Variable-range hopping in the critical regime

T. G. Castner
Department of Physics and Applied Physics, University of Massachusetts Lowell, Lowell, Massachusetts 01854
(Received 25 August 1999; revised manuscript received 28 December 1999)

Mott's theory of variable range hopping (VRH) is extended to the critical regime approaching the metal-insulator transition where the hopping energy $\Delta_h(T) < kT$. The Miller-Abrahams impedance between a pair of sites is minimized for an arbitrary value of Δ_h/kT. The theory features spatial dispersion of the dielectric response function $\epsilon(R,n)$, which introduces a new length scale r_s into the Miller-Abrahams resonance energy. In the regime $\Delta_h \ll kT$ the Mott exponent 1/4 changes to 2/7 and the numerical factor [18.1] in the Mott T_0 is reduced to 1.51. Efros-Shklovskii VRH is also modified by the inclusion of $\epsilon(R,n)$ and the T_0' depends critically on the spatial dispersion of $\epsilon(R,n)$ and this feature allows an explanation of the Ge:Ga data of Watanabe *et al.*, which exhibits $T_0' \propto (1-n/n_c)^{1.0}$. The new result for the Mott case allows the Si:As and Si:P results to be explained over a range of 10^6 in Mott T_0 values encompassing the crossover from high-temperature Mott VRH to conventional low-temperature Mott VRH. The reduced density dependence of the Mott VRH prefactor $\sigma_0(n)$ is also satisfactorily explained. In the critical regime for $1-n/n_c < 0.05$, Si:As and Si:P yield the localization length exponent $\nu \sim 1$.

For barely insulating disordered metal-insulator transition (MIT) systems the conductivity is determined by different phonon-assisted hopping processes at low T and at higher T there is also a thermally activated component from itinerant electrons excited above the mobility edge. The activated component $\sigma_a \propto T^p \exp[-\epsilon_2/kT]$ becomes negligible for $\epsilon_2/kT \gg 1$ except extremely close to n_c because $\epsilon_2(n) \rightarrow 0$ as $n \rightarrow n_{c-}$. For weakly compensated barely insulating samples the two most important hopping processes lead to Mott[1] and Efros-Shklovskii[2] (ES) variable range hopping (VRH) conduction. These are of the form $\sigma(n,T) = \sigma_0(n,T)\exp[-(T_0/T)^m]$ with $m = 1/4$ for Mott VRH and $m = 1/2$ for ES VRH. There are numerous observations of both Mott VRH and ES VRH in many systems ranging from amorphous semiconductors,[3] which led to Mott's original formulation, crystalline doped semiconductors,[4,5] amorphous semiconductor-metal alloys,[6] quasicrystals,[7] and even in the high-temperature superconducting materials[8] like $La_{2-x}Sr_xCuO_4$. In a number of cases the crossover between Mott VRH and ES VRH has been documented as a function of temperature, doping, or magnetic field. A particularly elegant treatment of the crossover has been given by Aharony, Zhang, and Sarachik[9] (AZS). A number of studies have documented the scaling behavior of the Mott and ES characteristic temperatures T_0 and T_0'. In general one finds $T_0 \gg T_0'$ and one observes ES VRH at the lowest T and Mott VRH at higher T. In the critical regime one finds the Mott hopping energy $\Delta_h = 1/4k(T^3T_0)^{1/4}$ becomes less than kT. The original Mott derivation, and also subsequent derivations by Ambegaokar, Halperin, and Langer[10] and by Pollak,[11] were intended for the regime $T_0 \gg T$. In the critical regime where the hopping energy $\Delta_h < kT$ and the mean hopping distance $R_h(n,T) < \xi(n)$, where ξ is the localization length, this calculation uses the Miller-Abrahams[12] (MA) pair approximation and the impedance $Z(R,\Delta)$ between a pair of sites with spacing R and energy difference Δ. A key feature of the approach of this paper is the introduction of spatial dispersion[13] of the dielectric response $\epsilon(n,R,T)$ that is crucial in the critical regime where $\epsilon(R \rightarrow \infty, n)$ diverges as $n \rightarrow n_{c-}$ but $\epsilon(R \rightarrow 0, n) \rightarrow \epsilon_h$ where ϵ_h is the dielectric constant of the host semiconductor. Since the original theories were motivated by the amorphous semiconductor results where the Mott T_0 was of order 10^8 K and the exponential temperature dependence was the dominant effect, insufficient attention was paid to the prefactor $\sigma_0(n,T)$ and the characteristic length L that entered the result in the percolation approach for which $\sigma = G_c/L$, where G_c is the critical conductance associated with the resistor network of MA. In the critical regime where $(T_0/T)^{1/4} < 1$ it is important to carefully consider the prefactor $\sigma_0(n,T)$.

BACKGROUND

The subject of VRH conduction is now three decades old and new experimental results continue to emerge. The theory has been discussed in detail by many theorists and is reviewed by Efros and Shklovskii.[14] With increased interest in MIT studies in the last two decades a large number of studies have documented both Mott and ES VRH as $n \rightarrow n_{c-}$ and some of these studies have documented the critical behavior of the Mott and ES characteristic temperatures. The critical behavior of VRH was reviewed[15] a decade ago, but a number of new, significant experimental results in the last five years have given important new information and presented some new puzzles to be explained. In addition, the emergence of finite-T scaling (fTs) has important implications for both VRH conduction and thermally activated conduction of itinerant electrons above the mobility edge.

A controversy developed over the proper interpretation of the Si:P data for $\sigma(n, T \rightarrow 0)$ and the scaling behavior of $\sigma(n > n_c, T=0) = \sigma_0(n/n_c - 1)^s$. Paalanen *et al.*,[16] obtained $s \sim 0.50$ and saw distinct $\delta\sigma = \sigma(n,T) - \sigma(n,0) = m(n)\sqrt{T}$ behavior to within 0.0008 of n_c. However, Stupp *et al.*,[17] claimed a crossover from $s \sim 0.55$, where $m(n) < 0$ to s

 ©2000 The American Physical Society

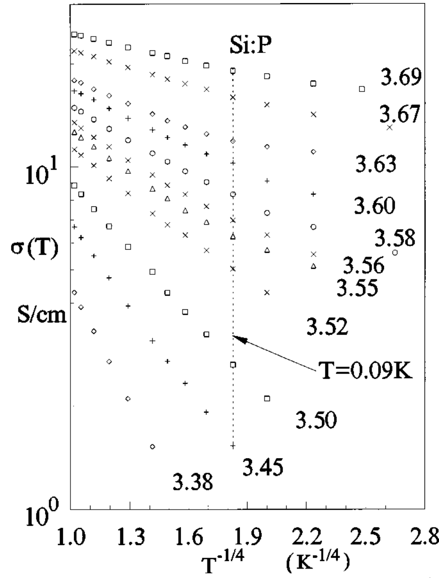

FIG. 1. The Si:P results of Stupp *et al.* (Ref. 17) replotted for comparison with Mott's VRH expression. The results for $T > 0.09$ K (vertical dashed line $T = 0.09$ K) are a good fit to Mott's 1/4 law. The sample dependent upward deviations (particularly pronounced for 3.63) result from thermal decoupling of the electrons from the lattice and thermal reservoir.

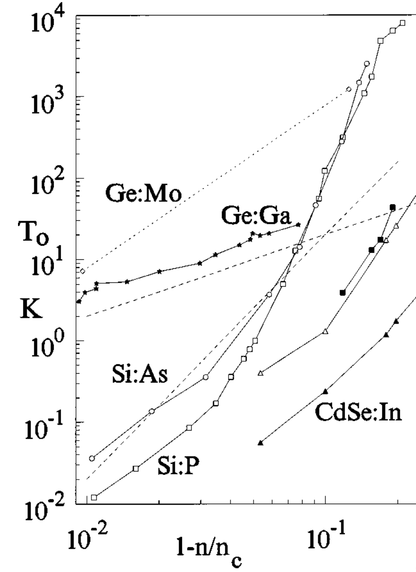

FIG. 2. Mott (open symbols) and ES (filled symbols) T_0 and T_0' values versus $1-n/n_c$ for Si:P (□ and ■), Si:As (○), CdSe:In (△ and ▲), GeGa (★), and Ge:Mo (◇). The rapid upturn in T_0 for Si:P and Si:As for $1-n/n_c > 0.04$ represents entry into the crossover regime where the coefficient $D(n)$ in T_0 increases rapidly. Note the Ge:Ga T_0' increases only linearly with $1-n/n_c$. The lower dashed line has a slope of three while the middle dashed line below the Ge:Ga data has a slope of one.

~ 1.3 when $m(n)$ became positive very close to n_c. Stupp *et al.*, claimed the critical regime was restricted to $m(n) > 0$. Their explanation reduced n_c from 3.74 to $3.52 \times 10^{18}/\text{cm}^3$ (from here on the $\times 10^{18}/\text{cm}^3$ will be dropped). This 6% reduction in n_c is a large reduction and the exponent s is strongly coupled to n_c. The Stupp *et al.*, plot of $\sigma(n,T)$ versus $\sqrt{T}$ (their Fig. 1) showed $m(n)$ increasing substantially with decreasing T suggesting a poor fit to $\sqrt{T}$ behavior. Their data for samples 3.38 to 3.69 is replotted in Fig. 1 in a Mott VRH plot ($\ln \sigma$ vs $T^{-1/4}$). The results suggest a good fit to the Mott law for $T > 90$ mK ($T^{-1/4} < 1.82\,\text{K}^{-1/4}$), but there are sample-dependent deviations (3.63 is particularly pronounced) that indicate thermal decoupling of the electrons from the thermal reservoir and this has been confirmed by a detailed analysis of the Kapitza resistance between the reservoir and sample and the thermal time constant between reservoir and sample. Stupp *et al.*, obtained a Mott $T_0 \sim 2.3$ K for the 3.52 sample they claimed as n_c, but $T_0 \propto [N(E_F)\xi(n)^3]^{-1}$ [$T_0 \propto (1-n/n_c)^3$ close to n_c] and T_0 must scale to zero as $n \to n_c$. The data in Fig. 1 for $n \geqslant 3.56$ are in the regime $\Delta_h/kT < 0.25$ at $T = 0.5$ K and $T_0 \sim 0.012$ K for $n = 3.69$. It will be shown below that this reanalysis leads to an n_c in good agreement with the Paalanen *et al.*, results. The Stupp *et al.*, data in Fig. 1 is very similar to the Si:As results[18] of Shafarman, Koon, and Castner (SKC, see Figs. 2 and 3).

A. The Mott and Efros-Shklovskii characteristic temperatures

Figure 2 shows values of Mott T_0's and ES T_0''s versus reduced density for Si:As, the results of Hornung and von Löhneysen[19] for Si:P (also the results of Stupp *et al.*, as replotted in Fig. 1 with a detailed analysis by Castner[20]), the new Ge:Ga results of Watanabe *et al.*,[21] and the CdSe:In results of Zhang *et al.*[22] The first three results are for very weakly compensated samples, while the CdSe:In samples have larger compensation, but are an excellent example because they show the crossover from Mott to ES VRH as the temperature is reduced below a crossover temperature T_x and yield values of both T_0 and T_0'. The new DC Si:P results overlap in reduced density the earlier Si:P results of Hess *et al.*,[23] measured at 400 MHz. The sharp difference in the $T_0(n)$ scaling results of Hess *et al.*, with the Si:As results has already been discussed.[15] The new DC Si:P results yield T_0 values more than 200 times larger than the Hess *et al.*, values for $1-n/n_c > 0.15$ and 100 times smaller for $n \sim 0.98 n_c$. 400 MHz corresponds to $T = 19$ mK, which is $3\Delta_h/4$ for $T_0 \sim 1$ K and $T = 50$ mK, while for $T_0 \sim 15$ K and $T = 60$ mK 400 MHz is about $\Delta_h/3$. When $h\nu$ is an appreciable fraction of Δ_h the VRH theory needs to take account of photon-assisted VRH which qualitatively will enhance $\sigma(n,T)$ and reduce the T_0 values. Hess *et al.*, claim $T_0 \sim (n_c/n - 1) = (n_c/n)(1-n/n_c)$ which is close to the ES $T_0' \sim (1-n/n_c)$ results for Ge:Ga but is very different than the DC Si:P and Si:As results.

