Directed inelastic hopping of electrons through metal-insulator-metal tunnel junctions

Yizi Xu,* D. Ephron, and M. R. Beasley
Department of Applied Physics, Stanford University, Stanford, California 94305
(Received 13 March 1995)

We have used the metal/amorphous silicon/metal tunnel junction as a model system to explore the role of localized states in electron transport through thin insulating layers. We measured the tunneling conductance as a function of temperature T, bias voltage V, and barrier thickness d. The data show marked deviations from the classical WKB tunneling theory in the limit of low T and V with d intermediate between the decay length in the barrier and the Mott variable range hopping length. The data are instead consistent with directed inelastic hopping along statistically rare but highly conductive "chains" of localized states. The most effective chains for a given set of conditions (T,V,d) contain a definite number of localized states, $N>1$, configured in a nearly optimal way in space and energy. The conductance of the lowest-order hopping channel (all chains with $N=2$) exhibits the characteristic voltage and temperature dependences $G_2^{\rm hop}(V)\propto V^{4/3}$, and $G_2^{\rm hop}(T)\propto T^{4/3}$, respectively, as predicted by theory. Higher-order channels ($N>2$) also conform to the theoretical predictions remarkably well. The physical nature of these highly conductive channels and their implications for conduction through thick tunnel barriers and thin dielectrics is discussed.

I. INTRODUCTION

Electronic conduction in noncrystalline insulating materials has been a subject of considerable interest because the electronic states near the Fermi level are generally localized. The electron wave functions decay exponentially over a distance known as the localization length, α^{-1}. The dominant mechanism for electronic conduction in bulk noncrystalline insulating solids at sufficiently low temperatures is hopping via these localized states. Mott predicted[1] that the typical length of a hop, the variable range hopping length $l_{\rm VRH}$, increases with decreasing temperature as $T^{-1/4}$. The conductivity is proportional to the probability of such a hop,

$$\sigma_{\rm VRH}\propto\exp(-2\alpha l_{\rm VRH})=\sigma_0\exp[-(T^*/T)^{1/4}]\ .$$

This unique temperature dependence, often referred to as Mott's $T^{-1/4}$ law, describes transport in numerous insulating and semiconducting materials over a wide range of temperatures.

In the opposite limit of conduction perpendicular to a thin amorphous layer not much thicker than the localization length, the existence of localized states is not expected to alter the conduction process significantly. Direct tunneling from one electrode to the other is the dominant conduction process. Experiments confirm these expectations in tunnel junctions with thin amorphous silicon barriers.[2,3] As the barrier thickness increases, however, the conductance shows temperature and bias dependences that are markedly different from the predictions of the WKB tunneling model. A straightforward application of this model using the average barrier height approximation, without properly including the effects of the localized states in the barrier, yielded an average barrier height of about 20 meV, which is unphysically low.[3] The authors of that study themselves pointed out the difficulty of reconciling such a low barrier height with other physical properties of the barrier material. The question then arises of how the presence of the localized states (LS's) influences the conductance when the sample size (in the direction of transport) is much greater than the localization length but smaller than $l_{\rm VRH}$. It is the onset of transport via these states—both elastic and inelastic—and its ultimate evolution into variable range hopping, that are the focus of this paper.

The observation of a crossover from direct to resonant tunneling behavior, manifested distinctly in the dependence of the conductance on the barrier thickness, demonstrated the role of the localized states in elastic tunneling.[4] The peaks due to transport via individual localized states were resolved, and their temperature dependence studied, yielding important information about the inelastic processes involved.[4,5] The importance of the inelastic effects was further recognized in Ref. 5. More recent work[6] has focused on a quantitative understanding of inelastic hopping via localized states in different thickness regimes. The dominant conduction mechanism crosses over as a function of barrier thickness from direct tunneling to resonant tunneling to directed hopping along quasi-one-dimensional chains of localized states with nearly optimal configurations and finally, in the bulk limit with $d\gg l_{\rm VRH}$, to variable range hopping.

Pollak and Hauser[7] first recognized theoretically the potential importance of hopping along special one-dimensional chains when the variable range hopping length becomes comparable to the sample size. Based on a series of experiments by Hauser and co-workers,[8] in which systematic deviations from Mott's $T^{-1/4}$ law were observed in the low-temperature conductance of samples with dimensions on the order of $l_{\rm VRH}$, they suggested that this regime favors hopping along chains containing localized states that are nearly equidistantly positioned across the sample.[7] They further argued qualitatively that, although these chains are rare, they are exponentially more conductive than the percolative hopping paths predicted by the VRH model for the bulk limit. The

model of Pollak and Hauser was further developed theoretically by two groups. Initial quantitative results were obtained for the temperature dependence (Tartakovskii *et al.*[9]) and the electric field dependence (Levin, Ruzin, and Shklovskii[10]) of the hopping conductance in the limit that the chains contain many localized states ($N \gg 1$).

Motivated by the amorphous silicon tunnel barrier studies,[4,5] Glazman and Matveev[11] proposed a microscopic model for hopping via two or more localized states forming optimal conduction chains. The model predicts a $T^{4/3}$ dependence of the conductance (for $eV \ll kT$) and a $V^{4/3}$ dependence of the conductance (for $kT \ll eV$) in the case $N=2$. Increasing temperature, bias voltage, or barrier thickness favors hopping along chains with more localized states ($N>2$), resulting in an increasingly nonlinear dependence of conductance on temperature and bias voltage. In this limit, the theory gives a quantitative basis to the original model of Pollak and Hauser. In this paper, we report results of our work on metal/amorphous silicon/metal tunnel junctions, which quantitatively confirm these more modern theories.

Hopping conduction at intermediate length scales has also been studied by Popovic, Fowler, and Washburn[12] in a two-dimensional electron-gas (2DEG) system formed in n-channel metal-oxide-semiconductor field-effect transistors (MOSFET's). Application of a gate voltage changes the density of carriers in the 2DEG and tunes the disorder of the system from weak to strong. Fluctuations in the conductance and interesting temperature dependences were observed in the strongly disordered regime, and were interpreted by the authors as a consequence of the conductance being dominated by hopping channels of the Glazman-Matveev type.

The rest of this paper is organized as follows. In Sec. II we review the relevant theoretical ideas that we will later use to analyze our experimental results. In Sec. III we give a brief account of our experimental procedures. In Sec. IV we focus on the experimental observation of the $T^{4/3}$ and $V^{4/3}$ dependences of the hopping conductance due to the $N=2$ hopping channel. The crossover to higher-order hopping channels ($N>2$) and to VRH-like behavior will also be discussed. Then we will mention briefly the low-temperature anomalies observed in the thin barrier regime. Finally, we offer some conclusions in Sec. V.

II. THEORETICAL OVERVIEW

A. Resonant tunneling via a single localized state

A one-dimensional model often captures the essential physics of a tunneling process. The transfer Hamiltonian method[13,14] is particularly well suited to deal with this situation. A term representing the barrier weakly couples two electrodes with free electron energy spectra. The presence of localized states in the barrier opens up the possibility of alternative mechanisms for transferring electrons between the electrodes other than direct tunneling.

The conductance due to an elastic electron transition from the left electrode, where the electron has initial energy ε, to the right electrode with final energy ε', via a localized state with energy ε_1, has been calculated by many authors[15,16] and is given by the expression

$$g(\varepsilon \rightarrow \varepsilon') = \frac{e^2}{2\pi\hbar} \frac{4\Gamma_L \Gamma_R}{(\varepsilon-\varepsilon_1)^2+\Gamma^2} \delta(\varepsilon-\varepsilon') \ . \tag{1}$$

The term $(e^2/2\pi\hbar)=4\times10^{-5}\ \Omega^{-1}$ on the right-hand side is the quantum of conductance. The δ function enforces energy conservation. The Lorentzian shape stems from the quantum mechanically coherent nature of the process. Γ/h is the inverse lifetime of an electron on the localized state, $\Gamma=\Gamma_L+\Gamma_R$, where[15]

$$\Gamma_{L(R)} = E_0 \exp(-2\alpha z_{L(R)}) \ , \tag{2}$$

for a 1D model. E_0 is a measure of the effective depth of the localized state, α^{-1} is the localization length, and $z_{L(R)}$ is the distance from the localized state to the left (right) electrode. The conductance (1) attains a maximum, $e^2/2\pi\hbar$, when $\Gamma_L=\Gamma_R$ and $\varepsilon_1=\varepsilon=\varepsilon'$. For a uniform barrier height, $\Gamma_L=\Gamma_R$ implies that $z_L=z_R=d/2$, where d is the barrier thickness. When $\alpha d \gg 1$, the width of this resonance is sharp, being confined in space to within α^{-1} of the center of the barrier, and in energy to within Γ of the energy of the resonant state. Thus a localized state will contribute to the zero-bias conductance of a junction only if its energy is within $\sim\Gamma$ of the Fermi energy.

When there are many localized states within the barrier, the total resonant conductance is given by the sum of the conductances through all of the resonant states. Assuming a uniform distribution of the localized states in space and in energy near the Fermi level with a density g, the integrated resonant tunneling conductance is simply

$$G_1^{\rm res} = \frac{\pi e^2}{\hbar} g S \alpha^{-1} E_0 e^{-\alpha d} = \hat{G}_1^{\rm res} e^{-\alpha d} \ , \tag{3}$$

where S is the area of the junction. The subscript 1 denotes that each resonant tunneling event occurs via a single localized state. This thickness dependence differs characteristically from that of direct tunneling,

$$G_0^{\rm dir} = \hat{G}_0^{\rm dir} e^{-2\alpha d} \ .$$

Note that for $g>0$, $G_1^{\rm res} > G_0^{\rm dir}$ for sufficiently large d. Equation (3) is obtained by integrating Eq. (1) over all possible configurations of LS's, and assumes that S is not so small that the details of the configurations are important. A simpler, more physically transparent approximation is to multiply the width of the maximal peak in Eq. (1) in space and energy, $S\alpha^{-1}\Gamma(z_L=z_R=d/2)$, by g to obtain the number of LS's within the width of the resonance, and then to multiply this quantity by the height of the peak, e^2/h.

B. Resonant tunneling via chains of two or more localized states

In the case of a thick barrier and a high density of localized states, the probability of an electron tunneling resonantly through two or more consecutive localized

states is non-negligible. Larkin and Matveev[15] calculated the conductance G_2^{res} for this process. They found that $G_2^{res} \propto \exp(-\alpha d)$, which has the same thickness dependence as G_1^{res}. The two localized states have to be aligned with each other in energy to within the width of the combined resonance, and also in space in order for the double resonance to form. Although elastic multiple-site resonant conduction channels should in principle predominate in the limit of very thick barriers at extremely low temperatures, their contribution to the conductance will in practice be masked by the phonon-assisted inelastic hopping processes described in Sec. II D.

C. Inelastic-scattering effects

Stone and Lee[17] addressed the role of inelastic processes in resonant transmission from a scattering point of view. They argued that inelastic scattering reduces the height of the resonant transmission peak, but conserves the total integrated transmission through the resonant state.

Microscopic theories which treat the dynamics of resonant tunneling quantum mechanically in the presence of the electron-phonon interaction (EPI) have been developed independently by Glazman and Shekhter[16] and Wingreen, Jacobsen, and Wilkins.[18] The main results are essentially identical, although different techniques were employed in the calculations. Though Wingreen, Jacobsen, and Wilkins present a more systematic and generalized treatment, the physical system in Ref. 16 is directly related to ours. The key parameter governing the strength of the EPI is the dimensionless quantity

$$\beta \equiv \frac{\Lambda^2}{(M v_s^2)(\hbar\omega_D)} .$$

Here Λ is the deformation potential, M is the ion mass, v_s is the speed of sound in the material, and ω_D is the Debye frequency. For later reference we relate M to the mass density $\rho = M q_D^3$, where $1/q_D$ is the interatomic spacing, $q_D = \omega_D / v_s$. The parameter β may then be rewritten as

$$\beta = \frac{\Lambda^2 (\hbar\omega_D)^2}{\hbar^3 \rho v_s^5} . \tag{4}$$

Glazman and Shekhter show[16] that the presence of the phonon field has two distinct effects. First, the lattice vibrations couple to the energy level of the localized state and broaden the resonance peak. Second, the resonant conduction redistributes itself into elastic part and inelastic parts, which represent tunneling in which the energy of the incident and final electron states differ. When the EPI is strong, i.e., $\beta \gg 1$, the transmission probability on resonance approximately equals the ratio of the elastic resonance width to the inelastic broadening: $\Gamma_e / \varepsilon_\Lambda$, where $\varepsilon_\Lambda \approx (\hbar\omega_D)\sqrt{\beta} \gg \hbar\omega_D$. If one regards ε_Λ / h as equivalent to some inelastic rate Γ_i, then the conclusion of Stone and Lee,[17] $\sigma_{peak}^{res} \approx \Gamma_e / \Gamma_i$, reemerges. When the EPI is weak, as it is in our system, the detailed calculation indicates that the inelastic resonant effects introduce a temperature-dependent correction to the resonant tunneling conductance

$$\frac{\delta G(T)}{G_1^{res}} \approx \beta \left[\frac{k_B T}{\hbar\omega_D} \right]^2 . \tag{5}$$

Here G_1^{res} is the temperature-independent resonant tunneling conductance given in Eq. (3) in the absence of the EPI.

