rev. Lett. **5**, 147 (1960).
[3] R. Meservey and P. Tedrow, Phys. Rep. **238**, 174 (1994).
[4] J. Moodera, L. Kinder, T. Wong, and R. Meservey, Phys. Rev. Lett. **74**, 3273 (1995).
[5] J. Spratt, R. Schwarz, and W. Kane, Phys. Rev. Lett. **6**, 341 (1961).
[6] M. Heiblum, Solid-State Electron. **24**, 343 (1981).
[7] W. Kaiser and L. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[8] L. Bell and W. Kaiser, Phys. Rev. Lett. **61**, 2368 (1988).
[9] D. Smith and S. M. Kogan, Phys. Rev. B **54**, 10354 (1996).
[10] D. Smith, M. Kozhevnikov, E. Lee, and V. Narayanamurti, Phys. Rev. B **61**, 13914 (2000).
[11] L. Schowalter and E. Lee, Phys. Rev. B **43**, 9308 (1991).
[12] M. Ke, D. Westwood, C. Matthai, B. Richardson, and R. Williams, Phys. Rev. B **53**, 4845 (1996).
[13] A. Bauer, M. Cuberes, M. Prietsch, and G. Kaindl, Phys. Rev. Lett. **71**, 149 (1993).
[14] E. Lee, V. Narayanamurti, and D. Smith, Phys. Rev. B **55**, R16033 (1997).
[15] E. Lee, Phys. Rev. B **59**, 15332 (1999).
[16] P. de Andres, F. Garcia-Vidal, K. Reuter, and F. Flores, Prog. Surf. Sci. **66**, 3 (2001).
[17] I. Appelbaum, R. Sheth, I. Shalish, K. Russell, and V. Narayanamurti, Phys. Rev. B **67**, 155307 (2003).
[18] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **44**, 247 (1949).
[19] C. Mead, J. Appl. Phys. **32**, 646 (1961).
[20] J. Simmons, J. Appl. Phys. **34**, 1793 (1963).
[21] M. Prietsch and R. Ludeke, Phys. Rev. Lett. **66**, 2511 (1991).
[22] L. Bell, Phys. Rev. Lett. **77**, 3893 (1996).
[23] D. Smith, M. Kozhevnikov, E. Lee, and V. Narayanamurti, Phys. Rev. B **61**, 13914 (2000).