Several striking features are illustrated in Fig. 2. The Si:As and Si:P T_0 behavior is similar featuring a sharp upward break to a very steep slope for $1-n/n_c > 0.04$. This steep slope is the intermediate regime where $\Delta_h \sim kT$ and represents the crossover between high-temperature Mott VRH ($\Delta_h < kT$) and conventional low-temperature Mott VRH ($\Delta_h > kT$). The crossover parameter $x = \Delta_h/2kT = (T_0/T)^{1/4}/8$ requires nearly four orders of magnitude in T_0 for a given T to get from $x \sim 0.3$ to $x \sim 3$, which is in qualitative agreement with the Si:P and Si:As results in Fig. 1. The slopes are similar for $1-n/n_c < 0.03$, however in this regime the ratio $T_0(\text{As})/T_0(\text{P}) = [N(E_F)\xi_0^3]_P/[N(E_F)\xi_0^3]_{\text{As}}$

is of order 3. The new ES T_0''s for Si:P for $1-n/n_c>0.12$ are nearly two orders of magnitude smaller than the Mott T_0's for the same value of n/n_c. One of the significant facts about the Si:As and Si:P results is that the mean hopping distance $R_h(n,T)=3/8\xi(n)(T_0/T)^{1/4}$ [$\xi(n)$ the localization length] goes through a minimum for $n\sim0.95\pm0.01n_c$ at a fixed T [for Si:As see $R_h(n,T=1\text{ K})$ values in Table IV in SKC]. *This minimum cannot be explained by conventional Mott VRH theory*. Specific heat results[24] show $N[E_F(n)]$ decreasing monotonically below n_c implying $R_h \propto\{\xi(n)/N[E_F(n)]T\}^{1/4}$ at fixed T will increase monatonically with n in the regime $0.8<n<0.99n_c$. The importance of this minimum in $R_h(n,T=\text{const})$ versus n has been overlooked. It is exactly this regime where spatial dispersion of $\epsilon'(n,R)$ is important.

The Mott T_0 for CdSe:In is approaching the same range of values as for Si:P and Si:As for $1-n/n_c<0.04$ and the ratio T_0/T_0' is approximately constant for $1-n/n_c<0.1$, but increases for larger values of $1-n/n_c$. The T_0 values for CdSe:In also exhibit a break to a steeper dependence on $1-n/n_c$ for $1-n/n_c>0.1$. The Ge:Ga results for the ES T_0''s are particularly striking because $T_0'\propto(1-n/n_c)^{\alpha}[\alpha\sim1.0]$ representing a much weaker dependence on $1-n/n_c$ than the other data. Furthermore, the values of T_0' are 100 or more larger than for Si:P and Si:As for $1-n/n_c<0.1$. For a given material in the critical regime one expects $T_0\gg T_0'$ and this is confirmed for the Si:P and CdSe:In results. Another important difference in the Ge:Ga results[21] are the strong deviations from ES VRH for $1-n/n_c<0.01$ where the VRH exponent $m[\ln\sigma(T)=-(T_0/T)^m]$ decreases rapidly from $m\sim1/2$ thru Mott's $m=1/4$ to $m\sim1/8$. The large difference in the value of $1-n/n_c$ where ES VRH starts to dominate between Ge:Ga and Si:P (and also Si:As, but where ES VRH has not been documented) suggests the critical regime for $n<n_c$ may be much smaller for Ge:Ga than for Si:P and Si:As.

The Coulomb gap plays a crucial role in ES VRH and the energy width of the Coulomb gap is $\Delta_{CG}\sim e^3N(E_F)^{1/2}/[\epsilon'(n,R,T)]^{3/2}$. If spatial dispersion of $\epsilon'(n,R,T)$ is ignored, one expects Δ_{CG} to decrease rapidly as $\epsilon'(n,T)$ diverges as $n\to n_c$. However, the important tunneling results of Massey and Lee[25] for Si:B demonstrate an initial decrease in Δ_{CG} from 0.85 to $0.93n_c$, but this is followed by successive increases in Δ_{CG} for 0.96 and $0.99n_c$. Spatial dispersion of $\epsilon'(n,R,T)$ can qualitatively account for this reversal. The broadening of Δ_{CG} very close to n_c can also account for the striking Ge:Ga results that demonstrate ES VRH for $0.9<n<.99n_c$.

Deutscher, Levy, and Souillard[26] have proposed a new VRH mechanism for disordered systems such as granular metals based on ''superlocalization'' of the wave function at the percolation threshold based on random walks on fractals. They find a hopping conductivity $\sigma(T)\propto\exp[-(T_0/T)^{(\zeta/\zeta+D)}$ where ζ is an exponent characterizing the wave function decay at E_F and D is the fractal dimensionality of the percolation cluster. Using the Alexander-Orbach[27] conjecture $D/\zeta=4/3$ leads to a VRH exponent 3/7, which is close to the ES result, but Coulomb interactions have been neglected. Unlike the usual situation discussed herein where the crossover is from Mott VRH to ES VRH as the temperature is reduced, the Deutscher, Levy, and Souillard proposal leads to 3/7 VRH at high T and 1/4 VRH at low T, and therefore is not relevant for the crystalline doped systems.

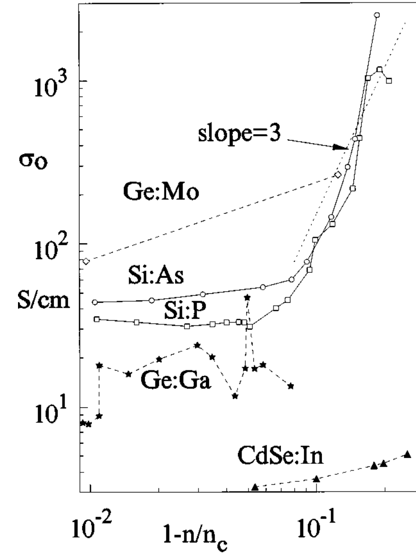

FIG. 3. Mott (open symbols) and ES (filled symbols) conductivity prefactors σ_0 versus $1-n/n_c$ for Si:P (□), Si:As (○), CdSe:In (▲), Ge:Ga (★), and Ge:Mo (◇). In all cases with sufficient data $\sigma_0(n)$ approaches a constant as $n\to n_c$. The dashed line with slope three is the theoretical prediction for $\sigma_0(n)$ for $z_m\ll1$.

B. The VRH conductivity prefactor $\sigma_0(n,T)$

The second issue for any VRH theory to address is the prefactor $\sigma_0(n,T)$ and specifically the T and n dependence of σ_0. In many cases experimentalists have fit their data with constant prefactors (T independent) that yield reasonable fits to the data, however, which may not be in agreement with theoretical predications. Pollak[11] obtained $\sigma_0\propto a^{-1}(T_0/T)^{1/4}$ for the Mott VRH prefactor. Allen and Adkins gave $\sigma_0\propto a^{-3}(T_0/T)^{1/2}$ for Mott VRH for $qR>1$ while ES gave $\sigma_0\propto(T_0'^3/T)^{1/2}$ for ES VRH, although the relevance of these results to the critical regime is unclear. There have been numerous analyses of prefactors by a variety of approaches (random resistor networks, etc.). These analyses have been reviewed by Efros and Shklovskii.

In Fig. 3 the prefactors are shown for five experimental systems ranging from the amorphous semiconductor-metal alloy system Ge:Mo, to *n*-type Si:As and Si:P, to *p*-type Ge:Ga, and finally to compensated *n*-type CdSe:In. The first three cases are for Mott VRH, while Ge:Ga and CdSe:In are for ES VRH. The two points for Ge:Mo are less than a factor of 2 larger than the Si:As values although an $(n_c)^{1/3}$ dependence from $\sigma_0\propto(e^2/hd_c)$ $[d_c=(n_c)^{-1/3}]$ would suggest a ratio of 8.5. The Si:As and Si:P results are similar in dependence on $1-n/n_c$ and both seem to approach a constant value as $n\to n_{c-}$. The ratio is in approximate agreement with the $[n_c(\text{As})/n_c(P)]^{1/3}$. The onset of a significant increase for $1-n/n_c>0.05$ suggests the onset of a different regime and is in good agreement with the T_0 versus $1-n/n_c$ results shown in Fig. 1. There is more scatter in the Ge:Ga results and the trend as $n\to n_{c-}$ is less obvious because of this scatter. Nevertheless, one can argue that σ_0 is

not changing rapidly. Based on the $n_c^{1/3}$ dependence the Ge:Ga values should be 0.36 times those for Si:P but are a little larger. The CdSe:In σ_0 for ES VRH approaches a value $3S/\mathrm{cm}$ as $n \to n_{c-}$ and gradually rise with increasing $1-n/n_c$, but without showing the sharp upward break shown by Si:As and Si:P. The magnitude is smaller than expected based on $\sigma_0 \propto n_c^{1/3}$ since $n_c = 2.8\times 10^{17}/\mathrm{cm}^3$ and the n_c dependence would yield a ratio (to Si:P) of 0.4, as compared to the experimental ratio of 0.06 at $1-n/n_c = 0.06$. Unfortunately, values of σ_0 for Mott VRH are not available. Comparison of the two ES VRH cases for σ_0 for Ge:Ga and CdSe:In appear to demonstrate a breakdown of the $n_c^{1/3}$ dependence of σ_0 since $n_c(\mathrm{CdSe{:}In})/n_c(\mathrm{Ge{:}Ga}) \sim 1.5$, but the CdSe:In results are for compensated n-type samples while the Ge:Ga results are for weakly compensated p-type samples.

Mansfield[28] states that close to n_c, hopping processes are complicated by correlated many-electron hopping effects (CMEH) above a certain impurity density N_s. The complex subject of CMEH has been discussed in detail by Pollak and co-workers,[29] but their discussion seems most relevant to medium and heavier compensation. The Mansfield criterion for CMEH effects involves the Coulomb interaction $E_{ij} = e^2/\epsilon'(n,R_{ij})R_{ij}$ for hopping between two sites separated by R_{ij}. Equating $2R_{ij}/a$ and E_{ij}/kT Mansfield obtains $N_s = (3/4\pi)(2\epsilon' kT/e^2 a)^{3/2}$ and finds $N_s < 0.01 n_c$ for Si and Ge at 2K using $\epsilon' = \epsilon_h$ (the host values of ϵ_h at LT are 11.4 for Si and 15.4 for Ge). When one employs $\epsilon' \sim \epsilon(n,T)$ in the critical regime N_s can increase by orders of magnitude, however one also needs to take account of spatial dispersion of $\epsilon'(n,R,T)$ in calculating N_s. For small R the screening is that of the host so that the Hubbard U for two electrons on the same site is large, while the long-range part of the Coulomb interaction is fully screened by $\epsilon(n,T)$ which diverges as $n \to n_c$. The role of $e-e$ interactions and CMEH depend strongly on the hopping distance R_h and on T. The viewpoint adopted in this paper is pragmatic, namely, when the data suggest Mott VRH it will be assumed $\sigma(n,T)$ is dominated by one-electron hops. When the data suggest ES VRH is dominant, then $e-e$ interactions are clearly important. The Mansfield criterion is qualitatively consistent with a crossover from Mott VRH to ES VRH as T is lowered, but there are other criteria based on the width of the Coulomb gap. The extension of VRH theory discussed below doesn't include CMEH effects.