Integrating the *temperature-independent* transmission terms for elastic resonant and direct tunneling over the thermally broadened Fermi surfaces of the electrodes at finite temperatures also leads to an increase of the conductance by an amount proportional to T^2 to first order.[19] These corrections are negligible in the experimental results presented in this paper.

D. Inelastic hopping via chains of two or more localized states

Employing a similar microscopic model, Glazman and Matveev[11] investigated the process of inelastic tunneling through two localized states. The model is shown schematically in Fig. 1. Each of the two localized states, labeled 1 and 2, is assumed to couple elastically to the

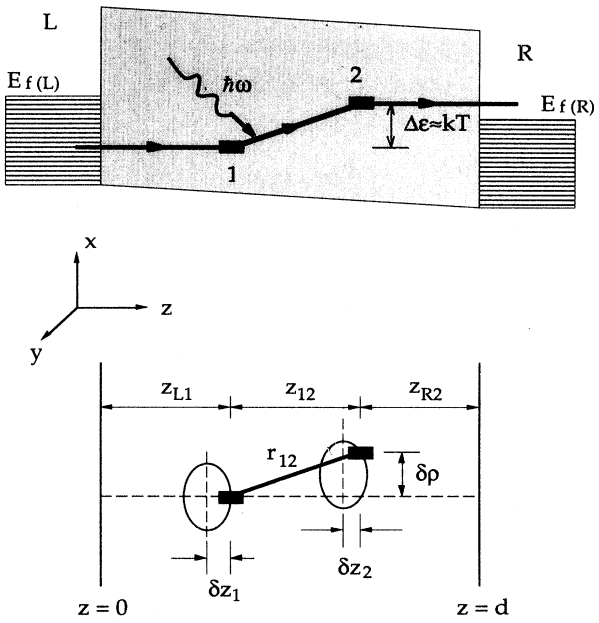

FIG. 1. Upper panel: a schematic energy vs position (z direction) diagram showing an electron tunneling from the left electrode (L) to the right electrode (R) through two localized states (LS's) labeled 1 and 2. The hop between the LS's is inelastic and is associated with the emission (or absorption) of a phonon. The most important pairs of the LS's are those within $k_B T$ of the Fermi level with energy separation $\Delta\varepsilon \approx k_B T$. Lower panel: a schematic representation of the positions (x, y, and z) of the two LS's forming a nearly optimal hopping chain of the upper panel. The first LS must be located within δz_1 of its ideal longitudinal position, but may be located anywhere laterally across the junction. The second LS should be within $\delta\rho$ of the first LS in the lateral direction, and within δz_2 of its ideal longitudinal position.

nearest electrodes only, labeled L and R. The coupling between the two localized states is assumed to be inelastic, involving interactions with phonons.

The authors use the Hamiltonian

$$H=\sum_k \varepsilon_k a_k^+ a_k+\sum_p \varepsilon_p a_p^+ a_p+\varepsilon_1 a_1^+ a_1+\varepsilon_2 a_2^+ a_2 +\sum_k (T_{k1}^* a_k^+ a_1+\text{c.c.})+\sum_p (T_{p2} a_p^+ a_2+\text{c.c.}) +(T_{12} a_1^+ a_2+\text{c.c.})+\sum_q \hbar\omega_q b_q^+ b_q +T_{12} a_1^+ a_2 \sum_q (\lambda_q b_q^+-\lambda_q^* b_q)+\text{c.c.} ,$$

where ε and a^+ are the energies and operators for the creation of electrons, and the indices k, p, 1, and 2 refer to the left electrode, right electrode, first LS, and second LS, respectively. The matrix elements T connect the states indicated by the subscripts, and b_q^+ creates a phonon with energy $\hbar\omega_q$; λ_q describes the electron-phonon coupling. Because the hybridization of the two LS's is negligible, transport is necessarily inelastic and the Hamiltonian reduces to a set of coupled kinetic equations, which the authors solved in certain limiting cases. They first considered *the low-bias limit, $eV \ll k_B T$*, and found the conductance due to hopping via a single pair of localized states with a fixed configuration

$$g(1;2)=\frac{e^2/\hbar}{2k_B T}\left[\frac{1}{\Gamma_{L1}}+\frac{1}{\Gamma_{R2}}+\frac{2}{\gamma}\right]^{-1} . \tag{6}$$

The factor of $k_B T$ in the denominator is the result of the expansion in the small parameter $eV/k_B T$. As the three constituent transitions are incoherent, their rates add in a serial fashion, a direct consequence of using the kinetic equation approach. Γ_{L1} and Γ_{R2} are the elastic transition rates between the left electrode and state 1, and between state 2 and the right electrode, respectively:

$$\Gamma_{L1}\approx E_0\exp(-2\alpha z_{L1}),\quad \Gamma_{R2}\approx E_0\exp(-2\alpha z_{R2}) . \tag{7}$$

Referring to the lower panel of Fig. 1, z_{L1} and z_{R2} are the distances between the two localized states and the respective electrodes. E_0 is the binding energy of the localized states, and may be taken to be the average barrier height.

γ is the transition rate between the two localized states, which are coupled inelastically. It is given by[11]

$$\gamma\approx\Delta\varepsilon\frac{\Lambda^2 E_0^2}{\hbar^3\rho v_s^5}\exp(-2\alpha r_{12}) , \tag{8}$$

where $\Delta\varepsilon$ is the energy difference between the two states, Λ is the average deformation potential, ρ is the mass density of the barrier material, v_s is the speed of sound, and r_{12} is the distance between the two localized states:

$$r_{12}=[z_{12}^2+(\delta\rho)^2]^{1/2} . \tag{9}$$

Here z_{12} is their separation along the axis normal to the plane of the junction, and $\delta\rho$ is their lateral separation. At this point it is convenient to introduce the dimensionless quantity

$$\lambda\equiv\frac{\Lambda^2 E_0^2}{\hbar^3\rho v_s^5} , \tag{10}$$

and to express γ in the form

$$\gamma\approx\lambda\left[\frac{\Delta\varepsilon}{E_0}\right][E_0\exp(-2\alpha r_{12})] , \tag{11}$$

in which the broadening of elastic transition width [cf. Eq. (7)] due to electron-phonon interaction is manifested.

Since the conductance in Eq. (6) is manifestly that of three resistors in series, each representing a tunneling or hopping transition whose rate is exponentially small in the separation between the states connected by the transition, the maximum is attained when all three rates are equal. This occurs if the two localized states lie on a single line perpendicular to the plane of the junction [i.e., $\delta\rho=0$ in Fig. 1] and (approximately) divide the barrier thickness d into equal parts [$\delta z_{1(2)}=0$ in Fig. 1]. Substituting into Eq. (6) yields

$$g_{\max}(1;2)\approx\frac{e^2/\hbar}{k_B T}\left[\lambda\frac{\Delta\varepsilon}{E_0}\right]^{1/3} E_0\exp\left[-\frac{2\alpha d}{3}\right] . \tag{12}$$

Equation (12) is a sharp maximum due to the exponential dependences of the transition rates on the coordinates of the LS's. Thus we can deduce heuristically the average zero-bias conductance due to hopping via two localized states, $G_2^{\text{hop}}(T)$, for a dense distribution of localized states by multiplying the maximum conductance for a single pair, Eq. (12), by the width of the maximum (in space and energy coordinates) and by g^2; the last two factors together give the probable number of pairs in the barrier with configurations within the spatial and energy widths of the maximum. Referring to Fig. 1, the width of the maximum with respect to the coordinates of the first state is α^{-1} in the z direction, S parallel to the plane of the junction, and $k_B T$ in energy (in order to contribute to the zero-bias conductance). The location of the first state fixes the position of the chain, so the second state contributes $\alpha^{-1}(\delta\rho)^2 k_B T$, where $(\delta\rho)^2=(\alpha^{-1}d)$. Collecting terms yields

$$G_2^{\text{hop}}(T)=\nu_{2,T}(gS\alpha^{-1}k_B T)(g\alpha^{-2}d\,k_B T)\left[\frac{e^2/\hbar}{k_B T}\right] \times\left[\lambda\frac{k_B T}{E_0}\right]^{1/3} E_0\exp\left[-\frac{2\alpha d}{3}\right] \propto(k_B T)^{4/3} . \tag{13}$$

$\Delta\varepsilon$ has been set equal to $k_B T$, and we have made the implicit assumption that the dispersion of phonon modes is linear—appropriate for low-energy acoustic phonons. The factor of $\exp(-2\alpha d/3)$ reflects the fact that the electron's traversal of the barrier requires three steps. The actual integration over the coordinates and energies of the two states gives[11] $\nu_{2,T}\approx 50$. The subscript 2 denotes the $N=2$ channel, and the subscript T denotes the limit $k_B T \gg eV$.

Equation (13) predicts that the hopping conductance

due to the $N=2$ channel has the unique temperature dependence $(k_BT)^{4/3}$. Note that this specific power law is the result of three different factors: the first-order expansion in the small parameter eV/k_BT, which contributes T^{-1}; the integration over a linear phonon dispersion combined with the proportionality of the phonon coupling strength to the phonon energy, together contributing $T^{1/3}$; and the energy requirements that the first localized state be within $\sim k_BT$ of the Fermi level, so that electrons are available to tunnel onto it, and that the second state be within $\sim k_BT$ of the first, so that phonons are available to couple them, contributing T^2. Thus the deduction of the power law $T^{4/3}$ on physical grounds is robust and independent of many of the approximations inherent in the theory. For example, the assumption of a uniform barrier height and the approximate forms of the coupling constants in Eq. (7) affect the prefactor in Eq. (13), but not the prediction that $G_2^{\rm hop}\sim T^{4/3}$.

In any actual conductance measurement this temperature dependence will show up as an addition to the temperature-independent resonant tunneling conductance. The relative importance of the two terms is manifested in their ratio

$$\frac{G_2^{\rm hop}(T)}{G_1^{\rm res}}=\nu_{2,T}(g\alpha^{-2}dk_BT)\left[\lambda\frac{k_BT}{E_0}\right]^{1/3}\exp\left[\frac{\alpha d}{3}\right]\propto(k_BT)^{4/3}\ . \tag{14}$$

The correction is more pronounced at higher temperatures and for thicker barriers, unlike the correction given in Eq. (5), which is thickness independent. The inelastic hopping conductance becomes exponentially larger than the elastic tunneling conductance as the barrier thickness increases. The exponential advantage conferred by breaking the barrier traversal into three rather than two segments compensates for the statistical rarity of realizing such nearly optimal configurations of chains involving two localized states. Physically, the reason that these inelastic chains dominate the elastic resonant chains with the same number of localized states ($N>1$) at finite temperature is that the probability of realizing the inelastic channels is greatly enhanced by the significantly relaxed requirement on the alignment of the states in energy. This freedom in configuration space more than compensates for the coupling to the phonon field.