VARIABLE RANGE HOPPING IN THE CRITICAL REGIME

The first step in this calculation is the determination of the impedance between two sites. MA found the impedance between sites i and j, separated by the distance R_{ij}, for isotropic envelope functions, but neglecting contributions from excited states because $\Delta_{ex}/kT \gg 1$, to be

$$Z_{g,ij}(Ba^4/R_{ij}^2)\exp[2R_{ij}/a]h_{ij}h_g, \tag{1a}$$

where

$$h_{ij} = \{1+\exp[-\beta(E_i-\mu)]\}\{1+\exp[-\beta(E_j-\mu)]\}, \tag{1b}$$

and

$$h_g = (1/\beta\Delta)\{\exp\beta(E_i-\mu) - \exp\beta(E_j-\mu)\}, \tag{1c}$$

and

$$B = (9n_v/4)(\rho_0 s^5 \hbar^4/E_1^2)(\epsilon_h^2/e^6). \tag{1d}$$

$E_i = E_j + \Delta(\Delta > 0)$ with μ the chemical potential. ρ_0 is the crystal density, s the velocity of sound, E_1 a deformation potential, ϵ_h the host dielectric constant, a the donor Bohr radius for a spherical envelope function, and n_v is the number of equivalent conduction-band valleys. After some algebra it can be shown

$$h_{ij}h_g = (4/\beta\Delta)\sinh(\beta\Delta/2)[\cosh\beta(E_j-\mu+\Delta/2) + \cosh(\beta\Delta/2]. \tag{2}$$

The states available for hopping lie within the order of kT of μ. For simplicity we will minimize $Z_{g,ij}$ for two cases: case 1 for $E_j = \mu$; case 2 for $E_j = \mu - \Delta/2$. The choice makes no difference for $\beta\Delta \gg 1$, but does change the result in a numerical way for $\beta\Delta \ll 1$. The MA result in Eq. (1a) is valid in the dilute limit where the resonance integral $\langle W\rangle = (2e^2/3\epsilon_h a^2)\mathrm{Re}^{-R/a}$ and $Z_g \propto \langle W\rangle^{-2}$ and the subscript ij is now dropped on R_{ij}. The crucial feature in the critical regime is the strong spatial dispersion of the dielectric response $\epsilon'(n,r,T)$ given by

$$\epsilon'(n,r,T)^{-1} = [\epsilon_h^{-1} - \epsilon(n,T)^{-1}]\exp(-r/r_a) + \epsilon(n,T)^{-1}, \tag{3}$$

where ϵ_h is the host dielectric constant, $\epsilon(n,T\to 0) = \epsilon_h + 4\pi\chi'(n)$ where $\chi'(n)$ diverges as $n \to n_c$. Significant VRH for n just below n_c implies $\epsilon(n,T)/\epsilon(n,0)$ can be much larger than one. r_s is the screening length and $r_s = (3/4\pi kn)^{1/3}$ where $k\sim 4$. This form in Eq. (3) has been used earlier by Haken[30] in treating the interaction of excitons with optical phonons and is also similar to the interaction term employed for polarons in ionic crystals because of the $[\epsilon_h^{-1} - \epsilon(n)^{-1}]$ factor. It has also been employed in a donor polarizability enhancement[13] calculation as $n \to n_{c-}$. For $r \ll r_s$, $\epsilon'(n,r,T) \to \epsilon_h$ while for $r \gg r_s$, $\epsilon'(n,r,T) \to \epsilon(n,T)$. This implies the localized wave function is characterized by a Bohr radius $a = a_B\epsilon_h(m/m^*)$ for $r \ll r_s$ and a Bohr radius of order the localization length $\xi(n) = a_B\epsilon(n,T \to 0)(m/m^*)$ for $r \gg r_s$ suggesting a wave function

$$\psi(r,r_s,a,\xi) = c(\pi a^3)^{-1/2}e^{-r/a} + (a/r_s)^{3/2}(\pi\xi^3)^{-1/2}e^{-r/\xi}, \tag{4}$$

where c is a coefficient determined by normalization. The coefficient $(a/r_s)^{3/2}$ is an approximate coefficient estimated from the probability the electron is inside a sphere of radius r_s. For $a/r_s \leq 1/3.2 (a/r_s)^3 \leq 0.03$, which guaranties most of the probability density of $|\psi|^2$ lies within r_s. The coefficient $c = -(a/r_s)^{3/2}\lambda + [1-(a/r_s)^3(1-\lambda)]^{1/2}$ where $\lambda = 8/[(\xi/a)^{1/2} + (a/\xi)^{1/2}]^3$. In the critical regime $c > 0.9$ and $c \to 1$ as $\xi(n) \to \infty$ as $n \to n_{c-}$. The matrix elements $L(R), J(R)$ and overlap are calculated in the appendix. There it is shown $S(R) \leq 0.06$ in the relevant regime and that $J(R) \ll L(R)$. The MA expression $W = L - SJ$ in the critical

regime becomes $W\sim L$. The largest term in L results from cross terms between inner and outer portions of $\psi(n,r)$. One finds

$$W\sim L_{ct}\approx\{8e^2a^2[\epsilon_h^{-1}-\epsilon(n)^{-1}]/(r_s\xi)^{3/2}(1+a/r_s)^2\}\times e^{-R/\xi}[1-2a^2/R\xi]. \tag{5}$$

There are three other smaller contributions to L_{ct} proportional to $e^{-R(1/\xi+1/rs)}$, $e^{-R/a}$, and $e^{-R(1/a+1/rs)}$, respectively, that are much smaller than the contribution in Eq. (5). In the critical regime where $\epsilon(n)\gg\epsilon_h$, the various lengths satisfy $a<r_s<R_h(n,T)<\xi(n)$ with $a\ll R_h$ and ξ. The quantity $2a^2/R\xi\leq 0.025$. *In the critical regime $W(R)\propto e^{-R/\xi}$ with no significant R dependence in the prefactor. This crucial result removes any significant T dependence of the prefactor* $\sigma_0(n)$. The MA result $W\propto[e^2R/\epsilon_h a^2]e^{-R/a}$ is removed in the critical regime because ϵ_h is replaced by $\epsilon(n)$ and $\epsilon(n)\gg\epsilon_h$. Spatial dispersion of $\epsilon(n,r)$ introduces the screening length r_s and suggests the more complex $\psi(r,r_s,a,\xi)$ in Eq. (4). As shown in the appendix this simple form of spatial dielectric dispersion changes the leading term to $\langle W\rangle^2\sim\{8ce^2/[\epsilon_h(r_s\xi)^{3/2}(1+a/r_s)^2]\}^2e^{-2R/\xi}$ when $\epsilon(n,T)\gg\epsilon_h$. The MA matrix element for phonon absorption is given by

$$H'=iE_1(\hbar q n_q/2\rho_0 Vs)\langle(W/\Delta)(e^{i\mathbf{q}\cdot\mathbf{Rb}}-e^{i\mathbf{q}\cdot\mathbf{Ra}})\langle b|e^{i\mathbf{q}\cdot\mathbf{rb}}|b\rangle+[1-(W/\Delta)^2]S(\mathbf{q})\rangle, \tag{6}$$

where the overlap matrix element $S(\mathbf{q})$ is given in the appendix. The calculation of $S_{ab}(\mathbf{q})$ is tedious with numerous terms, but the results show that, as in the dilute case, W/Δ is much larger than $[1-(W/\Delta)^2]|S(\mathbf{q})|$. This conclusion is contrary to that in SKC because the earlier consideration neglected dielectric dispersion and didn't use the $\psi(r,a,\xi)$ in Eq. (4). The modified MA impedance, neglecting the conventional MA terms in the regime $\epsilon(n)\gg\epsilon_h$ becomes

$$Z_g(R,\Delta,T)=[(B_{cr}r_s^3\xi^3/a^4)(1+a/r_s)^4]e^{2R/\xi}f(z),$$

$$\text{case \#1:}\quad f(z)=\sinh(2z)/2z,$$

$$\text{case \#2:}\quad f(z)=(\sinh z+1/2\sinh 2z)/2z, \tag{7}$$

where $z=\beta\Delta/2$ and $B_{cr}=(256c^2/9)B$. The quantity $z_m=\Delta_m/2kT$ obtained after minimization of the MA impedance is the crossover parameter between low-T Mott VRH with $z_m>3$ and the new case of high-T Mott VRH encountered in the critical regime where $z_m<1$.

A. The Mott VRH case

Mott used the ansatz $\Delta=3/4\pi R^3N(E_F)$ and minimized the exponent of Z (or $\sigma\propto(1/Z)$) and obtained values of R_m/ξ and Δ_m. In the critical regime where the exponent becomes small ($R_m/\xi<1$ and $\Delta_m/kT<1$) one must minimize $Z_g(R,\Delta,T)$ with respect to R, but still employing the Mott ansatz for $\Delta(R)$. For the two cases the minimization leads to

$$(R_m/\xi)^4=[9/16\pi N(E_F)a^3kT]g(z_m),$$

$$\text{case \#1:}\quad g(z_m)=(\tanh z_m+\coth z_m-1/z_m),$$

$$\text{case \#2:}\quad g(z_m)=(\tanh(z_m/2)+\coth z_m-1/z_m). \tag{8}$$

In both cases for $z_m\gg 1\, g(z_m)\rightarrow 2$ and one obtains Mott's results with $R_m/\xi=(3/8)(T_0/T)^{1/4}$, $\Delta_m=1/4k(T^3T_0)^{1/4}$, and $T_0=512/27\pi N(E_F)\xi^3k$. In the opposite limit for $z_m\ll 1$ relevant to the critical regime, one obtains $g(z_m)=4z_m/3-16z_m^3/45+\cdots$ for case #1 and $g(z_m)=5z_m/6-23z_m^3/360+\cdots$ for case #2. Keeping only the linear term and using the Mott ansatz for $z=\beta\Delta(R)/2$ one finds

$$(R_m/\xi)^{3.5}=[3/4\sqrt{2}\pi N(E_F)\xi^3kT]. \tag{9}$$

For $z_m\ll 1$ $f(z_m)\sim 1$ for both cases #1 and #2 and $Z_{g,m}\propto\exp[2R_m/\xi]$, which leads to a smaller value of the numerical coefficient D in $T_0=D/N(E_F)\xi^3k$ ($6/\pi$ for case #1 and $4.74/\pi$ for case #2), or more than an order of magnitude smaller than the T_0 for $z_m>1$. It is easily shown that for $z_m<1$ case #2 yields the minimum value of D. The Mott exponent of 1/4 has changed to 2/7 (14% increase). The quantity Δ_m differs from the conventional Mott case and is given by

$$\Delta_m=3/4\pi N(E_F)R_m^3=(6/\pi D)kT_0^{1/7}T^{6/7}, \tag{10}$$

where the coefficient $6/\pi D=1.2648$ (case #2) which is five times the case for $z_m\gg 1$. This difference is important and increases the magnitude of the phonon wave number $q=\Delta/hc_s$ by a large factor (5) compared to the usual Mott case. In addition, in the regime $z_m<1$ $R_m/\xi(n)$ and Δ_m/kT have different T dependencies, however, only the former effects the T dependence of $Z(R_m)$. The prefactor of $\sigma=1/Z_{g,m}L_c$ is independent of T assuming the characteristic length L_c is independent of T. Since $1/Z_{g,m}$ can be viewed as the critical conductance G_0, the characteristic length L_c is a macroscopic length determined by sample dimensions. A substantial body of experimental data (Refs. 16–18 and 20–22) supports the notion $\sigma_0(n)$ is independent of T. From Eq. (6) the density dependence is $\sigma_0(n)\propto\xi(n)^{-3}$ as long as $1-\epsilon_h/\epsilon(n,T)\approx 1$.