The above statements are correct, and the actual hopping conductance will be given by Eq. (13), only when the junction area S is sufficiently large that the junction is in the self-averaging regime. If the probability of forming a single chain of states within the given junction area is negligibly small, then the enhanced conductance of such a chain cannot compensate for its nonexistence. A quantitative estimate of this requirement follows from dividing the average conductance of the $N=2$ hopping channel, Eq. (13), by the maximum conductance of a single ideal hopping chain, Eq. (12), yielding an expression for the effective number of chains:

$$N_{2,T}^{\rm eff}\approx\nu_{2,T}(S\alpha^{-1}g\,k_BT)(\alpha^{-2}dg\,k_BT)\ . \tag{15}$$

When this number is small, the hopping conductance is no longer a self-averaging quantity, and strong sample-dependent fluctuations in the conductance should appear. Note, however, that Eq. (15) is only a rough guideline because it calculates the equivalent number of *ideally* configured chains (i.e., the minimum number of chains) necessary to account for the averaged conductance of *all* possible configurations. Thus the actual crossover from the self-averaging to the fluctuating regime occurs only when $N_{2,T}^{\rm eff}$ is less than some constant less than 1.

Taking the physical reasoning after Eq. (14) one step further, we can see that as the temperature or barrier thickness further increases, the dominant conduction channel crosses over to hopping along chains containing a progressively larger number of localized states ($N>2$) that match a nearly optimal configuration analogous to the case of $N=2$, though they are statistically even more rare. Quantitatively, for the averaged hopping conductance of the chains consisting of exactly N localized states, Glazman and Matveev estimated[11]

$$\frac{G_N^{\rm hop}(T)}{G_1^{\rm res}}=\nu_{N,T}(g\alpha^{-2}d\,k_BT)^{N-1}\left[\lambda\frac{k_BT}{E_0}\right]^{(N-1)/(N+1)}\times\exp\left[\frac{N-1}{N+1}\alpha d\right]\ . \tag{16}$$

The essential features of this equation are the following.

(a) A distinctive power-law temperature dependence $(k_BT)^{N-2/(N+1)}$.

(b) An exponential thickness dependence, $\exp[(N-1)(\alpha d)/(N+1)]$, that greatly favors chains in which N is large.

(c) A factor $\nu_{N,T}(g\alpha^{-2}dk_BT)^{N-1}$ which, roughly speaking, is the ratio of the number of nearly optimal chains consisting of N localized states per unit area to the number of LS's per unit area that contribute optimally to the resonant conductance. This factor becomes diminishingly small as N increases. Thus points (b) and (c) compete to determine the value of N that dominates for a given set of conditions T, V, and d.

(d) A coefficient $\nu_{N,T}$, that according to Ref. 11 varies approximately as N^{2N}, reflecting the increasing spatial flexibility of chains consisting of many localized states. Note that this estimate for the dependence of $\nu_{N,T}$ on N neglects factors of order unity that may multiply the base value of N. These small factors accumulate quickly. Also, the relative values of the sequence of prefactors, $\{\nu_{N,T}\}$, as well as their absolute values, are sensitive to the approximations discussed preceding Eq. (14).

The total conductance at any temperature is the sum of the contributions from all the channels:

$$G_{\sum}(T)=G_0^{\rm dir}+G_1^{\rm res}+\sum_{N=2}G_N^{\rm hop}(T)\ . \tag{17}$$

We use the term "channel" to refer to the total contribution to the conductance of all chains consisting of a given number of LS's. For thicker barriers at low temperatures the hopping conductance is usually dominated by the first few terms in Eq. (17). With increasing temperature, the dominant channel shifts from $N=2$ to $N=3$ to $N=4$, etc. Reference 11 suggested the criterion that when

$N \geq (\alpha d)^{1/3}$, N is more conveniently treated as a continuous variable. Making this approximation recovers the results of Tartakovskii *et al.*,[9] which will be discussed in Sec. II E.

The second limiting case considered in Ref. 11 is *the high-bias and low-temperature limit* $eV \gg k_B T$. The authors gave an expression similar to that of Eq. (14) for the voltage-dependent correction to the conductance due to the $N=2$ hopping channel:

$$\frac{G_2^{\rm hop}(V)}{G_1^{\rm res}} = \nu_{2,V}(g\alpha^{-2}deV)\left[\lambda\frac{eV}{E_0}\right]^{1/3}\exp\left[\frac{\alpha d}{3}\right] \propto V^{4/3} . \quad (18)$$

The equivalent role of eV and $k_B T$ in Eqs. (14) and (18) is not surprising, since these factors set the energy window for integration in the appropriate limits. The coefficient $\nu_{2,V}$ is, however, much smaller than $\nu_{2,T}$. In fact $\nu_{2,V}$ is only of order unity, as can be shown by writing out the intermediate steps leading up to Eq. (3.11) in Ref. 11. The criterion for self-averaging is

$$N_{2,V}^{\rm eff} \approx \nu_{2,V}(S\alpha^{-1}geV)(\alpha^{-2}dg\,eV) > 1 . \quad (19)$$

An expression analogous to Eq. (16) exists for the bias voltage dependence of the N-state hopping channel, $G_N^{\rm hop}(V)$, in the limit $eV \gg k_B T$. The coefficients $\{\nu_{N,V}\}$, however, vary with N much more slowly—approximately as $(N-1)^{(N-1)}$, as compared with N^{2N} for $\{\nu_{N,T}\}$. The physical reason for this difference is that, whereas in the low-bias limit an amount of energy $\sim k_B T$ is available for each hop, in the low-temperature limit the total energy available for inelastic hops, eV, apportions itself among $N-1$ hops.

E. Crossover from Glazman-Matveev model to Pollak-Hauser model

The computational scheme in Ref. 4 is most convenient at low temperatures and low bias, where the dominant hopping processes involve only a small number of localized states, $N \leq (\alpha d)^{1/3}$. At higher temperatures and bias the optimal N is most readily approximated as a continuous variable of temperature or voltage, depending on the appropriate limit. Reference 9 gives the optimal N as a function of temperature:

$$\tilde{N}_T = (2\alpha d/\delta_T)^{1/2} , \quad (20)$$

where δ_T is the solution of the transcendental equation

$$\delta_T = \ln\left[\frac{0.0365\times\delta_T}{g\alpha^{-1}d^2 k_B T}\right] . \quad (21)$$

The corresponding hopping conductance in this limit is[9,11]

$$G_{\rm Tart}(T) \propto \exp(-2\alpha d/\tilde{N}_T) = \exp[-2(2\alpha d\delta_T)^{1/2}] . \quad (22)$$

Tartakovskii *et al.*[9] obtained this expression in a more rigorous treatment of the one-dimensional hopping model originally proposed by Pollak and Hauser.[7] Glazman and Matveev[11] show that their model approaches Eq. (22) asymptotically as $N \Longrightarrow \infty$.

Levin, Ruzin, and Shklovskii[10] obtained an expression analogous to Eq. (22) for $eV \gg k_B T$:

$$G_{\rm Levin}(V) \propto \exp(-2\alpha d/\tilde{N}_V) = \exp[-2(2\alpha d\delta_V)^{1/2}] , \quad (23)$$

where δ_V is given by

$$\delta_V = \ln\left[\frac{0.234}{g\alpha^{-2}deV}\right] . \quad (24)$$

Equations (22) and (23) have a limited range of applicability in terms of sample thickness, temperature, and bias. Consider, for example, that Eqs. (21) and (24) have meaningless negative solutions for δ_T and δ_V if the barrier thickness is too large or if the temperature or bias is too high. Under these extremes the model of highly directed quasi-1D transport breaks down and a percolative model, such as variable range hopping, becomes more appropriate. However, exactly how essentially directed transport crosses over to diffusive variable range hopping in the bulk limit depends on the dimensionality of the system[20] and is not entirely clear at present.

F. Variable range hopping—bulk limit

Mott[1] recognized that two important physical considerations determine the bulk hopping conductance. A hop has to be short enough so as to have a nonvanishing overlap between wave functions on the two localized states involved. On the other hand, it has to be long enough to find a state not too different in energy. These two competing considerations fix a typical hopping length $l_{\rm VRH}$, which is a function of temperature. For a uniform distribution of localized states it was found[1] that $l_{\rm VRH}$ varies as $T^{-1/4}$. Consequently the variable range hopping conductance acquires the form

$$G_{\rm VRH}(T) = \hat{G}_0 \exp[-(T^*/T)^{1/4}] , \quad (25)$$

where T^* is given by

$$k_B T^* = 23/g\alpha^{-3} . \quad (26)$$

The constant 23 is a result of numerical calculations within the context of percolation theory.[21]

Shklovskii[22] proposed an analogous current-field relation for hopping in the limit of high bias,

$$J(E) = \hat{J}_0 \exp[-(E^*/E)^{1/4}] , \quad (27)$$

where $J(E)$ is the current density as a function of the local electrical field E, which we assume to be uniform for relatively thick barriers, $E = V/d$. E^* in Eq. (27) is a characteristic electrical field which is related to the characteristic temperature T^* in Eq. (26),

$$E^* = k_B T^*/e\alpha^{-1} , \quad (28)$$

where e is the absolute value of the charge of the electron. The validity of Eq. (27) may be quite limited, how-

ever. Under high bias the width of the impurity band available for hopping increases. The assumption of a uniform distribution of localized states in energy is therefore no longer valid. Also, high bias changes the shape of the barrier and reduces the average barrier height, increasing the local value of the localization length. Moreover, when the bias becomes comparable to the band gap, entirely different processes may contribute to the conduction. Despite these complications, Eq. (27) has a conceptual value and even some utility.

III. EXPERIMENTAL PROCEDURE

To make our junctions we begin by depositing metal/a-Si/metal trilayer structures *in situ* using electron-beam evaporation in a vacuum chamber with a base pressure of 2×10^{-7} Torr. The chamber has three independent electron-beam sources; each is feedback controlled to maintain the desired evaporation rate to within 0.1 Å/s ($\sim$1%).[23] The metal electrodes are either niobium or molybdenum. Using molybdenum allows us to measure the barrier conductance down to 1.4 K free of complications due to the onset of superconductivity in the electrodes. Single-crystal sapphire from Union Carbide is used as the substrate. These substrates are degreased and etched according to standard procedures and are immediately mounted into the evaporator. The substrates are maintained at 700 °C during the base electrode deposition from the central e-beam source, resulting in smooth, highly uniform, high-quality films with a residual resistivity ratio (RRR) greater than 60. After the samples cool to below 100 °C, coevaporation from the two peripheral e-beam sources, symmetrically situated with respect to the substrate holder, is carried out to form uniform, pinhole free amorphous silicon barriers across the entire film area. Two shutters are manipulated to produce samples with a range of known barrier thicknesses. The counter electrodes are then deposited at room temperature. The bottom electrodes are generally 2500 Å thick. The barrier thicknesses range from 20 to 1500 Å. The counter electrode is usually 500 Å thick. Additional details of the sample deposition and characterization can be found in Refs. 24 and 25.

Tunnel junctions, with areas varying from 4×4 to 90×90 μm^2, are defined using standard photolithography. For niobium electrodes we use a variant of the selected-area niobium annodization process (SNAP) to isolate the junction.[26] For molybdenum electrodes we use wet chemical etching followed by the deposition of a protecting 50 Å insulating layer of SiO_2 and then a protective insulating layer of photoresist. Deposition of a metal contact layer (either Ti/Au, Sn/Nb, or Cu/Mo) follows. Prior to this final deposition, a low-energy ion-beam source is used to clean the exposed surface area *in situ*. When finished each $1/4\times1/4$-in.^{2} chip has four junctions sharing a common ground plane. Four probe measurements of junctions on the same chip yield nearly identical I-V characteristics for junctions with areas as large as 90×90 μm^2, attesting to the extremely high degree of uniformity and pinhole-free nature of the tunnel barriers.

For junctions with resistance less than 100 kΩ at 4.2 K, ac techniques are used to measure the zero-bias resistance as a function of temperature. ac excitation voltages are kept below 20-μV rms, and the junction response is lock-in detected. For junctions with higher resistances, the roll-off frequency of the system (junction plus leads) becomes too low for this ac technique to be practical. We then rely on dc techniques alone.

I-V characteristics are measured using a battery-powered voltage source and an electrometer (Keithley 617), when appropriate. With this system we can achieve a current resolution of 0.1 pA, allowing us to measure junction resistances that span more than ten orders of magnitude. In all cases, the effects of the leads and the spreading resistance of the electrodes are carefully accounted for.