B. The Efros-Shklovski VRH case

One starts with the same MA expression for $Z_g(R,\Delta,T)$, but uses the ES result $\Delta=e^2/\epsilon(n,R)R$ for the energy difference, namely,

$$\Delta(R)=(e^2/R)\{[\epsilon_h^{-1}-\epsilon(n,T)^{-1}]\exp-R/r_s+\epsilon(n,T)^{-1}\}, \tag{11}$$

where a T dependence has been included since in the hopping regime the low-frequency dielectric response is strongly T dependent until the hopping is frozen out. The minimization is complicated by the spatial dispersion and for $\epsilon(n,T)\gg\epsilon_h$ it might appear that only the first term is important, but this depends critically on the magnitude of R/r_s. $dZ_g/dR=0$ yields for case #1

$$(R_m/\xi)^2=[e^2/4\epsilon(n,T)\xi kT]\{1+[\epsilon(n,T)/\epsilon_h-1]\times(1+R_m/r_s)\exp-R_m/r_s\}\times[\tanh z_m+\coth z_m-1/z_m]. \tag{12}$$

For $z_m\gg 1$ (both cases #1 and #2) and negligible dispersion $[\exp-R_m/r_s(1+R_m/r_s)\epsilon(n,T)/\epsilon_h\ll 1]$ one obtains a standard result $R_m/\xi=[e^2/2\epsilon(n,T)akT]^{1/2}$. This leads to the ES

VRH results with $R_m/\xi = 1/4(T_0'/T)^{1/2}$, $\Delta_m = 1/2k(TT_0')^{1/2}$, and $T_0' = [8e^2/\epsilon(n)\xi k]$. The numerical coefficient in T_0' is $2\sqrt{2}$ larger than that given by ES and gives a smaller ratio of T_0/T_0' that seems in better agreement with the experimental ratios. AZS also obtained the numerical factor 8 for T_0'. When spatial dispersion is crucial $\{[\epsilon(n,T)/\epsilon_h - 1]\exp -R_m/r_s(1+R_m/r_s) \gg 1\}$ where R_m/r_s is near unity $[\exp -R_m/r_s(1+R_m/r_s) \sim [1-1/2(R_m/r_s)^2 - 1/3(R_m/r_s)^3 + \cdots]$ one obtains the results for ES VRH but with $\epsilon(n,T)$ replaced by ϵ_h. The result $T_0' = 8e^2/\epsilon_h \xi k$ is highly relevant to the Ge:Ga results of Watanabe *et al.* In addition the Coulomb gap width Δ_{CG} is strongly affected by the spatial dispersion of $\epsilon(n,R)$. One finds $\Delta_{CG}(R\sim r_s) \gg \Delta_{CG}(R \gg r_s)$ because of the spatial dispersion of $\epsilon(n,R)$. ES VRH depends critically on $\epsilon(n,R_m,T)$. It is also well known from many experimental studies of $\epsilon(n,T)$ that ultralow temperatures are required close to n_c to freeze out hopping and achieve asymptotic values of $\epsilon(n,T\rightarrow 0)$. One specific example is the data of Katsumoto[31] for $Ga_{0.7}Al_{0.3}As$:Si showing $\epsilon(n,T) = \epsilon(n,0)[1 + aT^2 + bT^4 + \cdots]$. This behavior can be approximated by $\epsilon(n,T) = \epsilon(n,0)\cosh[kT/E_c(n)]$ where $E_c(n)$ is a characteristic energy that becomes small as $n \rightarrow n_{c-}$. It is quite likely that $\epsilon(n,T)$ is larger than $\epsilon(n,0)$ in the ES VRH regime, however if spatial dispersion is dominant, then it is ϵ_h that enters the ES parameter T_0'. This is one of the important differences between Mott VRH, which is independent of $\epsilon(n,T)$, and ES VRH.

C. The relation between the impedance $Z(R)$ and the macroscopic conductivity and resistivity

The above extension of VRH to the critical regime is based on the MA impedance $Z_g(R,\Delta)$ between a pair of sites separated by R and differing in energy by Δ. Minimization of $Z_g(R,\Delta)$ rather than just the exponential terms is a valid procedure for the entire range of $z = \Delta/2kT$ and is necessary when $z<1$. The shape of $Z(R)$ and how it changes as $n \rightarrow n_{c-}$, is important, particularly because fluctuations in the hopping length about R_m can become important near n_c. Figure 4 shows $Z(R)/Z(R_m)$, based on Eq. (7) [case #1] and the Mott ansatz, versus R/R_m for three T_0 values. The results show $Z(R)$ is asymmetrical and it rises very rapidly for $R < R_m$. There is a strong drop in curvature for $R \geq R_m$ as $T_0 \rightarrow 0$. For small T_0 value hops with $2R_m$ and $3R_m$ have only slightly larger values of $Z(R)$. The larger range of R values with nearly the same $Z(R)$ values for $R > R_m$, suggests values of $qR_{\rm eff}$ will be substantially larger than the qR_m listed in Table II. This helps to explain why $\langle \sin(qR)/qR\rangle$ is less than for a single value for $R=R_m$. The curvature for case #1 yields

$$d^2Z/dR^2]_{Rm} = Z(R_m)\{8/R_m\xi + (9+R_m^2) \times [1-(2z_m/\sinh 2z_m)^2]\}. \quad (13)$$

In the regime where z_m is enough smaller than $1[1-(2z_m/\sinh 2z_m)^2] \rightarrow 4z_m^2/9$ and using $R_m/\xi = 1/2(T_0/T)^{2/7}$, one finds

$$[d^2Z/dR^2 2]_{Rm} \sim 46Z(R_m)/R_m\xi \propto Z(R_m)(T/T_0)^{2/7}/\xi^2.$$

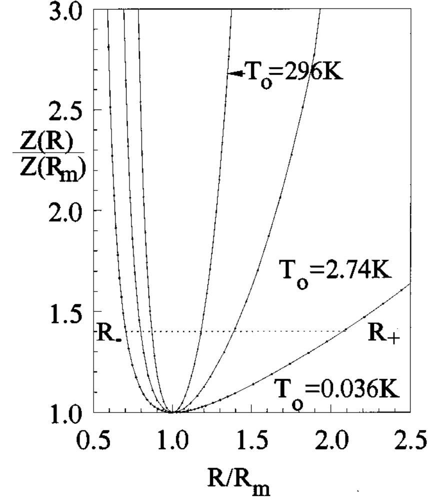

FIG. 4. The normalized impedance $Z(R)/Z(R_m)$ versus R/R_m based on the Mott ansatz at $T=1$ K for Si:As samples with $T_0 = 0.036$, 2.74, and 296 K, respectively. As $n \rightarrow n_{c-}(T_0 \rightarrow 0)$ the curvature for $R>R_m$ becomes dramatically smaller. $R_+(Z) > R_m$ and $R_-(Z) < R_m$ are shown for $T_0 = 0.036$ K.

$[1/Z(R_m)][d^2Z/dR^2]_{Rm}$ scales to zero as $n \rightarrow n_{c-}$ and as $(1-n/n_c)^{8\nu/7}$ and for fixed n exhibits a $T^{2/7}$ T dependence. The correction to Z from fluctuations about R_m is $[1/2d^2Z/dR^2]_{Rm}\langle (R-R_m)^2\rangle$. If $\langle (R-R_m)^2\rangle \propto dR_m^2$ [d = const] the correction will be negligible as $n \rightarrow n_{c-}$ since $R_m/\xi \rightarrow 0$. But if $\langle (R-R_m)^2\rangle \propto R_m\xi$ then the fluctuation correction is exactly proportional to $Z(R_m)$ and will not lead to different T or n dependencies for the total $Z(R)$. The magnitude of $qR_{\rm eff}$ can be increased considerably and will reduce the effects of the $\langle \sin(qR)/qR\rangle$ term. Figure 4 and the results from Eq. (13) are based on the Mott ansatz. This ansatz will breakdown for too small a R, but remains valid for R somewhat larger than r_s. In percolation analysis one seeks the critical value of Z, namely, Z_c assuming that there are values of Z both larger and smaller than Z_c. In the present calculation there are no values of $Z(R)$ less than $Z(R_m)$. The present analysis, based on the MA approach, makes no assumptions about the nature of the resistor network or about the probability distribution $p(Z)$. The carriers choose paths that minimize $Z(R)_{\rm total}$ and maximize σ.

Using $p(Z)dZ \sim 4\pi nR^2 dR$ leads to a divergent $p(Z)$ as $Z \rightarrow Z(R_m)$ since $dZ/dR \rightarrow 0$ at R_m. For a given $Z(R), p(Z)$ can be calculated from the result

$$p(Z) = d/dZ \int_{zm}^{z} p(Z')dZ' = (n/N_s) d/dZ \left[\int_{R_-}^{R_+} F(r) 4\pi r^2 dr \right], \quad (14)$$

where $F(r)$ is the fraction of localized electrons within $2\Delta(r)$ of E_F given approximately by $2\Delta(r)/[E_F - E_D(n)]$ where $\Delta(r)$ is given by the Mott ansatz and $E_D(n)$ is the donor binding energy. N_s is the total number of donors and $N_s = nV_s$ with V_s the volume of the sample. $R_+(Z)$ and $R_-(Z)$ are the two values of R corresponding to a specific value of $Z>Z_m$ (see Fig. 4). In the quadratic regime near $Z_m[0.95R_m < R < 1.05R_m]$ $p(Z) \propto (Z/Z_m - 1)^{-1/2}$. The integral in Eq. (14) yields $\ln(R_+/R_-)$. For larger values of Z/Z_m

TABLE I. The prefactors $\sigma_0(n,T)$, T_0, and T_0'.