IV. EXPERIMENTAL RESULTS

A. Overview

In Fig. 2 we show the zero-bias tunneling conductance plotted on a logarithmic scale versus barrier thickness for samples with molybdenum electrodes (closed symbols) and niobium electrodes (open symbols and the cross—different symbols indicate samples made in different evaporator runs). For junctions with molybdenum electrodes, this is the conductance measured at or below 4.2 K and at very low bias where the I-V curve is Ohmic.

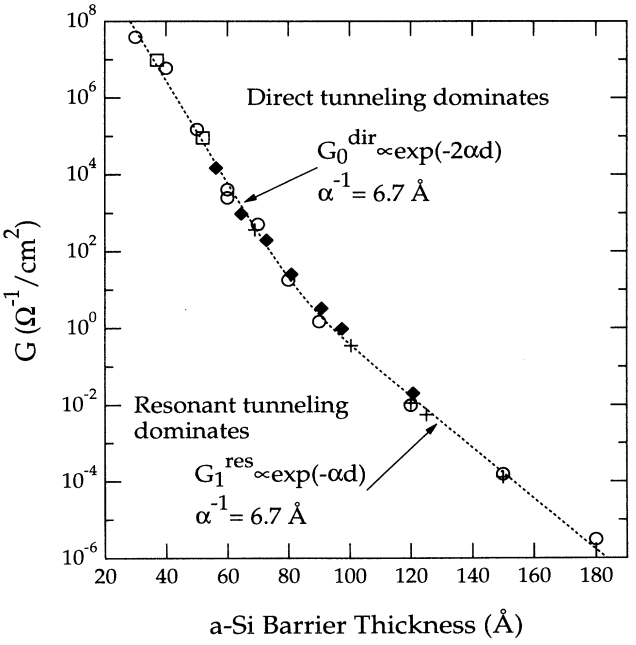

FIG. 2. Zero-bias tunneling conductance as a function of barrier thickness at 4.2 K for three sets of niobium junctions (open symbols and the cross) and one set of molybdenum junctions (closed diamond). The exponential decrease is a signature of tunneling. The change of the slope by a factor of 2 indicates the crossover from predominantly direct to predominantly resonant tunneling. The fit is to the form of Eq. (29), from which we infer the localization length $\alpha^{-1}=6.7$ Å.

For junctions with niobium electrodes, this is the conductance measured immediately above the superconducting gap of the electrodes (3 mV) in the case of thin barriers, e.g., $d < 80$ Å. For thicker barriers, the nonlinear effects are already quite noticeable even at 3 mV. However, in this case there is also substantial smearing of the gap structure so that the *I-V* characteristics at lower biases are practically Ohmic and are used instead.

The exponential decay of the Ohmic conductance with increasing barrier thickness, as demonstrated in Fig. 2, is a signature of quantum-mechanical tunneling. The crossover from predominantly direct to predominantly resonant tunneling causes the factor of 2 change in the slope.[4] The dashed line is a least-square fit to the data by the form

$$G^{\text{elastic}} = \hat{G}_0^{\text{dir}} e^{-2\alpha d} + \hat{G}_1^{\text{res}} e^{-\alpha d}/\alpha d \ , \tag{29}$$

where $\hat{G}_0^{\text{dir}}$, $\hat{G}_1^{\text{res}}$, and α are free parameters. The factor of αd in the denominator of the second term is appropriate for tunneling via a LS in three dimensions. From the fit we obtain $\alpha^{-1} = 6.7$ Å for the localization length. The prefactors agree to within a factor of 2 with their expected values.[15,24]

At higher temperatures phonon-assisted inelastic processes quickly dominate the tunneling conductance, resulting in a temperature-dependent conductance and nonlinear *IV*'s. Figure 3 shows the conductance on a logarithmic scale plotted versus $T^{-1/4}$ for junctions with different thicknesses. The behavior of the two thickest junctions ($d = 1550$ and 290 Å) follows closely the prediction of the VRH theory. Fitting the data of both samples using Eq. (25) in the region where they are linear on this plot yields a value of 1.24×10^8 K for T^*, and from Eq. (26) the density of the LS's in the amorphous silicon barrier $g = 8 \times 10^{18}$ eV^{-1}cm^{-3}. In the opposite limit, the junction with the 62-Å barrier shows little temperature dependence at low temperatures (apart from the structure due to the superconducting gap), indicating essentially elastic tunneling with a minuscule contribution from inelastic processes. However, the temperature-independent region diminishes quickly as the thickness rises above 70 Å. Moreover, the onset of nonlinearities becomes more precipitous with increasing d. The conductance for junctions in this intermediate thickness range conforms to neither limit, although at high temperatures the data seem asymptotically to approach VRH behavior. It is this regime in which the hopping conductance exhibits the more complicated behavior that is the focus of Secs. IV B and IV C.

Complementing Fig. 3, Fig. 4 shows the low-temperature conductance as a function of bias voltage for a series of tunnel junctions with barrier thicknesses ranging from about 50 up to 180 Å. As d increases, the Ohmic region quickly disappears and the nonlinearities become more pronounced.

B. Voltage dependence of the hopping conductance ($eV \gg k_B T$)

1. Observation of lowest-order hopping channels

Figures 5 and 6 show the conductance versus bias voltage up to 10 and 100 mV, respectively, for a junction with molybdenum electrodes, a barrier thickness $d = 98$ Å, and an area $S = 50 \times 50$ μm^2. The data were taken at 4.2 K, corresponding to a thermal energy of about 0.3 meV, so $eV \gg kT$ except very close to the origin. We choose to study this junction in detail because the

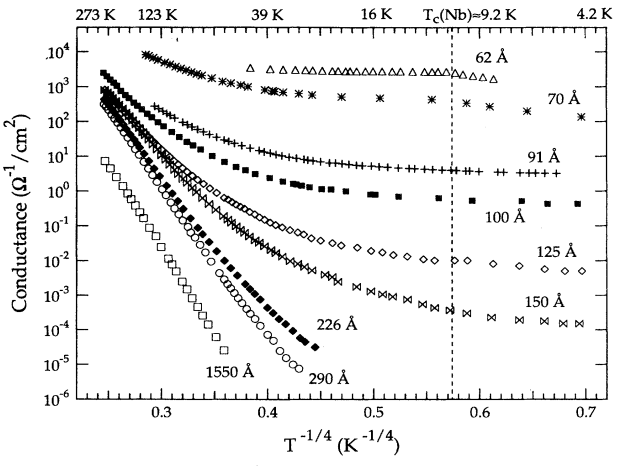

FIG. 3. Conductance vs temperature for junctions with various barrier thicknesses. The thick barrier limit is well described by the variable range hopping theory. The thin barrier limit is dominated by (temperature independent) elastic tunneling. The intermediate thickness range is the focus of Secs. IV B and IV C.

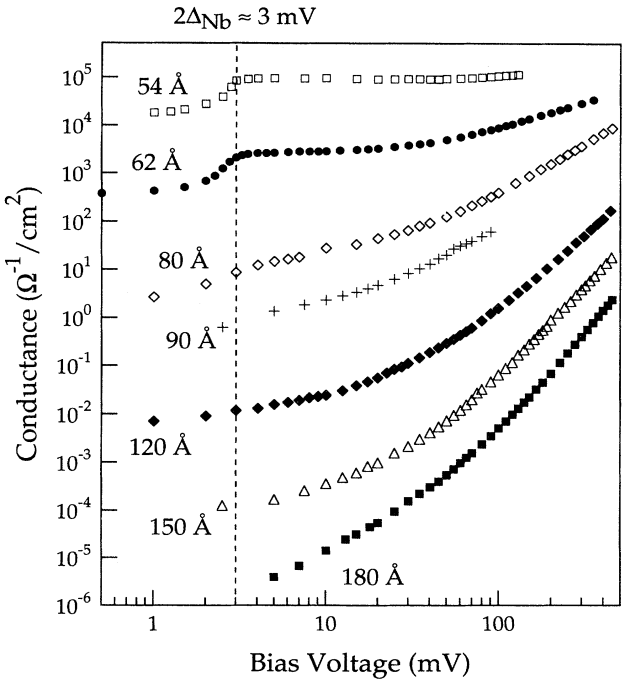

FIG. 4. Conductance vs bias for a series of junctions with niobium electrodes. Note that the threshold voltage for the onset of nonlinearity decreases with increasing barrier thickness, and the voltage dependence becomes increasingly nonlinear as the bias increases.

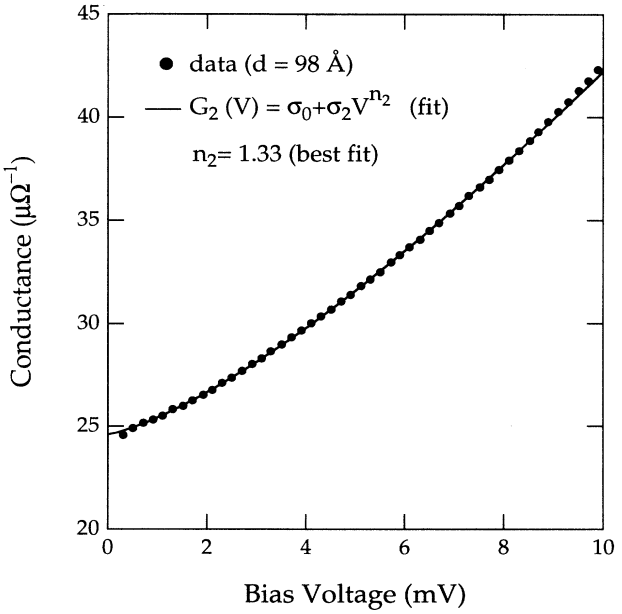

FIG. 5. Conductance vs bias for a $d=98$-Å sample. The solid line is the best fit to the form indicated, and gives 1.33 for the power exponent, indicating inelastic hopping via two LS's ($N=2$).

molybdenum electrodes remain normal at 4.2 K and because it is in the thickness regime where resonant tunneling begins to dominate the direct tunneling (cf. Fig. 2). Thus the influence of localized states, especially the lowest-order inelastic effects due to hopping, should be readily observable.

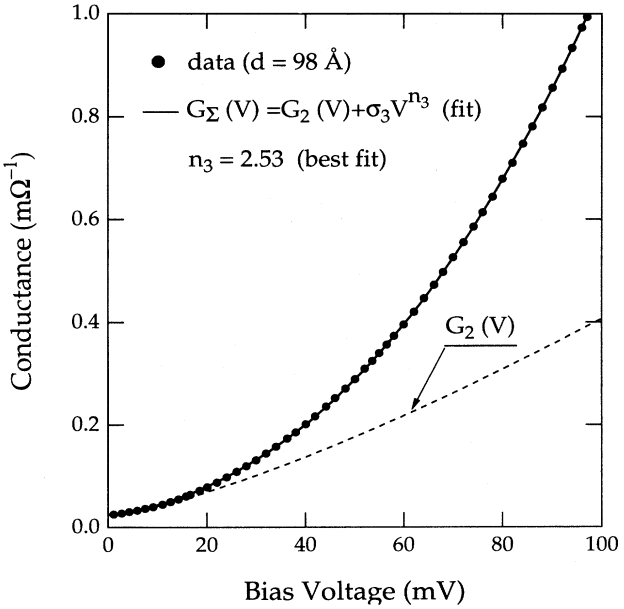

FIG. 6. Conductance vs bias for the same sample in Fig. 5 up to 100-mV bias. The dashed line is the fitting curve from Fig. 5 extrapolated to higher bias. The further increase of the conductance above the dashed line is due to the $N=3$ hopping channel. The solid line is a fit to the difference between the data and the dashed line.