	$z_m>1$	$z_m<1$
Mott	$\sigma_0 \propto (1/\xi^3 L_c)(T_0/T)^{1/4}$ $T_0 = 18.1/N(E_F)\xi^3 k$	$\sigma_0 \propto (1/\xi^3 L_c)$ $T_0 = D/N(E_F)\xi^3 k$, $1.51<D<1.9$
Efros-Shklovskii	$\sigma_0(1/\xi^3 L_c)(T_0'/T)^{1/2}$ $T_0' = 8e^2/\epsilon(n)\xi k$ (ND) $T_0' = 8e^2/\epsilon_h \xi k$ (D)	not physically relevant
Finite-T scaling	$AT^{x-py}\phi^p \exp[-b\phi^q/T^{qy}]$ $\phi = (n_c - n)$; $T_0 = b(1-n/n_c)^q$	Mott $qy = 1/4$, $x-py=0$, $p=1/2$ ES $qy=1/2$, $x-py=-1/2$, $p=3/2$

$$p(Z) \propto Z^{-1}\{\xi/2R_+ + (3x)^{-1}[2z_m - x^{-1} - (2R_m/\xi x^{4/3})]^{-1}\}, \tag{15a}$$

and

$$R_+(Z) = R_m + (\xi/2)[\ln(\sinh(2z_m)/2z_m) + \ln(Z/Z_m)], \tag{15b}$$

and

$$x = (R_m/R_-)^3, Z/Z^m = \exp[2(R-R_m)/\xi] \times [\sinh 2z_m x / x \sinh 2z_m]. \tag{15c}$$

The first term in the {} in Eq. (15a) decreases with increasing Z since R_+ increases with Z as $\ln(Z/Z_m)$. The second term in the {} decreases more rapidly with increasing Z. The first term dominates for larger Z/Z_m and the overall result is that $p(Z)$ decreases less rapidly than Z^{-2} for Z larger than a characteristic value than depends on T_0 [for $T_0 = 0.036$ K $Z_c \sim 2.0 Z_m$. This is a problem for $\xi > 2R_m$, namely, for the smallest values of T_0 such as the $T_0 = 0.036$ K curve in Fig. 4. The macroscopic conductivity is given by $\langle\sigma\rangle = L^{-1}\int_{Zm}^{Zs}[p(Z)/Z]dZ$, where Z_s is determined by the sample volume V_s, since $p(Z)$ is sharply peaked for small $Z/Z_m - 1$ and continues to drop faster than $Z^{-1}\langle\sigma\rangle \approx (LZ_m)^{-1}$. On the other hand, the macroscopic resistivity $\langle\rho\rangle = L\int_{Zm}^{Zs} Zp(Z)dZ$ is more complicated. The quadratic region with the sharply peaked $p(Z)$ provides a contribution kLZ_m [$k<1$ and dependent on the normalization of $p(Z)$]. However, there can be a contribution from the tail of $p(Z)$ for $Z > 2Z_m$ since $Zp(Z)$ doesn't decrease faster than $Z^{-t}(t > 1)$ with increasing Z. The tail contribution leads to $\langle\rho\rangle$ significantly larger than LZ_m and yields $\langle\sigma\rangle\langle\rho\rangle$ larger than one. The tail of $p(Z)$ depends sensitively on the asymptotic behavior of $\psi(r)$ [see Eq. (4)] for $r > R_m$. If $\psi(r) \propto \exp[-r/(R_m\xi)^{1/2}]$ for $r > R_m$ then $p(Z)$ would fall off faster for $Z > 2Z_m$ and $\langle\sigma\rangle\langle\rho\rangle$ would be closer to one. This is plausible since $\psi \propto \exp[ik_F r]$ where $k_F = [2m^*(E_F - E_c)/\hbar^2]^{1/2}$. For $n < n_c$ $E_F < E_c$ and $k_F \propto i(E_c - E_F)^{1/2} \propto i/\xi(n)^{1/2}$ leading to $\psi(r) \propto \exp[-r/(b\xi)^{1/2}]$ since $\xi(n) \propto [E_F - E_c]^{-\nu}$ and $\nu \approx 1.0$. The experimental data determine $\langle\rho\rangle$ and $\langle\sigma\rangle$ was obtained from $\langle\rho\rangle = \langle\sigma\rangle^{-1}$. While the data is in apparent agreement with $\langle\rho\rangle \propto LZ_m$ the result $\langle\sigma\rangle\langle\rho\rangle \approx 1$ has not been established experimentally. The asymptotic behavior of $\psi(r)$ for $r > R_m$ [and $p(Z)$ for $Z > 2Z_m$] remains a large uncertainty in the present treatment when $\xi \gg R_m$. In addition, the $p(Z)$ given by Eq. (15a) and the $Z(R)$ in Fig. 4 are based on the Mott ansatz (a continuum approach), which for a given $N(E_F)$ yields a unique Δ for a given R. This ignores the detailed random potential and the statistical distribution of donors.

The critical conductance distribution $p_c(g)$ has been calculated[32] numerically for the Anderson MIT. Slevin and Ohtsuki[32] demonstrated $p_c(g)$ decreases rapidly for small $g(g<0.02)$. Knowledge of $p_c(g)$ permits a determination of $p_c(Z) = p_c(g)dg/dZ \propto p_c(g)/Z^2$. For $g<0.02$, $p_c(Z)$ drops faster than Z^{-3} and removes any contribution to $\langle\rho\rangle$ from the large Z tail. This suggests when the random Anderson potential is properly incorporated into $Z(R)$ for $R > R_m$ one should obtain for $1 - n/n_c \ll 1$ a $p(Z)$ that should drop more rapidly than Eq. (15a).

The second moment of $Z(R)$ is related to $\langle(R-R_m)^2\rangle$ and this quantity can be calculated with this simple $p(Z)$. The integration is difficult since dZ/dR is complicated and changes sign. $t>1$ is required for a convergent second moment. An estimate for $t=2$

$$\langle(R-R_m)^2\rangle \sim \xi^2/2[1 - h(R_m/\xi)], \tag{16}$$

where $h(R_m/\xi)$ is a complicated function that increases toward 1 as R_m/ξ increases. Because of the asymmetry of $Z(R)$, Eq. (16) demonstrates the effective hopping length ($R_{\text{eff}} > R_m$) in the limit $\xi \gg R_m$ must remain less than $\xi/\sqrt{2}$ depending on the magnitude of $h(R_m/\xi)$.

D. Summary of prefactor $\sigma_0(n,T)$, T_0, and T_0'

The prefactor for ES VRH for $z_m \gg 1$ for both cases #1 and #2 takes the form $\sigma_0 \propto [1/\xi(n)^3 L_c](T_0'/T)^{1/2}$ (Table 1). The $T^{-1/2}$ of the prefactor agrees with that obtained by ES. Finite T-scaling of the form $\sigma(n,T) = AT^x f(|n-n_c|/T^y)$ has been used recently in the analysis of experimental data and its predictions for the prefactor of Mott and ES VRH should be examined. It yields a prefactor of the form $\sigma_0(n,T) = AT^{x-py}(n_c - n)^p$ where the p depends on the functional form of f, which can be determined from the data. The exponential part of f for Mott and ES VRH is different and is not governed by the form of the prefactor. Data from the metallic side for many doped weakly compensated semiconductors supports $x/y = s = 0.5$, which yields $p = 0.5 - m/y$. The different choices of m will determine p and the T and n dependence of the Mott and ES VRH prefactors and one expects in general m will be different for Mott and ES VRH.

TABLE II. Hopping parameters for Si:As samples.

n	$1-n/n_c$	$q\times10^{-5}$ cm^{-1}	qa[a]	qR_m[a]	$q\xi$[a]	R_m/ξ	L_{ct}/Δ_m
8.48	0.014	1.90	0.029	0.47	2.44	0.174	0.61
8.07	0.06	3.53	0.054	0.60	0.906	0.514	1.45
7.90	0.081	4.40	0.068	0.87	0.837	0.764	0.53
7.57	0.12	6.88	0.106	2.28	0.89	1.66	0.39

[a]Based on $a=15.4$ Å, $\xi_0=15.4$ Å, and $T=1$ K.

DISCUSSION

Before attempting to compare the experimental results with the theoretical expressions one should first compare the results with possible predictions of fTs. There are two classes of MIT systems, namely, the weakly compensated doped Si and Ge systems that show $\sigma(n>n_c,T)=\sigma_0(n/n_c-1)^s+m(n)T^{1/2}$, where $s\sim0.5$ and $m(n)$ gets large as $n\rightarrow n_{c-}$, and the amorphous semiconductor-metal alloys where $s\sim=1.0$ and $m(n)$ is nearly a universal constant for systems like Si:Nb. The former can be explained with $x=y/2\sim0.27$ and $s=x/y$. If one assumes the prefactor is independent of T then this requires $p=1/2$ and leads to $\sigma_0\alpha(n_c-n)^{1/2}$. However, the experimental results in Fig. 2 for σ_0 for Si:P and Si:As show no appreciable n dependence for $1-n/n_c<0.05$. If instead one sets $p=0$ to remove the n dependence one finds $\sigma_0\propto T^{1/4}$. This is inconsistent with the experimental results that would support a T^s prefactor with $s<0.04$ for Si:As and $s<0.02$ for Si:P. The Si:As and Si:P results appear to be incompatible with fTs (at least with $x\sim0.27$ and $y=2x$). They can be explained by the prefactor given in Table I for Mott VRH for $z<1$ if the characteristic length L_c is independent of T. As $1-n/n_c$ increases one eventually crosses over to $z>1$ and this leads to $\sigma_0\propto(T_0/T)^{1/4}$ for L_c independent of T and slowly varying with n. This can be explained with fTs with $x-py=-1/4$ and $p\sim1.0$. The Si:As and Si:P σ_0 are varying more rapidly with $1-n/n_c$, but it should be emphasized that the transition region between $z\ll1$ and $z\gg1$ is broad and one can be in the intermediate regime where T_0 itself (see Fig. 2) is changing much more rapidly than expected from $T_0\propto[N(E_F)\xi^3]^{-1}$ for reasonable values of the localization length exponent $\nu\sim1$. The results in Figs. 2 and 3 for Si:P and Si:As are either in the $z_m<1$ regime or in a broad transition regime. Note that for $T=1$ K, $z=1$ requires $T_0=4096$ K while for $T=0.1$ K, $z=1$ requires $T_0=410$ K. It is much easier to satisfy the condition $z_m>1$ for ES VRH since $z_m=1/4(T_0'/T)^{1/2}$. The Ge:Ga data, remarkably, exhibits ES VRH conduction to about $0.99n_c$. For $1-n/n_c=0.0109$ the Ge:Ga results yield $T_0'\sim4$ K, which gives $z_m=2.5$ at $T=40$ mK and $z_m=0.5$ at $T=1$ K. It is exactly in the regime for $1-n/n_c<0.01$, as seen in Figs. 6 and 7 in Ref. 20, that strong deviations from conventional ES VRH are observed. The two CdSe:In samples with $T^*=0.056$ K ($T^*=T_0'$) $[1-n/n_c=0.0536]$ and $T^*=0.24$ K $[1-n/n_c=0.10]$ only yield $z=1$ for $T=3.5$ and 15 mK, respectively, and the data were obtained in the regime where $z<1$, where corrections to conventional ES VRH are expected.

A. Magnitude of VRH parameters near n_c

Comparison of the theory with the experimental results requires a knowledge of the relative magnitude of the terms in Eq. (6). These parameters are shown for Si:As for four samples from SKC. The magnitude of qR_m is substantially larger than found in SKC because of the larger coefficient of q (factor of 5.1). qR_m and $q\xi$ in Table II are the minimum values based on $\xi_0=a^*$. If $\xi_0>a^*$ these values increase correspondingly. The quantities qa, qR_m, and $q\xi$ all get smaller as T is reduced and this causes $S(\mathbf{q})$ to get larger at lower T, unless $e^{-Rm/\xi}$ gets smaller more rapidly than the increase in the other terms in Eq. (A7). However, the ratio L_{ct}/Δ_m also gets larger with reduced T as $e^{-Rm/\xi}/T^{6/7}$ since $\Delta_m\propto T^{6/7}$. Unlike the dilute case, the ratio $W/\Delta\sim L_{ct}/\Delta$ can approach unity at $T=1$ K and is even larger at $T\sim0.1$ K for samples close enough to n_c. The crucial point is that W/Δ is always much larger than $[1-(W/\Delta)^2]S(\mathbf{q})$, and the term neglected by MA can also safely be neglected in the critical regime. In addition the angular average of the first term $\langle|H'|^2\rangle$ in Eq. (6) yields $\langle(2-2\cos\mathbf{q}\cdot\mathbf{R})\rangle=2[1-\sin(qR)/qR]$. In the dilute limit $qR_m\gg1$ and the oscillatory term could be ignored. In the critical regime qR_m becomes smaller than one for $T<1$ K for the 8.48 sample. In the limit that $qR_m\ll1$, $[1-\sin(qR_m)/qR_m]\propto(qR_m)^2$. The extra R_m^{-2} dependence in $Z(R_m)$ changes the theoretical results in Sec. III and leads to a much stronger T dependence of the prefactor $\sigma_0(n,T)$ than is observed in the data. The actual conducting path is not one of a series of hops of length R_m, but includes fluctuations about R_m. These fluctuations will produce an averaging of $\sin(qR)/qR$ that may make the oscillatory term unimportant, but this remains an issue. Also, in the regime $1-n/n_c<0.02$, where the theory may need corrections, the role of doping inhomogeneity becomes important. One expects the average over a doping distribution $\langle qR_m\rangle$ to be larger than qR_m, although it is difficult to estimate the magnitude of this effect. The Si:P data[17] in Fig. 1 appears to show a good fit to Mott VRH down to $T\sim0.1$ K. The data is a reasonable fit to Eq. (7) for $T>0.1$ K and $1-n/n_c>0.01$.

B. Comparison of localization length exponent and metallic scaling exponent

Returning to the Si:As and Si:P results in Fig. 2 one can compare the localization length exponent ν' obtained from $T_0^{1/3}\propto[N(E_F)]^{-1/3}\xi(n)^{-1}\sim[N(E_F,n_c)]^{-1/3}\xi_0^{-1}(1-n/n_c)^{\nu'}$ with the conductivity exponent s obtained from $\sigma(n>n_c,T\rightarrow0)=\sigma_0(n/n_c-1)^s$ (a new scaling expression[33] for the Boltzmann conductivity obtained from standard scattering theory and Anderson localization concepts yields $s=1/2$ and suggests the possibility $\nu'=2s$). Figure 5 shows $T_0^{1/3}$ for Si:As and Si:P versus $1-n/n_c$ for several different values of n_c. The finer grid of samples for Si:P yields nearly linear behavior for $n_c\sim3.75$ for $1-n/n_c<0.05$ and a steeper slope for $1-n/n_c>0.05$. The Si:As results (with a smaller density of points and more scatter) also shows approximate linear behavior for $n_c=8.60$, but only three-plus points are in the linear range. Both the Si:P and Si:As results show a significant drop in ν' as n_c is decreased. The scaling of $\sigma(n>n_c,T\rightarrow0)\propto(n/n_c-1)^s$, where the new scaling expression suggests $\nu'\sim2s$, and the experimental results for weakly

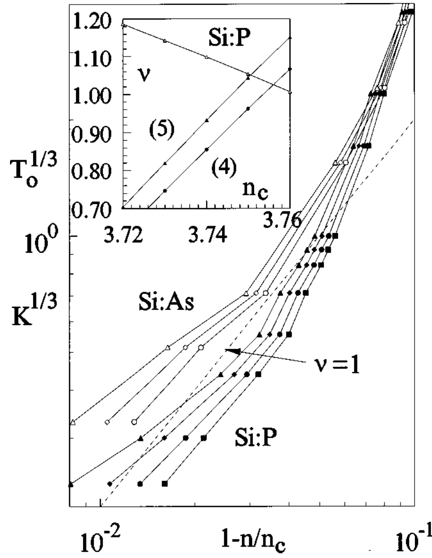

FIG. 5. $T_0^{1/3}$ versus $1-n/n_c$ for Si:P [3.73 (▲), 3.74 (◆), 3.75 (●), 3.76 (■)] and Si:As [8.57 (△), 8.58 (◇), 8.59 (○)] for several values of n_c. The dashed line $T_0^{1/3}\propto\xi(n)^{-1}\propto(1-n/n_c)$ for $\nu=1.0$ is shown for comparison. The inset shows ν versus n_c from $T_0^{1/3}$ (solid symbols) and $\sigma(n>n_c,T\rightarrow 0)=\sigma_0(n/n_c-1)^s$ (open symbols) based on a forced fit of the data with n_c as a variable. The intersection of the insulator ν and the metallic $2s$ yields a more accurate value of n_c.

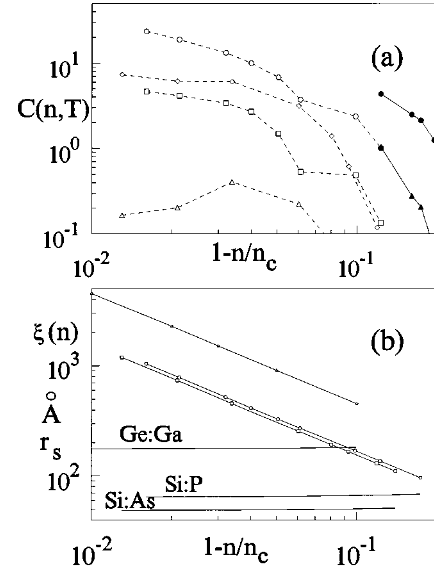

FIG. 6. (a) The correction factor $C(n,T)$ versus $1-n/n_c$ demonstrating the importance of the spatial dispersion of $\epsilon'(n,R,T)$ for Si:P [Mott (○) $T=1$ K, (□) $T=0.1$ K; ES (●) $T=1$ K, (■) $T=0.1$ K] and Si:As [Mott (◇) $T=1$ K, (△) $T=0.1$ K]. $C(n,T)$ drops as $R_m(n,T)$ increases with decreasing T. $C(n,T)$ is larger for Si:P than for Si:As for $1-n/n_c<0.04$. (b) The length $\xi(n)$ versus $1-n/n_c$ for Ge:Ga, Si:P, and Si:As based on ξ_0 equal 45.6, 16.7, and 15.4 Å, respectively. The r_s values are obtained from $r_s=(3\cdot 4.19/4\pi n_c)^{1/3}(n_c/n)^{1/3}$ for n_c equal 0.186, 3.75, and 8.59, respectively.

compensated Si and Ge MIT systems with $s\sim 0.5$ also produce a variation of s with n_c, but in this case s increases as n_c decreases. The opposite dependence of ν' on $T_0^{1/3}$ and s from $\sigma(n>n_c,T\rightarrow 0)$ and second-order phase transition theory suggests the plot in the inset in Fig. 5. Both the $T_0^{1/3}$ data and the $\sigma(n>n_c,T\rightarrow 0)$ data are force fit to a power law $[|n/n_c-1|^{\nu'}$ or $(n/n_c-1)^s]$ with n_c as a floating parameter. The functions $\nu'=F(n_c)$ and $2s=G(n_c)$ are shown in the inset for Si:P. Two straight lines of ν' versus n_c are shown for the 5 and 4 samples nearest n_c, respectively. An analysis of the Si:P results of Blaschette *et al.*[34] (10 metallic samples in the range $n=3.85$ to 4.37) and the four smallest values of T_0 shown in Fig. 5, leads to $n_c=3.756$, $\nu'=2s=1.024$, and $s=0.512$. The forced fit results for the 10 Si:P metallic samples yields standard deviations [for s] that were smaller than for the more limited Si:As results with only five metallic bar samples, but the minimum standard deviation for the Si:P case occurred for $n_c=3.75$, suggesting the best fit of the Blaschette *et al.*, data is for $s\sim 0.5$. The value of $s\sim 1.3$ suggested by Stupp *et al.*, resulted from using $n_c=3.52$, which is 6% below $n_c=3.756$ because the samples 3.54 to 3.69 (see Fig. 1) were identified as metallic rather than insulating. A similar analysis for five metallic Si:As bar samples (SKC $n=8.67$, 8.91, 9.06, 9.50, and 10.4) yield s versus n_c. The two curves intersect for $[\nu'=2s]$ for $n_c=8.58$ for five T_0 points and $n_c=8.592$ for four T_0 points. In both cases the intersection $\nu'=2s$ is close to 1.0 yielding 1.02 for $n_c=8.58$ and 0.98 for $n_c=8.592$. The new theoretical scaling result suggests $s\geqslant 0.5$. This analysis is consistent with the results in SKC (ignoring the Hall disk results), but the use of the criterion $\nu'=2s$ provides a considerably more sensitive determination of n_c.

C. The correction factor $C(n,T)$ for ES VRH

The correction factor $C(n,T)=[\epsilon(n,T\rightarrow 0)/\epsilon_h-1]\{\exp[-R_m(n,T)/r_s][1+R_m(n,T)/r_s]\}$ in Eq. (12) is shown in Fig. 6(a) demonstrating the effect of spatial dispersion of the dielectric response $\epsilon(n,R)$. The factor $\epsilon(n,T\rightarrow 0)/\epsilon_h-1=4\pi\chi'(n)/\epsilon_h$ is obtained from the experimental data for Si:P (Ref. 35) $[4\pi\chi'=7.0(n_c/n-1)^{-1.15}]$ and Si:As (Ref. 36) $[4\pi\chi'=8.0(n_c/n-1)^{-1.20}]$. The values of r_s are obtained from $4\pi r_s^3/3=k/(N_d-N_a)$ for $k=4.19$ $[r_s\sim n_c^{-1/3}]$. $k=1$ corresponds to a volume containing on average one donor while $k=4.19$ corresponds to a volume with 4.19 donors. The precise value of k isn't known but for $R/r_s>3$ the screening is nearly complete [unless $\epsilon(n)>50\epsilon_h$]. $R=3r_s$ corresponds to an average of 113 donors inside $3r_s$. $R_m(n,T)$ has been calculated with the Mott result $R_m=3/8\xi(n)(T_0/T)^{1/4}$ for $z_m(n,T)>1$ and the new HT Mott result $R_m=1/2\xi(n)(T_0/T)^{2/7}$ for $z_m(n,T)<1$ while $R_m=1/4\xi(n)(T_0'/T)^{1/2}$ was used for the ES case. The dashed lines (Si:P, ○$T=1$ K, □$T=0.1$ K; Si:As, ◇$T=1$ K, △$T=0.1$ K) based on the Mott R_m, demonstrate the correction is large close to n_c because of the large values of $4\pi\chi'$. $C(n,T)$ is smaller at 0.1 K than at 1 K because of larger values of $\exp(R_m/r_s)$ at 0.1 K and smaller values of $\exp R_m/r_s$. $C(n,T)$ can be much larger than shown in Fig. 6(a) because of the use of $\epsilon(n,T\rightarrow 0)$ rather than $\epsilon(n,T)$. $\epsilon(n,T)/\epsilon(n,T\rightarrow 0)$ can be large near n_c because it is hard to freeze out the hopping. When $C(n,T)>20-100$, the dielectric response that determines R_m and T_0' for ES VRH is $\epsilon(n,R_m)\sim\epsilon_h\exp(R_m/r_s)/(1+R_m/r_s)$. The solid lines in Fig. 6(a) show $C(n,T)$ based on ES R_m values obtained from the

Si:P data[19] in Fig. 2. The correction shows the same qualitative trends as for the dashed curves for the Mott R_m. However, the values of $C(n,T)$ at the same value of $1-n/n_c$ [0.123 and 0.173] are larger by a factor or 3–5 because of the smaller ES R_m values.