We expect that the conductance at low bias consists of a bias-independent term, which includes the direct and resonant tunneling contributions, and a bias-dependent term to account for the onset of hopping processes. Considering the theoretical prediction that the hopping term varies as $V^{4/3}$ in the limit $eV \gg kT$, we fit the data to the form

$$G_2(V)=\sigma_0+\sigma_2 V^{n_2} , \tag{30}$$

where σ_0, σ_2, and n_2 are free parameters. The best fit in Fig. 5 yields the values

$$\sigma_0=2.46\times10^{-5}\ \Omega^{-1},\quad \sigma_2/\sigma_0=334V^{-n_2} ,$$
$$\text{and} \tag{31}$$
$$n_2=1.33 .$$

We see that the conductance increases as $V^{4/3}$, in precise agreement with the prediction of Eq. (18), up to a bias of about 9 mV; above this voltage, deviations become apparent. One may question the objectiveness of this procedure, since the value of the parameter n_2 is clearly sensitive to the bias range chosen for the fitting. For example, if one fits the data with the same functional form, Eq. (30), but up to 15 mV, then a larger value of $n_2(=1.5)$ results. However, using this larger n_2 produces a poorer fit to the data in the lower-bias range. Also, we found that the value of n_2 is *not* sensitive to the range chosen for the fit when restricted *below* 9 mV, and its variation reflects merely the noise inherent in fitting to a small number of points. Moreover, the theory predicts that, as bias voltage increases, hopping chains with more than two localized states play a more important role. Thus a self-consistent procedure should examine the data over a wider bias range and account for the higher-order processes. Specifically, to Eq. (30) we added a term of the form $\sigma_3 V^{n_3}$. Keeping σ_0, σ_2, and n_2 fixed at the values listed in Eq. (31), and allowing σ_3 and the exponent n_3 to vary as free parameters, we fit the conductance data from 0 to 100 mV, and found that the best fit for n_3 is 2.53. This is very close to the power exponent $\frac{5}{2}$ predicted by the theory for the bias dependence of the channel with $N=3$ localized states (and hence our subsequent choice of notation).

The solid line in Fig. 6 shows the best fit with the addition of the $N=3$ term. The dashed line shows the fit to the first two terms only, extrapolated past 10 mV. (Thus the second fit uses a single term, $\sigma_3 V^{n_3}$, to fit to the difference between the data and the dashed line.) The near-perfect agreement between the data and the fits, the conformance of the free parameters n_2 and n_3 to the theoretically predicted values, and the entirely self-consistent procedure prove beyond reasonable doubt that the non-Ohmic conduction at low bias is attributable to hopping through the $N=2$ channel at first, and then additionally through the $N=3$ channel as the voltage increases. A similar analysis on numerous other samples with different barrier thicknesses yields nearly identical values for the power exponents of the first two hopping

channels, $n_2=1.33$ and $n_3=2.5$, and comparable values for the prefactors.

2. *Evaluation of prefactors $\nu_{2,V}$ and $\nu_{3,V}$*

In addition to the power exponents, the coefficients of the $V^{4/3}$ and $V^{5/2}$ terms also contain valuable information. The quality of our data permits a quantitative comparison of these coefficients with the theory. To proceed, we note that for $d=98$ Å, resonant tunneling accounts for about 91% of the total conductance, as inferred from Fig. 2, whereas direct tunneling accounts for about 9%. Thus, $\sigma_0^{\rm res}\approx 0.9\sigma_0$.

We can calculate $\nu_{2,V}$ and $\nu_{3,V}$ by comparing Eqs. (31) and (18). The values of d, α^{-1}, and g are already known from independent measurements (see Sec. IV A). Bending[24] thoroughly investigated the average barrier height and found $E_0\approx 0.3$ eV. To a good approximation[27,28] we take the values of the material parameters in the definition of λ [Eq. (10)] to be equal to their values in crystalline silicon, as summarized in Table I. Substitution yields $\nu_{2,V}\approx 2.7$ and $\nu_{3,V}\approx 35$.

For completeness, we estimate the effective number of optimal hopping chains with $N=2$ states, using Eq. (19). At $V=5$ mV, we find $N_{2,V}^{\rm eff}\approx 30$. For the $N=3$ chains at $V=25$ mV we find $N_{3,V}^{\rm eff}\approx 6$. Therefore, these junctions are suitable for observing the self-averaged hopping conductance of the lowest-order channels. In fact, all four junctions sharing this substrate exhibit nearly identical I-V characteristics in the bias range studied, without any sample-to-sample fluctuations. This exercise indicates, however, that if the junction areas were reduced to the order of 10×10 μm^2, we should expect to begin to see deviations from the self-averaging behavior.

3. *Observation of higher-order hopping channels*

Now that we have a quantitative understanding of the hopping conductance in the relatively low-bias range in the framework of the model of Ref. 11, we are in a better position to analyze the data at higher bias, where hopping channels with more localized states are expected to become progressively more important, resulting in increasingly nonlinear conductances. To study this regime experimentally, however, the junction analyzed in Sec. IV B 1 turned out to be unsuitable. The resistances of that junction and its sister samples drop quickly as the bias increases above 100 mV because of their large areas (50×50 μm^2), and become comparable to that of the metal contacts. The four-point configuration no longer yields the true junction resistance under these conditions.[29]

We therefore chose junctions with niobium electrodes, which remain superconducting at 4.2 K, to study quantitatively the voltage dependence of the hopping conductance over a much wider bias range, thus extending our comparison with the theory to include the higher-bias range. The areas of these junctions range from 4×4 μm^2 to 8×8 μm^2. In Fig. 7 we show conductance versus bias voltage up to 500 mV for four junctions, of which two (the upper two traces) were deposited in one vacuum deposition run, whereas the other two (the lower two traces) are from a different run. The anomaly in the uppermost trace reflects the superconducting gap in the electrodes.

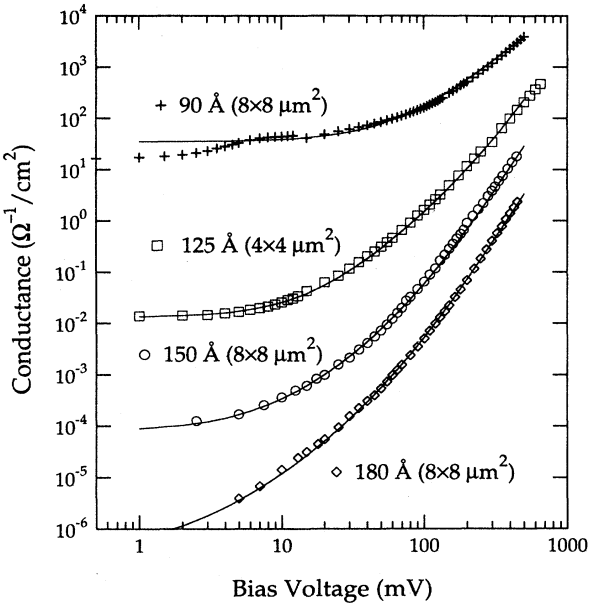

FIG. 7. Fits to the data for four samples with the prefactors of the various power terms treated as fitting parameters. The coefficients $\{\nu_{N,V}\}$ for the various channels that emerge from the fit are listed in Table II.

We fit the total hopping conductance of all four junctions according to the following procedure (solid lines). First, we determine the resonant tunneling conductance from the zero-bias conductance. Second, we write down the theoretical ratio of the hopping conductance of each channel to the resonant tunneling conductance using Eq. (16) with k_BT replaced by eV, and $\nu_{N,T}$ replaced by $\nu_{N,V}$. Third, we sum the contributions from the different channels up to a maximum number, $N_{\rm max}$, that is appropriate for the thickness of each individual junction. The introduction of this upper cutoff is necessary for a finite-sized junction. $N_{\rm max}$ is calculated from the criterion that, when the probability of formation of a chain containing N localized states is much smaller than unity, then the most likely situation is that there will be no such chains. The exponential term in Eq. (16) cannot compensate for a prefactor of exactly zero, and thus channels with $N>N_{\rm max}$ will not contribute to the conductance of junctions with a finite area. Finally, we fit all four curves using the materials parameters listed in Table I and allowing $\nu_{2,V}$, $\nu_{3,V}$, $\nu_{4,V}$, and $\nu_{5,V}$ to vary as free parameters. (Here the powers of V are fixed at their predicted values.) The resulting best fits are given in Table II.

The agreement between the experimental data and the total fit is remarkable, considering the approximations in the theory, the range of data covered—nearly ten orders of magnitude in conductance—and the similarity and reasonableness of the values of the various coefficients found. As discussed in Sec. II D, the value of $\nu_{2,V}$ should be of the order of 1, and we expect $\nu_{N,V}$ to vary approxi-

TABLE I. List of amorphous silicon-specific materials parameters used in this work.

Quantity	Symbol	Units	Value
density of localized states	g	$eV^{-1}cm^3$	8×10^{18}
localization length	α^{-1}	Å	6.7
deformation potential	Λ	eV	2.0
average barrier height	E_0	eV	0.3
mass density	ρ	$g\,cm^{-3}$	2.33
speed of sound	v_s	$cm\,s^{-1}$	6.6×10^5
reduced e-p coupling const.	λ	[dimensionless]	7.2×10^3

mately as $(N-1)^{(N-1)}$. The theoretical estimates, tabulated in the last column of Table II, coincide with the experimental values surprisingly well. The success of the fitting is actually somewhat puzzling since the (probably oversimplified) method of determining the crossover from the self-averaging to the fluctuation regime discussed previously would suggest that these small junctions should be in or near the fluctuation regime. The variation of the values in Table II may reflect this point. We conclude that a more rigorous statistical analysis, taking into account the tails of the distribution of the conductance, e.g., Eq. (6), as a function of the configuration of the LS's, would be desirable—especially for chains with higher values of N. Also, for biases $eV > 50$ mV, the distortion of the barrier shape is non-negligible, but is not accounted for here. Thus the values of $\{\nu_{N,V}\}$ resulting from the fitting most likely reflect the barrier distortion to some degree. The surprising success of the fits to some extent reflects a fortunate cancellation of complications neglected in the original model. We should also point out that these values of $\nu_{N,2}$ and $\nu_{N,3}$ found for the small area junctions with Nb electrodes are about five times smaller than the corresponding values found in Sec. IV B 2 for the large area Mo junctions. We are not certain if this difference is a consequence of the electrode material (affecting, for example, the effective tunneling density of states), the relative sizes of the junctions, or some other factor.

4. *Thick barriers—crossover to* VRH *behavior*

In this section we analyze data for junctions with much thicker barriers in the high-bias regime. Figure 8 is a semilogarithmic plot of the current density versus $E^{-1/4}$, for a series of junctions with barrier thicknesses ranging from 270 to 990 Å, measured at 4.2 K. The current density is simply the total tunneling current divided by the junction area, and the electric field is taken to be the bias voltage divided by the barrier thickness. The upper bound of the abscissa corresponds to about 20-mV bias for the 270-Å barrier junction, whereas the lower bound of the abscissa corresponds to about 4 V for the 990-Å barrier. Some of the basic assumptions of our analysis, e.g., a uniform density of localized states and a bias- and position-independent localization length, may no longer be valid under such extreme bias conditions.

One striking feature of Fig. 8 is that the data for the $d=660$- and 990-Å barriers fall essentially on top of each other, suggesting that the $d=660$-Å sample has already attained the bulk limit, wherein resistance scales linearly with barrier thickness. Moreover, if we fit the data of the 990-Å sample to the functional form

$$J(E)=\hat{J}_0\exp[-(E^*/E)^{\eta}]\ ,$$

we obtain $\eta=0.25$ and $E^*=1.8\times10^{11}$ V/cm as the best fit (solid line). (Note that this fitting procedure heavily weights the points at high biases when displayed on a logarithmic plot.) Despite the high-bias voltage involved, the data seem to be adequately described by Eq. (27) over a limited range. Moreover, using Eq. (28) and the experimentally determined values $T^*=1.24\times10^8$ K and $\alpha^{-1}=6.7$ Å, we calculate $E^*_{\text{theory}}=1.2\times10^{11}$ V/cm, in reasonably good agreement with the experimentally obtained value.

The conductance at low bias, especially for the thinner samples in Fig. 8, exhibits systematic deviations from the variable-range-hopping-like behavior. This is a clear indication of the breakdown of simple variable-range hopping model. Similar observations presumably led Pollak and Hauser to propose the existence of statistically rare but highly conductive quasi-1D hopping chains. We now give a more quantitative argument which specifies this crossover in terms of the bias dependence of the hopping conductance.

Equation (27) may be written as $j(E)\propto\exp(-2\alpha l_{\text{VRH}})$, implying a voltage-dependent typical hopping length l_{VRH}:

TABLE II. The voltage coefficients of the inelastic hopping channels, $\{\nu_{N,V}\}$, for $N=2$–5, calculated from the best fits to the conductance data of four samples. A blank space indicates that the term was not used in the fit (see text). The last column shows a theoretical estimate based on Refs. 9–11.

Sample name	9113A3	9113B8	8643A7	8643A6	Theoretical estimate
d (Å)	90	125	150	180	Theoretical estimate
S (μm^2)	8×8	4×4	8×8	8×8	$\sim(N-1)^{N-1}$
$\nu_{2,V}$	0.6	0.5	0.4	0.5	~1
$\nu_{3,V}$	4	7	2.5	3	~4
$\nu_{4,V}$		25	25	20	~27
$\nu_{5,V}$		220	300	200	~256

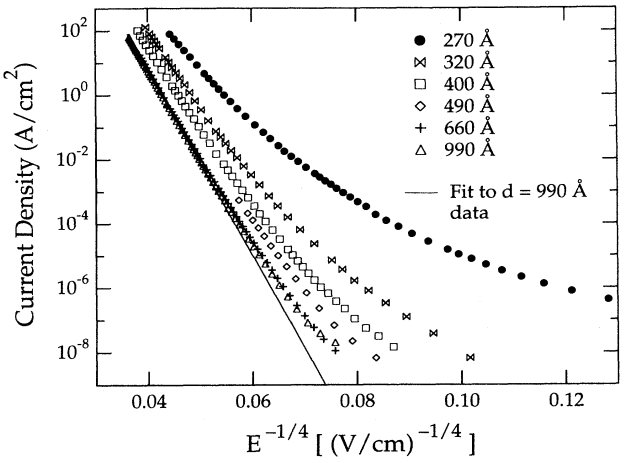

FIG. 8. Current density vs $E^{-1/4}$ for a series of samples with thick barriers measured at 4.2 K. The behavior of the thickest barriers at high bias conforms to the VRH model (Ref. 22). Thinner barriers show substantial deviations, especially at low bias. This may be understood in terms of a crossover in the dominant conduction process from VRH to directed hopping along chains.

$$l_{\mathrm{VRH}}=\frac{\alpha^{-1}}{2}(E^{*}d/V)^{1/4} . \tag{32}$$

For the $d=990$-Å sample at $V=0.5$ V, using $E^{*}=1.8\times10^{11}$ V/cm, we find that $l_{\mathrm{VRH}}\approx145$ Å. On the other hand, the Pollak-Hauser hopping chains will, according to Levin, Ruzin, and Shklovskii,[10] [cf. Eqs. (23) and (24)], give an averaged conductance of the form $G_{\mathrm{Levin}}(V)\propto\exp(-2al_c)$, where the characteristic hopping length l_c is

$$l_c=(2\alpha^{-1}d\delta_V)^{1/2},\quad \delta_V=\ln\left[\frac{0.234}{g\alpha^{-2}deV}\right] . \tag{33}$$

The conductance of the 990-Å sample first becomes measurable at a bias of 0.5 V, which is the lowest open triangle in Fig. 8. Note that the actual sample conductance at this point is about two orders of magnitude larger than it would have been had the conductance followed the VRH behavior (solid line) extrapolated all the way down to 0.5 V. For $d=990$ Å and $V=0.5$ V, and using $g=8\times10^{18}$ eV^{-1} cm^{-3} we find $l_c\approx60$ Å $\ll l_{\mathrm{VRH}}$. Since the conductance in both cases decreases exponentially with these hopping lengths, the Pollak-Hauser hopping channel predicts exponentially larger conductance over field-driven variable range hopping. This comparison demonstrates why the Pollak-Hauser hopping mechanism dominates over VRH at the lower end of the $d=990$-Å curve. However, we cannot presently account for the crossover from VRH to directed hopping along quasi-1D chains quantitatively by any available theory. For example, for the same 990-Å sample, at $V=1$ V the argument of the logarithmic function in Eq. (33) drops below one, and δ_V becomes a meaningless negative number. Part of the difficulty is that the high-bias regime violates many of the assumptions of the models discussed in Sec. II.

In conclusion, we lack a theory that can account properly for the crossover from directed conduction along quasi-one-dimensional chains of localized states to the three-dimensional diffusive percolation of VRH. This issue arises again in the context of the temperature dependence of the hopping conductance, to which we now turn.

C. Temperature dependence of hopping conduction ($eV\ll k_BT$)

1. Observation of $N=2$, 3, 4, 5, and 6 channels

We now analyze hopping conductance as a function of temperature. Figure 9(a) shows the zero-bias conductance as a function of temperature for a sample with barrier thickness $d=120$ Å and with a junction area of 90×90 μm^2. For $d=120$ Å, $G_1^{\mathrm{res}}>50G_0^{\mathrm{dir}}$, so we can neglect direct tunneling entirely. We first show a fit (the dashed line) to the data between 1.4 and 8 K to the form

$$G_2(T)=\sigma_0+\sigma_2T^{n_2} , \tag{34}$$

in complete analogy with our previous analysis of the bias dependence [cf. Eq. (30)], with σ_0, σ_2, and n_2 free to vary. The best fit yields $n_2=1.33$ (dashed line).

Anticipating a crossover to the $N=3$ channel, we next fit to the form

$$G_{\sum}(T)=(\sigma_0+\sigma_2T^{1.33}+\sigma_3T^{2.5}) , \tag{35}$$

with σ_0, σ_2, and σ_3 free to vary from 1.4 to 15 K (solid line). This function improves the accuracy of the fit in the lower-temperature range by altering slightly the parameters σ_0 and σ_2 to take into account the nonnegligible contribution of the $N=3$ channel. In order to account for the further increase of the conductance at higher temperatures, shown in Fig. 9(b), we add to Eq. (35) one more term, $\sigma_4T^{3.6}$, in accordance with the prediction of Ref. 11. σ_3 and σ_4 are now fitting parameters while σ_0 and σ_2 retain their previous values, since the higher-temperature data should not be sensitive to their values. The solid line in Fig. 9(b) is the best fit. The extrapolated contribution from the sum of the resonant tunneling background and the $N=2$ and 3 hopping channels (dashed line) shows that the crossover to the $N=4$ channel takes place at about 25 K.

Note that the procedure described above is robust with respect to reasonable changes in the limits of the successive fits, and thus is not arbitrary. Also, although the overall fit makes use of four parameters, each successive fit requires only one free parameter to fit the difference between the data and the extrapolation of the previous fit from lower temperatures. Though more cumbersome, this procedure explicitly demonstrates at each step that the onset of the higher-order hopping channels obeys the predicted power-law form, whereas a single multiparameter fit to the whole curve would not inspire the same degree of confidence in the results. Alternatively, one can also check the successive power exponents by allowing the exponent and the prefactor to vary at each step, as was done earlier for the bias dependence. This procedure

always yields exponents within a few percent of the predicted values. We have fixed the values of the exponents here in order to obtain the most procedurally consistent values of the prefactors for later comparison, as we did in Sec. IV B 2.

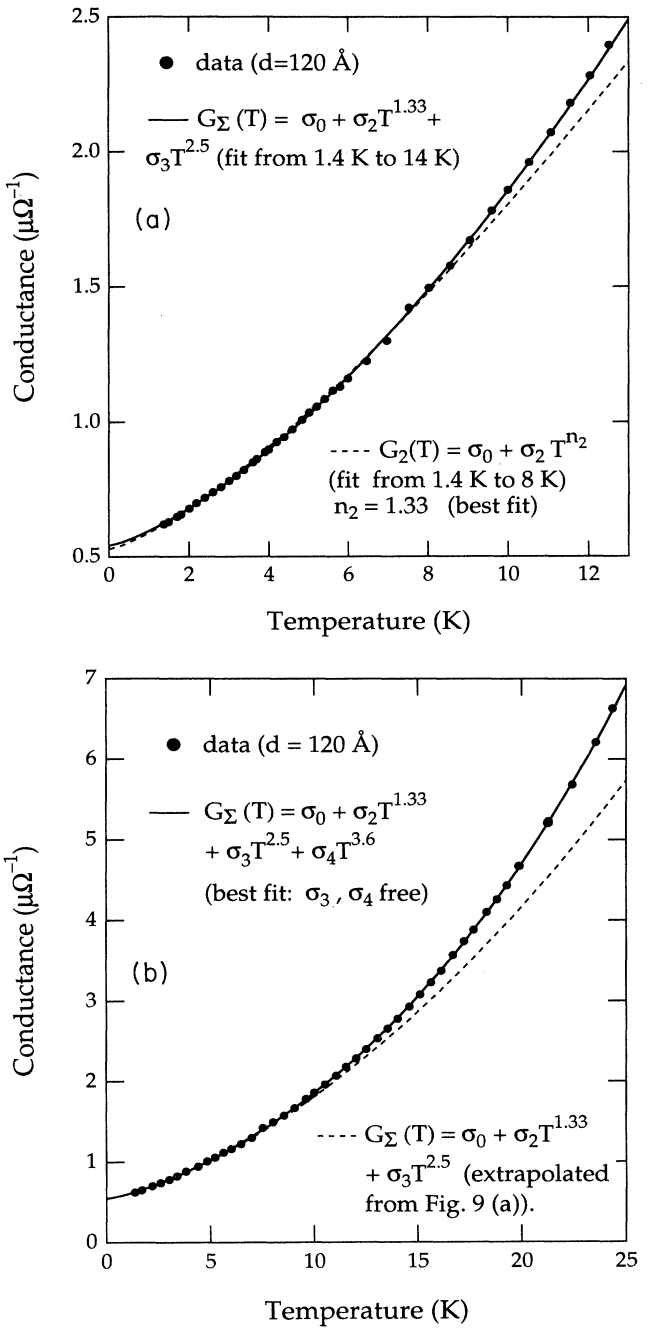

FIG. 9. (a) Conductance vs temperature for a $d = 120$-Å sample. The dashed curve is the best fit to the data from 1.4 to 8 K. The best fit yields $n_2 = 1.33$. The solid line is a fit to the data from 1.4 to 14 K, with n_2 fixed at 1.33 and with n_3 fixed at 2.5. (b) Data of the same sample up to 25 K. One more term with a $T^{3.6}$ dependence is added to account for the further increase in the conductance. The solid line is the best fit to the data with the power exponents fixed to their theoretically predicted values.

We applied the identical procedure to two more samples with $d = 90$ and 98 Å, and a similar procedure to one sample with $d = 150$ Å. Figure 10 shows the results for the $d = 98$-Å sample. We analyzed in great detail the bias dependence of this sample in Secs. IV B 1 and IV B 2. The fits were restricted to the range above 8 K to avoid the anomalous low-temperature behavior, to which we shall turn briefly in Sec. IV D. We note that the conductance behavior below 25 K is explained entirely by $N = 2$

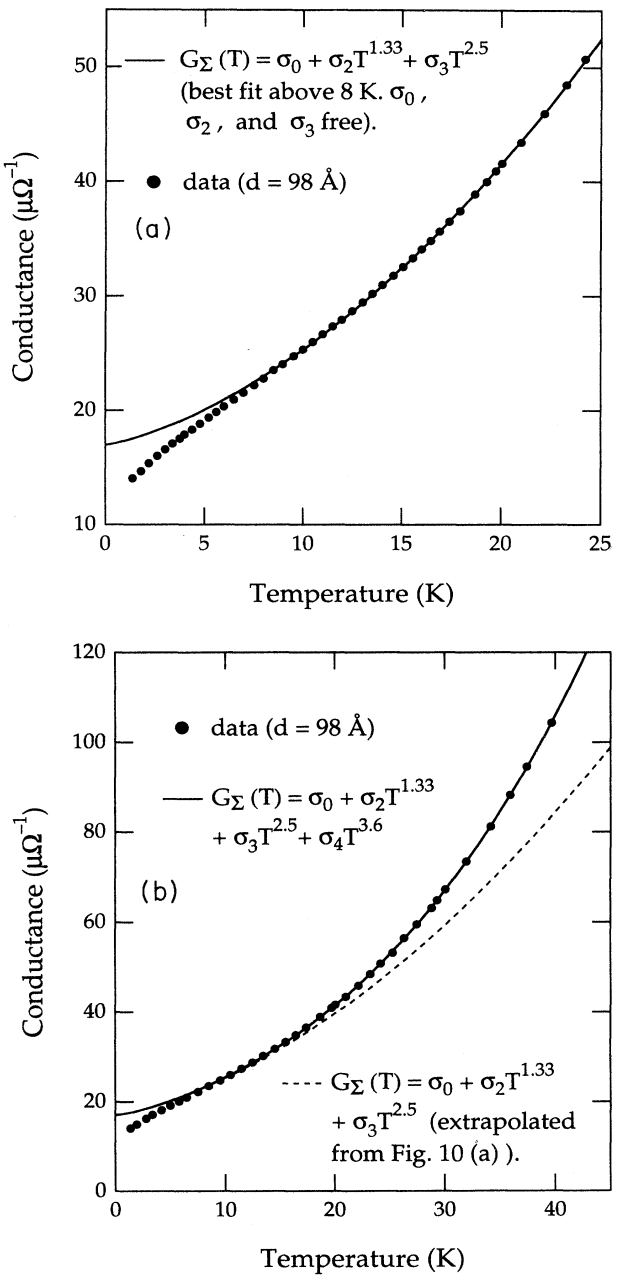

FIG. 10. The data for $d = 98$ Å presented in the same manner as in Fig. 9. A comparison with Fig. 9(b) reveals that the crossover to higher-order channels takes place at higher temperatures as the barrier becomes thinner.