Figure 6(b) shows localization lengths $\xi(n)$ versus $1-n/n_c$ for Ge:Ga, Si:P, and Si:As based on $\xi=a^*$ values of 45.6, 16.7, and 15.4 Å, respectively. The Ge:Ga value is smaller than used in Ref. 20 and is based on $\epsilon_h=15.4$ (see Faulkner[37] and Castner *et al.*[38]) the acceptor binding energy 11.0 meV, and the relation $a^*=a^*_{em}[E_{em}/E_A]^{1/2}$ based on the Whittaker envelope function. If one employed $r_s\sim a^*$, as done in the derivation of the Mott criterion $[n_c^{1/3}a^*\sim 1/4]$, one would obtain $k=\pi/48\sim 0.065$. The values of R_m/r_s would be much too large and $\exp(-R_m/r_s)[1+R_m/r_s]$ would be much too small. There can't be much screening from the neighboring donors for $k\ll 1$. In order for spatial dispersion to be important one must have $r_s/a^*\sim 4$. The values of r_s shown in 6(b) are based on $k=4.19$ and these yield r_s/a^* values of 3.845, 3.85, and 3.17 for Ge:Ga, Si:P, and Si:As, respectively. For $1-n/n_c<0.1$ the values of $C(n,T)$ are larger for Si:P than for Si:As because (1) ξ_0/r_s is larger for Si:As, and (2) for $1-n/n_c<0.05$ $T_0(\mathrm{As})/T_0(P)\sim 3$. However, the values of ξ_0/r_s are virtually identical for Ge:Ga and Si:P. The results in Fig. 5 suggest the origin of the ES VRH in Ge:Ga for $0.9n_c<n<0.99n_c$ results from very large values of $C(n,T)$ (corresponding to large values of the Coulomb gap width) which results from much larger values of $\epsilon(n,T)/\epsilon(n,0)$ at a given T for Ge:Ga than for Si:P. The smaller activation energies for doped Ge than for doped Si make it harder to freeze out the hopping in Ge. At a fixed T the increase in $\epsilon(n,T)/\epsilon(n,0)$ as $n\to n_c$ leads qualitatively to an increase in Δ_{CG} as $n\to n_{c-}$, in agreement with the experimental results of Massey and Lee[25] for Si:B.

Experimental values of $\epsilon(n,T\to 0)$ are not available for Ge:Ga, but based on the results in Fig. 6 it is plausible to assume that $C(n,T)\gg 1$ for $1-n/n_c<0.1$. In this case the relevant dielectric constant to use in Eq. (7) and in T_0' will be much closer to ϵ_h than to $\epsilon(n,T)$. Dielectric studies of n-type Ge demonstrate that even at $N_d\sim 0.44n_c$ temperatures well below 0.1 K are required to freeze out the hopping contributions to $\epsilon(n,T)$. For $0.9n_c<n<0.99n_c$ even 10 mK may not be cold enough to freeze out all the hopping. This suggests $C(n,T)\gg 1$ for the Ge:Ga results and explains why the Watanabe *et al.*, T_0' results are in agreement with $T_0'=2.8(e^2/\epsilon_h\xi_0)[1-n/n_c]^{1.0}$. The Ge:Ga results support $\nu'=1.0$. These results are strikingly different than the ES VRH results for Si:P and CdSe:In in Fig. 2. The latter two systems have $T_0'<T_0$ at the same value of n/n_c and the scaling of T_0' (see Fig. 2) is more rapid than linear in $1-n/n_c$. Zabrodskii and Zinov'eva[39] have observed for Ge:As ES VRH with $T_0'\propto(1-n/n_c)^{2.0}$ which can be interpreted with $\nu'=1.0$ and $4\pi\chi'\propto(1-n/n_c)^{-1}$ although these authors suggested -1.3 for the exponent of $4\pi\chi'$ and $\nu'=0.7$. The Ge:Ga results suggest $T_0'\sim e^2/\epsilon_h\xi(n)$ and no collapse of the Coulomb gap in the range $0.9n_c<n<0.99n_c$. This is explained by strong spatial dispersion of $\epsilon(n,R)$ and a large $C(n,T)$, but the parameters are not well known for Ge:Ga. Since the results for Ge:As are closer to the Si:P and CdSe:In results, the reason for the different behavior of Ge:Ga doesn't depend on a difference between Ge systems and the others in Fig. 2.

D. The minimum in $R_m(n,T=\mathrm{const})$ versus n for Mott VRH

What of the minimum in $R_m(n,T=\mathrm{const})$ versus n? One uses $T_0=[D(n)/N(E_F,n)\xi(n)^3k]$ where $D(n)$ varies rapidly with n in the crossover regime between conventional Mott VRH and HT Mott VRH. The logarithmic derivative of R_m is

$$d\ln R_m/d\ln(n)=1/4[\nu'(1-n/n_c)^{-1} -d\ln N(E_F,n)/d\ln(n) +d\ln D(n)/d\ln(n)]. \tag{17}$$

The minimum in $R_m[dR_m/dn=0]$ at $n\sim 0.95\pm 0.01n_c$ cannot be explained with the density-of-states term because $d\ln N(E_F,n)/d\ln(n)$ is of order unity. $D(n)\to 18.1$ as $n\to 0.8n_c$ and $D(n)\sim 1.51$ as $n\to n_c$. The functional dependence of $D(n)$ isn't known, but it is approximated by an S-shaped curve that is steepest in the region in Fig. 2 where T_0 is changing most rapidly with n, which means at n slightly below the minimum. An estimate $d\ln D/d\ln n$ for the entire range $\Delta n=0.2n_c$ is $(n/D)(\Delta D/\Delta n)\sim-(0.9/9.8)(16.6/0.2)=-7.6$. Thus $d\ln D/d\ln(n)$ need only be 2.5 times larger to account for the minimum. The rapid drop in the coefficient $D(n)$ in the Mott characteristic T_0 in the crossover regime, explains the minimum in $R_m(n,T=\mathrm{const})$ versus n.

E. The effect of doping inhomogeneity on the prefactor $\sigma_0(n)$

In the critical regime where $\epsilon(n,T)\gg\epsilon_h$ and $z_m<1$ the prefactor $\sigma_0(n)\propto L_{ct}^2\propto 1/\xi^3$ with r_s virtually constant $\sigma_0(n)\propto(1-n/n_c)^{3\nu}$, which is reasonable agreement for $1-n/n_c>0.1$, but is in sharp contrast to the results in Fig. 3 showing $\sigma_0(n)$ for Si:P and Si:As approaching a constant as $n\to n_c$. The rapid variation of $\sigma_0(n)$ for $1-n/n_c>0.1$ is faster than the prediction for $\nu=1$, which may result from the increasing importance of the neglected MA terms. Whenever a quantity approaches zero at n_c one must take account of doping inhomogeneity. For a normalized $[\int P(n)dn=1]$ Gaussian doping density distribution $P\propto\exp-[(n-\underline{n})/an_c]^2$, where a is measure of the spread in n about the average density $\underline{n}$. The average $\langle\sigma_0\rangle=\int\sigma_0(n)P(n)dn$, where for $\sigma_0=A(1-n/n_c)^t$ one obtains a result in terms of Hermite polynomials, namely,

$$\langle\sigma_0\rangle=A\sqrt{\pi}[\underline{x}^4+3a^2\underline{x}^2+3a^4/4]\quad\text{for}\quad t=4, \tag{18a}$$

$$\langle\sigma_0\rangle=A\sqrt{\pi}[\underline{x}^3+3a^2\underline{x}/2]\quad\text{for}\quad t=3, \tag{18b}$$

where $\underline{x}=1-\underline{n}/n_c$. The $t=3$ case reduces the dependence of $\langle\sigma_0(n)\rangle$ on $\underline{x}$, but doesn't yield a constant as $\underline{x}\to 0$. The $t=4$ case qualitatively fits the behavior in Fig. 3 reasonably well, but doesn't explain the shallow minimum seen in the Si:P results. Using Eq. (12a) and the Si:As data in Fig. 3 for $0.091<1-\underline{n}/n_c<0.148$ yields $a=0.02$, certainly a plausible estimate of the doping inhomogeneity, but this yields too small a value of $\langle\sigma_0(\underline{n}=n_c)\rangle$. A change in the asymptotic behavior of $\psi(r)$ for $r>R_m$ from $e^{-r/\xi}$ to $\exp[-r/(b\xi)^{1/2}]$ (b

a characteristic length of order r_s) will reduce the exponent t to $t/2$ for $\xi \gg R_m$, which is qualitatively consistent with the Si:P and Si:As data in Fig. 3. The physical significance of $\langle \sigma_0 \rangle$ finite at n_c, is that its magnitude is a measure of the inhomogeneity. For a truly homogeneous system the theory yields $\sigma_0(n) \to 0$ as $n \to n_c$. Some have argued a finite σ_0 as $n \to n_c$ is consistent with a Mott minimum metallic conductivity. The result $\sigma_0(n) \to 0$ is qualitatively consistent with the fTs result, which for $z_m < 1$ has a T-independent prefactor scaling to zero as $n \to n_c$. This inhomogeneity analysis provides a second reason why the Ge:Ga results[21] don't exhibit Mott VRH for $0.9 < n/n_c < 0.99$. The more homogeneous doping achieved with the NTD approach leads to a smaller Mott prefactor $\langle \sigma_0 \rangle$ in the critical regime.

F. Beyond the pair approximation (?) and validity issues when $\Delta_m < kT$

The validity of the pair approximation itself in the critical regime is not obvious and this question is difficult to address quantitatively. When the localization length $\xi(n) \gg d$ close to n_c there are of order $(\xi/d)^3$ donor sites within a localization radius [for $n = 0.98 n_c (\xi/d)^3 > 1000$]. Consider localized eigenstates

$$\psi_a(\mathbf{r} - \mathbf{R}_a, E_a) = \sum_i^m c_{a,i} \phi_i(\mathbf{r} - \mathbf{R}_i)$$

and

$$\psi_b(\mathbf{r} - \mathbf{R}_b, E_b) = \sum_j^m c_{b,j} \phi_j(\mathbf{r} - \mathbf{R}_j)$$

which differ in energy with $\Delta = E_b - E_a$. Since ψ_a and ψ_b are appreciable over many sites ($m > 1000$) the hopping should be viewed as between eigenstates rather than between specific sites. The $c_{a,i}$ and $c_{b,j}$ oscillate in sign and are particularly complex in many valley semiconductors because the $\phi_i(\mathbf{r} - \mathbf{R}_i)$ are $1s - A_1$ (ground state impurity band) which are symmetric linear combinations of the six (Si) and four (Ge) conduction band valleys. These functions oscillate in sign and the dependence of $c_{i,a}$ on $|\mathbf{R}_{ia}|$ features both oscillatory and exponential decay components. Calculation of the phonon matrix element $\langle \psi_b | V_q e^{\mathbf{iq}\cdot\mathbf{r}} | \psi_a \rangle$ for hopping leads to an expression of the form

$$\langle \psi_b | V_q e^{\mathbf{iq}\cdot\mathbf{r}} | \psi_a \rangle = \sum_i \sum_j c_{b,j} \cdot c_{i,a} V_q \exp(\mathbf{iq} \cdot \mathbf{R}_{ij}) F_{ij}(\mathbf{q} \cdot \mathbf{R}_{ij}), \tag{19}$$

where $F_{ij} = \int \phi_j(\mathbf{r} - \mathbf{R}_j) \exp[\mathbf{iq} \cdot (\mathbf{r} - \mathbf{R}_{ij})] \phi_i(\mathbf{r} - \mathbf{R}_i) d\mathbf{r}$. This is a complex summation for a dense array of random donors and is well beyond the scope of the present work. However, it is at least plausible that the important functional dependence is of the form $\sqrt{q} G(\mathbf{R}_{ab}) \exp[-|\mathbf{R}_{ab}|/\xi(n)]$, but where $G[\mathbf{R}_{ab}, a, r_s \xi(n)]$ is a weak function of $\mathbf{R}_{ab}$ [as in the case of Eqs. (5) and (A3c)] rather than a simple polynomial in $|\mathbf{R}_{ab}|/\xi$ found in the dilute limit by MA. The crucial new feature of this paper is the introduction of spatial dispersion in $\epsilon'(n,R)$ which introduces the new characteristic length r_s and removes R in the MA expression for the resonance energy W. This new length r_s removes the cancellation in $W = L - SJ$ and makes $L \gg SJ$, which in turn makes the form of the overlap S much less important. The overlap for the functions ψ_a and ψ_b will be reduced compared with the pair expression in Eq. (A6).