and 3 channels only [Fig. 10(a)], whereas for the $d=120$-Å junction the $N=4$ channel is needed [cf. Fig. 9(b)] to fit the data in the same temperature range. Subsequently the crossover to the dominance of the $N=4$ channel takes place at around 40 K [Fig. 10(b)], much higher than the corresponding crossover temperature for the $d=120$ Å junction [about 25 K, cf. Fig. 9(b)]. The thinner sample with $d=90$ Å manifests very similar behavior, and follows the trend that the crossover to higher-order channels takes place at higher temperatures as the thickness decreases, as predicted by Eq. (16).

The data for the thickest sample $d=150$ Å appear in Fig. 11 (the solid circles), plotted versus $T^{-1/4}$ for later convenience. The solid line is the sum of the $N=3$ through $N=6$ channels, fit over an intermediate temperature range with the four coefficients from $\nu_{3,T}$ to $\nu_{6,T}$ free to vary simultaneously. The $N=2$ channel is not included since its contribution is negligible even at 4.2 K for this thickness. The dashed line includes two more channels (up to $N=8$) using Eq. (36), which is discussed below, to calculate the coefficients for the two additional channels, which are then simply added on to the solid line—no additional fitting is performed. No reasonable variation of $\nu_{7,T}$ and $\nu_{8,T}$ will bring the curve into line with the data. Thus the directed chain model appears to break down at about $N=6$ for this junction. We will consider the implication of these results in Sec. IV C 2.

The values of the prefactors of the various hopping channels can be used, in combination with the theoretical predictions in Eqs. (14) and (16), to yield the relative weights (the temperature coefficients $\{\nu_{N,T}\}$) of the various channels. The extraction of the coefficients $\{\nu_{N,T}\}$ is similar to that carried out in Sec. IV B 2 for $\{\nu_{N,V}\}$, and the resulting values for the four samples discussed above are listed in Table III.

The *only* coefficient calculated explicitly in Ref. 11 is $\nu_{2,T}\approx 50$. The strikingly good agreement of the data with the prediction of Glazman and Matveev for this term strengthens the conclusion that their model accurately describes our experimental system. For $N>2$ they only estimate that $\{\nu_{N,T}\}$ should vary as N^{2N}. In order to normalize this estimate to the value calculated for $\nu_{2,T}$, we insert a factor of π, i.e., $\nu_{N,T}\approx\pi N^{2N}$. This estimate is

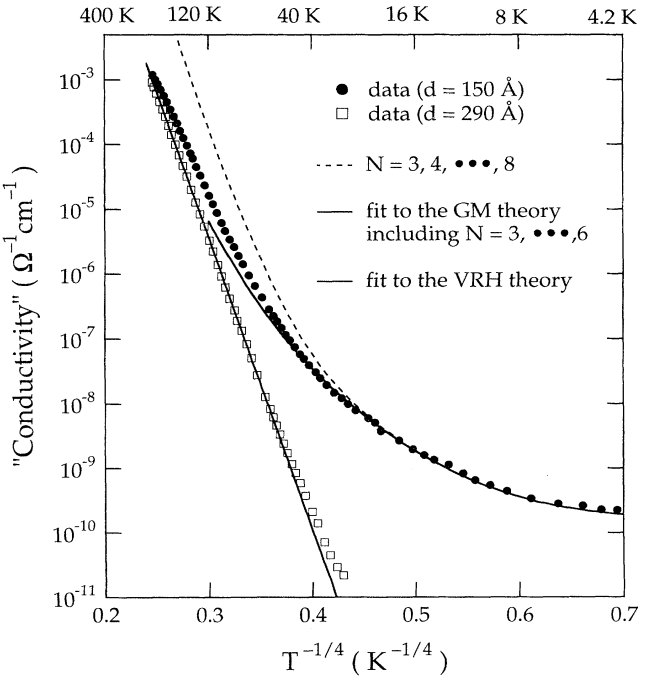

FIG. 11. Conductivity (Gd/S) plotted vs $T^{-1/4}$ for two samples. Where the data for the 290-Å sample appear linear on this plot, the sample has attained its bulk behavior. The solid lines are fits to the directed hopping model for the $d=150$-Å sample and to the theory of VRH for the $d=290$-Å sample. The dashed line shows the failure of additional terms in the directed hopping model to explain the high-temperature data.

tabulated as a function of N in the penultimate column of Table III. The disparity between this estimate and the experimentally determined coefficients, along with the observation that the integral over the coordinates of the localized states [Eq. (3.2) in Ref. 11] should generate a factor of π for every *inelastic* coupling, led us also to try the form

$$\nu_{N,T}\approx\pi^{N-1}N^{2N}\ , \tag{36}$$

tabulated in the final column of Table III. Equation (36)

TABLE III. Temperature coefficients of the inelastic hopping channels, $\{\nu_{N,T}\}$, calculated from fitting to the conductance vs temperature data of various samples. The penultimate column contains rough theoretical estimates from Ref. 11, except for the entry marked by the dagger (†), which is explicitly calculated in Ref. 11. The final column is a modified form of the estimate [Eq. (36)] that fits the observed values better.

Sample name	9206B4	9206B3	9207B5	9113B7	Theoretical estimate	
d (Å)	90	98	120	150	(Ref. 11)	(This work)
S (μm^2)	40×40	50×50	90×90	20×20	πN^{2N}	$\pi^{N-1}N^{2N}$
$\nu_{2,T}$	52	43	52		~50†	50
$\nu_{3,T}$	7.6×10^3	5.8×10^3	5.3×10^3	3.0×10^3	2.3×10^3	7.2×10^3
$\nu_{4,T}$	12×10^6	8.3×10^6	5.6×10^6	1.6×10^6	0.21×10^6	2.0×10^6
$\nu_{5,T}$				7.0×10^8	0.31×10^8	9.6×10^8
$\nu_{6,T}$				1.0×10^{12}	0.0068×10^{12}	0.66×10^{12}

is in much better agreement with the data.

As emphasized previously, the values of the coefficients associated with the various hopping channels, unlike the power exponents, are sensitive to the approximations of the model. For example, the transition rate of a hop contains power-law factors of the distance r between the two localized states involved besides the exponential dependence on r shown in Eq. (8). This implies that $\nu_{N,T}$ will depend weakly on d. A nonuniform barrier height and charging effects (see Sec. IV B) neglected in Ref. 11, among other approximations, will also scatter the values of the coefficients. Moreover, small errors in the values of the physical parameters in Table I, because they enter into expressions of the prefactors of the power-law terms raised to different powers in the various channels, will propagate through the calculations and ultimately lead to a spread in the values $\{\nu_{N,T}\}$ tabulated in Table III. Thus, the range of values found for each coefficient ($N>2$) in Table III, as well as the discrepancies with the theoretical estimates, is not surprising. To expect that Eq. (36) would do better than giving the correct order of magnitude for $\nu_{N,T}$ would be to stretch the theory beyond its limits.

2. *Thick barriers—crossover to* VRH *conduction*

The limit of high temperatures and thick barriers exposes the limitations of the directed hopping chain model. (We have already seen in Sec. IV C 1 that this model breaks down for the $d=150$-Å junction at about $N=6$.) This regime lends itself to a quantitative analysis more readily than does the high-bias voltage limit because the temperature does not introduce the nonidealities that the voltage does (see Sec. II F). Figure 11 shows the conductance of two junctions multiplied by their respective barrier thicknesses (i.e., the conductivity) plotted against $T^{-1/4}$. The nearly linear variation of the 290-Å data on this plot at high temperatures suggests that the sample has already attained the bulk limit, at least above 100 K. The solid line is a fit of that part of the data to the form

$$\ln G_{\rm VRH}(T)=\ln\hat{G}_0-(T^*/T)^{\eta}\ ,$$

with $\hat{G}_0$, η, and T^* as free parameters. The logarithm insures that every point receives equal weight, as measured by vertical distance on this graph. The best fit yields $T^*=1.24\times10^8$ K and $\eta=0.25$. The exponent confirms that Mott VRH is the predominant conduction mechanism in the limit of thick barriers and high temperatures. The value of T^* together with Eq. (26) is the source of our estimate for the density of localized states $g=8\times10^{18}$ $\mathrm{eV}^{-1}\,\mathrm{cm}^{-3}$. If we take

$$l_{\rm VRH}=\frac{\alpha^{-1}}{2}(T^*/T)^{1/4}\ ,$$

in analogy with Eq. (32), then $l_{\rm VRH}$ ranges from about 90 Å at room temperature up to about 170 Å at 20 K. Thus the $d=290$-Å data suggest that for $d>2l_{\rm VRH}$, Mott's law provides a reasonably good description of the transport. The approach of the $d=150$-Å data to that of the thicker junction near room temperature also supports this conclusion.

That a percolative model explains the transport across a film only twice as wide as the critical percolation length is somewhat surprising at first. Nonetheless, Shklovskii has demonstrated[20] that in three dimensions the crossover from directed hopping along rare but highly effective chains to VRH, associated with the diffusive regime, occurs fairly abruptly when $d\sim l_{\rm VRH}$. In two dimensions, in contrast, he predicts that severely twisted but nonetheless nonpercolative directed paths should be important in the crossover regime. Nevertheless, we lack a theory that describes precisely how the directed rectilinear chains evolve into a percolative network in our junctions. In Fig. 11, the range where the $d=150$-Å data points diverge from the $N=2,\ldots,6$ fit (solid line) until they converge toward the VRH fit (also a solid line) illustrates this crossover regime. The additional channels ($N>6$) included in the dashed line fail to explain the data at high temperatures, most likely because the probability of finding even a single directed chain of this order within the finite area of the junction is vanishingly small, so the correct prefactor is actually 0.

With increasing temperature, highly directed chains of high order ($N>2$) apparently continue to dominate the conductance until the required configurations become essentially impossible to realize in junctions with a given finite area [cf. Eq. (15)] and its discussion, at which point VRH takes over rather abruptly. If these two mechanisms dominate over temperature ranges that have little overlap, as Shklovskii's argument[20] suggests, their conductance should add in parallel. In fact we can simply add the two solid lines in Fig. 11 to fit the 150-Å data reasonably well over the entire temperature range. However, a firm conclusion on this point is not really possible based solely on this graph.

D. Low-temperature zero-bias anomalies (ZBA)

We have already presented evidence of anomalous behavior at low temperatures in Fig. 10. We observe similar behavior in all of our junctions with $d<120$ Å. In all cases, below some temperature, $T_{\rm ZBA}$, the resistance begins to rise rapidly, departing from the low-temperature extrapolation of the fits to the model of Ref. 11 that work exceedingly well at higher temperatures. $T_{\rm ZBA}$ decreases with increasing barrier thickness and ranges from about 10 K for $d=60$ Å to about 500 mK for $d=120$ Å (thicker junctions were not measured below 4.2 K). Figure 12 shows the data for a $d=62$-Å junction with Nb electrodes in an applied field of 2.5 T, sufficient to quench the superconductivity in the electrodes. (Based on similar data in other samples, the magnetic field also sharpens the resistive upturn but does not appear to affect $T_{\rm ZBA}$ significantly.)

We have carried out a preliminary study of this anomalous low-temperature behavior as a function of magnetic field, temperature, barrier thickness, and bias. In order to explain these data, we have considered a variety of possible theories, including the Kondo effect,[30] hybrids of the Kondo effect that account for the LS's in the barrier

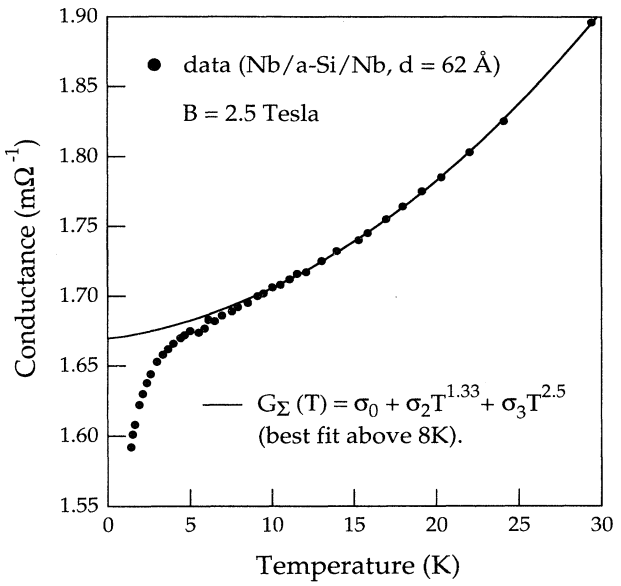

FIG. 12. Conductance vs temperature for a thin barrier junction showing a low-temperature anomaly. A 2.5-T field perpendicular to the ground plane was used to quench the superconductivity in the niobium electrodes. The solid curve is a fit to the data above the anomaly including the $N=2$ and 3 channels.

as well as at the interfaces,[31] and charging effects. We have not yet found a satisfactory quantitative explanation.

We have also studied the correlations of the hopping conductivity caused by the presence of a large on-site Coulomb interaction between electrons that share a localized site.[32] (Each uncoordinated dangling *a*-Si bond—a single localized *site*—gives rise to two spin-degenerate localized *states*.) These correlations manifest themselves most dramatically in the magnetoresistance at low temperatures, but still above T_{ZBA}. Thus we know that many-body effects play a significant role in our system at very low temperature and bias. More work is needed to remove the word anomalous from the description of the low-temperature behavior found in Figs. 10 and 12.

V. SUMMARY AND CONCLUSION

In this paper we have presented a detailed study of transport via localized states in tunnel junctions, using amorphous silicon tunnel barriers as a model system. Both elastic and inelastic processes are present and can be distinguished. The importance of hopping conduction through optimally configured channels (in space and energy) as temperature, voltage, or barrier thickness are increased has been clearly demonstrated. The presence of these channels is purely statistical and a consequence of the disordered nature of the material. The importance of such channels was first anticipated by Hauser and Pollack. The total conduction is a sum over channels containing $N=2,3,\ldots$ localized states. Characteristic of this process is that, as temperature or voltage increase, the dominant contribution to the conduction comes from channels with larger and larger numbers of localized states (i.e., increasing numbers of hops).

The dependence of the conductance on temperature, voltage, and barrier thickness is found experimentally to be in remarkably good agreement with the recent theory of Glazman and Matveev. In particular, we have demonstrated that the leading correction ($N=2$) to the zero-bias conductance due to hopping varies as $T^{4/3}$ and $V^{4/3}$ in the limits $kT \gg eV$ and $eV \gg kT$, respectively. The next-higher-order correction ($N=3$) has also been clearly identified. Fits to the sum over even larger numbers of channels is found to be in good agreement with the data over nine orders of magnitude in conductance.

These results clarify both the consequences of the presence of localized states in tunnel junctions and the nature of transport in disordered insulators at short length scales. The crossover from hopping along the directed channels studied here to fully developed variable range hopping is not completely understood, however, and may have some fundamental interest.

In this paper we have not discussed the correlated nature of these hopping processes due to the on-site Coulomb interaction. We have considered only single-electron processes. Correlations in the hopping are present, have been observed, and are also found to be in accord with recent experiments and theory.[32] A comprehensive understanding of these effects is not yet at hand, however. For example, as seen in the results presented here, zero-bias anomalies are present that are clearly due to the presence of localized states. A satisfactory understanding of these anomalies is not available. The potential of the amorphous silicon barriers studied here as a model system for studying zero-bias anomalies generally has not been explored.

Finally, we note that the effects of localized states on superconductive tunneling and the Josephson coupling have not been addressed in this paper. Relatively little systematic work has been done on these issues.[33] This situation may change in view of considerable interest in the role of localized states in the barriers of Josephson junctions made with the high-temperature superconductors that has arisen recently.[34–36]

ACKNOWLEDGMENTS

We are pleased to acknowledge many helpful discussions with Leo Glazman and Konstantin Matveev throughout the course of this work. This work was supported by the Office of Naval Research.

*Present address: School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853.

[1]N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials,* 2nd ed. (Oxford University Press, New York, 1979).

[2]D. A. Rudman and M. R. Beasley, Appl. Phys. Lett. **36**, 1010 (1980).

[3]R. Meservey, P. M. Tedrow, and J. S. Brooks, J. Appl. Phys. **53**, 1563 (1982).

[4]S. J. Bending and M. R. Beasley, Phys. Rev. Lett. **55**, 324 (1985).

[5]M. Naito and M. R. Beasley, Phys. Rev. B **35**, 2548 (1987).

[6]Yizi Xu, A. Matsuda, and M. R. Beasley, Phys. Rev. B **42**, 1492 (1990).

[7]M. Pollak and J. J. Hauser, Phys. Rev. Lett. **31**, 1304 (1973).

[8]J. J. Hauser, Phys. Rev. Lett. **29**, 476 (1972); Phys. Rev. B **8**, 3817 (1973); J. J. Hauser and A. Staudinger, *ibid.* **8**, 607 (1973).

[9]A. V. Tartakovskii, M. V. Fistul, M. E. Raikh, and I. M. Ruzin, Fiz. Tekh. Poluprovodn. **21**, 603 (1987) [Sov. Phys. Semicond. **21**, 370 (1987)].

[10]E. I. Levin, I. M. Ruzin, and B. I. Shklovskii, Fiz. Tekh. Poluprovodn. **22**, 642 (1988) [Sov. Phys. Semicond. **22**, 401 (1988)].

[11]L. I. Glazman and K. A. Matveev, Zh. Eksp. Teor. Fiz. **94**, 332 (1988) [Sov. Phys. JETP **67**, 1276 (1988)].

[12]D. Popovic, A. B. Fowler, and S. Washburn, Phys. Rev. Lett. **67**, 2870 (1991).

[13]J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).

[14]R. E. Prange, Phys. Rev. **131**, 1083 (1963).

[15]A. I. Larkin and K. A. Matveev, Zh. Eksp. Teor. Fiz. **93**, 1030 (1987) [Sov. Phys. JETP **66**, 580 (1987)].

[16]L. I. Glazman and R. I. Shekhter, Zh. Eksp. Teor. Fiz. **94**, 292 (1988) [Sov. Phys. JETP **67**, 163 (1988)].

[17]A. Douglas Stone and P. A. Lee, Phys. Rev. Lett. **54**, 1196 (1985).

[18]Ned S. Wingreen, Karsten W. Jacobsen, and John W. Wilkins, Phys. Rev. Lett. **61**, 1396 (1988); Phys. Rev. B **40**, 11 834 (1989).

[19]J. G. Simmons, J. Appl. Phys. **35**, 2655 (1964).

[20]B. I. Shklovskii (private communication).

[21]B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Berlin, 1984).

[22]B. I. Shklovskii, Fiz. Tekh. Poluprovodn. **6**, 2335 (1972) [Sov. Phys. Semicond. **6**, 1964 (1973)].

[23]R. H. Hammond, IEEE Trans. Magn. **MAG-11**, 201 (1975); J. Vac. Sci. Technol. **15**, 382 (1978).

[24]S. J. Bending, Ph.D. thesis, Stanford University, 1985.

[25]Yizi Xu, Ph.D. thesis, Stanford University, 1992.

[26]A. S. Barrera and M. R. Beasley, IEEE Trans. Magn. **MAG-23**, 866 (1987).

[27]R. A. Street (private communication).

[28]J. Bardeen and W. Shockley, Phys. Rev. **80**, 72 (1950).

[29]I. Gieaver, in *Tunneling Phenomena in Solids,* NATO Advanced Study Institute, Risö, Denmark, 1967, edited by Elias Burstein and Stig Lundqvist (Plenum, New York, 1969), pp. 19–30.

[30]J. Appelbaum, Phys. Rev. Lett. **17**, 91 (1966).

[31]For one example, see L. I. Glazman and M. E. Raikh, Pis'ma Zh. Eksp. Teor. Fiz. **47**, 378 (1988) [JETP Lett. **47**, 452 (1988)].

[32]D. Ephron, Yizi Xu, and M. R. Beasley, Phys. Rev. Lett. **69**, 3112 (1992); D. Ephron, M. R. Beasley, H. Bahlouli, and K. A. Matveev, Phys. Rev. B **49**, 2989 (1994); H. Bahlouli, K. A. Matveev, D. Ephron, and M. R. Beasley, *ibid.* **49**, 14 496 (1994).

[33]I. A. Devyatov and M. Yu. Kupriyanov, Pis'ma Zh. Eksp. Teor. Fiz. **59**, 187 (1994) [JETP Lett. **59**, 200 (1994)].

[34]J. Halbritter, Phys. Rev. B **48**, 9735 (1993).

[35]J. N. Eckstein, I. Bozovic, and G. F. Virshup, MRS Bull. **XIX**, 44 (1994).

[36]I. I. Vengrus, M. Yu. Kupriyanov, O. V. Snigiref, A. G. Maresov, and S. I. Krasnosvobodtsev, Pis'ma Zh. Eksp. Teor. Fiz. **60**, 372 (1994) [JETP Lett. **60**, 381 (1994)].

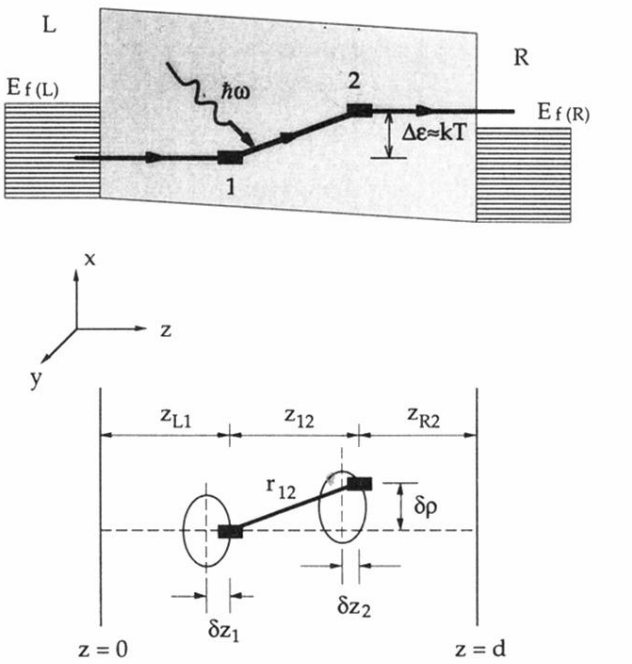

FIG. 1. Upper panel: a schematic energy vs position (z direction) diagram showing an electron tunneling from the left electrode (L) to the right electrode (R) through two localized states (LS's) labeled 1 and 2. The hop between the LS's is inelastic and is associated with the emission (or absorption) of a phonon. The most important pairs of the LS's are those within $k_B T$ of the Fermi level with energy separation $\Delta\varepsilon \approx k_B T$. Lower panel: a schematic representation of the positions (x, y, and z) of the two LS's forming a nearly optimal hopping chain of the upper panel. The first LS must be located within δz_1 of its ideal longitudinal position, but may be located anywhere laterally across the junction. The second LS should be within $\delta\rho$ of the first LS in the lateral direction, and within δz_2 of its ideal longitudinal position.