In analyzing whether Mott VRH is meaningful when $T_0 < T$ and $R_m < \xi$, one needs to consider the following. The number of states in a hopping volume in the hopping energy interval Δ_m is $N(E_F)(4\pi R_m^3/3)\Delta_m$. By Mott's ansatz this quantity is unity independent of whether $T_0 < T$ or $T_0 > T$. The number of neutral donors (number of localized electrons n_l) in a hopping volume is $n_l(4\pi R_m^3/3) = N_{d,hv} \gg 1$ at temperatures where most of the electrons are in states below the mobility edge E_c. This inequality is an essential feature of Mott VRH. This is satisfied over a substantial range of T, but breaks down at higher T when thermal excitations excite electrons above the mobility edge. In this case $n_l = n[1 - f_a(n,T)]$ and the density of electrons thermally activated above E_c is $n f_a(n,t)$. The quantity $f_a(n,T)$ has been determined for Si:As samples[40] [$f_a = 0.1$ for the 7.39, 7.57, 7.90, and 8.41 at $T = 37$, 32, 27, and 24 K, respectively]. For $f_a = 0.1$ a significant portion of the conductivity results from the activated component $\sigma_a(n,T)$. The smallest value of $N_{d,hv}$ occurs at the minimum in R_m discussed after Eq. (8). For Si:As this occurs close to $0.94 n_c$ [the 8.07 sample in Table IV in SKC] and leads to $N_{d,hv} \sim 155$ at 1 K and 9 at 25 K, but by 25 K $d \ln \sigma(T)/d \ln T$ already shows large deviations from the Mott result $1/4(T_0/T)^{1/4}(z_m \gg 1)$ [or $2/7(T_0/T)^{2/7}(z_m \ll 1)$] due to large values of $d \ln \sigma_a(T)/d \ln T$. However, there is a substantial range of T where $N_{d,hv} \gg 1$ and in this T regime it is easy to show that $R_m/\xi \ll 1.5[n/N(E_F)kT]$ for $z_m \gg 1$ and $R_m/\xi \ll 0.35[n/N(E_F)kT]^2$ for $z_m \ll 1$. For the above, Si:As samples one can prove that both $R_m/\xi > 1$ for $z_m > 1$ and $R_m/\xi < 1$ for $z_m < 1$. In the T regime (see Fig. 6 in SKC) where Mott VRH determines the logarithmic derivative $d \ln \sigma/d \ln T \sim 1/4(T_0/T)^{1/4}$ [or $2/7(T_0/T)^{2/7}$, which would be hard to distinguish from the Mott result] both of the above inequalities can be satisfied because of the large values of $N_{d,hv}$.

For $1 - n/n_c < 0.02$ the comparison of the theory and the experimental results starts to become problematic. The sample homogeneity might only be of order 2%. The 8.48 Si:As and the 3.69 Si:P samples show VRH exponents somewhat less than the Mott value $m = 2/7$. For $1 - n/n_c < 0.01$ the Ge:Ga results[21] show m dropping from $m \sim 1/2$ through 1/4 toward $m \sim 1/8$. For $T_0 \ll T$ the changes in $\sigma(n,T)$ are small and an accurate determination of the parameters m, T_0, and σ_0 becomes progressively more difficult as $n \to n_c$. Uncertainties in the magnitude of the parameter qR_{eff} (larger than qR_m) as $n \to n_c$, make quantitative estimates of the validity of theory difficult. In particular, there are serious questions about the correct asymptotic behavior of $\psi(r)$ in the regime $\xi \gg R_m$. The difficulties with the theory parallel those of the experimental analysis when $1 - n/n_c$ becomes sufficiently small and are complicated by sample homogeneity.

SUMMARY

In summary, the minimization of the MA impedance provides an extension of the procedure used by Mott that pro-

duces the original Mott VRH results in the limit $\Delta_m > kT$, but also yields a new high-temperature result (with $m = 2/7$) in the regime $\Delta_m < kT$. In the latter regime the T-independent prefactor $\sigma_0(n) \propto \xi(n)^{-3}$ is in reasonable agreement with the Si:P and Si:As data for $1-n/n_c > 0.1$. Doping inhomogeneity leads to a nonzero $\langle \sigma_0(\underline{n}=n_c) \rangle$. An important new feature of this approach arises from the spatial dispersion of $\epsilon'(n,R,T)$ which introduces a new screening length r_s. The large change (18.1 to 1.51) in the coefficient $D(n)$ in the Mott T_0 between the LT and HT Mott VRH regimes, is strongly supported by the Si:As and Si:P data. The rapid decrease in $D(n)$ accounts for the minimum in the mean hopping distance $R_m(n,T=\text{const})$ for $n \sim 0.95 \pm 0.01 n_c$. The spatial dispersion in $\epsilon'(n,R,T)$ permits an explanation of the Ge:Ga ES VRH results for $1-n/n_c > 0.01$ and can also explain the increase in the Coulomb gap width at fixed T observed in the tunneling results of Massey and Lee.

ACKNOWLEDGMENTS

The author gratefully acknowledges the values of T_0, T_0', σ_0, and σ_0' provided by H. v. Löhneysen, M. Watanabe, and S. Yoshizume, and also support by NSF Grant No. DMR-9803969.

APPENDIX: EVALUATION OF THE RESONANCE INTEGRAL W

The integrals involved in MA's quantity $W = L - SJ$ are readily evaluated for the case of spatial dispersion of the density-dependent dielectric response given much earlier by the expression $\epsilon(r,n)^{-1} = \epsilon_h^{-1} \exp[-r/r_s] + \epsilon(n)^{-1}(1-\exp[-r/r_s])$, where $\epsilon(n) = \epsilon_h + 4\pi\chi'(N)$, where $\chi'(n)$ is the diverging dielectric susceptibility.

The integral $L(R) = e^2 \int [\psi_A \psi_B / \epsilon(r_B,n) r_B] d\tau$ for two sites a distance $R = R_A - R_B$ apart, is given by $L = L_i + L_o + L_{ct}$, where L_i results from the inner portions [see Eq. (4)] of $\psi_A \psi_B$, L_o from outer portions and L_{ct} from the cross terms between the inner and outer portions of ψ.

$$L(R) = e^2 \int \psi_A(r_A) \psi_B(r_B) ([\epsilon_h^{-1} - \epsilon(n)^{-1}] \exp - r_B/r_s + \epsilon(n)^{-1}) d\tau / r_B . \tag{A1}$$

The second term in the square brackets yields results similar to that obtained by MA for a spherical envelope function, but with ϵ_h replaced by $\epsilon(n)$. In the critical regime where $\epsilon(n) \gg \epsilon_h$, this term is very small compared to the first term. Using elliptical coordinates with $\mu + \nu = 2r_A/R$, $\mu - \nu = 2r_B/R$, and $d\tau = (R^3/8)(\mu^2 - \nu^2) d\mu d\nu d\phi$ one obtains for the first term in (A1) the results for L_i, L_o, and L_{ct}, namely,

$$L_i = c^2 [2e^2 r_s / \epsilon_h a^2 (1 + a/2r_s)] f(n) e^{-R/a} [1 - [r_s / R(1 + a/2r_s)](1 - e^{-R/rs})], \tag{A2a}$$

$$L_o = [4e^2 a^3 / \epsilon_h r_s \xi^3 (1 + 2r_s/\xi)] f(n) e^{-R/\xi} [1 - [2r_s^2 / R\xi (1 + 2r_s/\xi)](1 - e^{-R/rs})], \tag{A2b}$$

$$L_{ct} = [8ce^2 a^2 / \epsilon_h (r_s \xi)^{3/2}] f(n) \{ e^{-R/\xi} / (1 + a/r_s)^2 - 2a^2 / R\xi (1 + a/r_s)^4 [e^{-R/\xi} - e^{-R(1/a + 1/rs)}] - e^{-R/a} / [1 - (a/r_s{}^2] + 2a/R[1 - (a/r_s)^2]^2 [e^{-R/a} + e^{-R(1/rs + 1/\xi)}] \}, \tag{A2c}$$

where $f(n) = 1 - \epsilon_h/\epsilon(n)$ and very small terms of order $(a/\xi)^2$ and $a^2/r_s\xi$ have been omitted. For comparison the MA-like matrix elements, but with ϵ_h replaced by $\epsilon(n)$, are shown to be

$$L_{i,\text{MA}} = c^2 [e^2/\epsilon(n) a] e^{-R/a} (1 + R/a), \tag{A3a}$$

$$L_{o,\text{MA}} = (a/r_s)^3 [e^2/\epsilon(n)\xi] e^{-R/\xi} (1 + R/\xi), \tag{A3b}$$

$$L_{ct,\text{``MA''}} = [8ce^2/\epsilon(n)(r_s\xi)^{3/2}][(e^{-R/\xi} - e^{-R/a}) \times (1 + 2a/R)]. \tag{A3c}$$

In the critical regime $L_{ct,\text{``MA''}}$ is the largest of these three just as $L_{ct}$ in Eq. (A2) is the largest of the terms containing $f(n)$. These cross terms do not have the polynomial form shown in Eqs. (A3a) and (A3b). The ratio of the leading terms in $L_{ct}$ and $L_{ct,\text{``MA''}}$ (the $e^{-R/\xi}$ terms) is given by

$$L_{ct}/L_{ct,\text{``MA''}} = [\epsilon(n)/\epsilon_h - 1] / [(1 + a/r_s)^2 (1 + 2a/R)]. \tag{A4}$$

Based on the Si:As data[36] for $\epsilon(n, T \to 0)$ and the parameters in Table II, one finds this ratio to be 2.0 at $0.8n_c$, 3.9 at $0.88n_c$, 8.9 at $0.94n_c$ and 87.3 at $0.99n_c$. In the entire range of $1-n/n_c$ shown in Figs. 2 and 3, the MA resonance integral W is determined predominantly by $L_{ct}(R) \propto \epsilon^{-R/\xi}$. *The prefactor dependence of L_{ct} on R is negligible.*

Applying the same approach to evaluating $J(R) = e^2 \int [|\psi_A|^2 / \epsilon(r_B,n) r_B] d\tau$ and ignoring the $\epsilon(n)^{-1}$ term that is the same as the MA result, one obtains for $J_{ct}(R)$

$$J_{ct}(R) = [8ce^2 a^2 / \epsilon_h (r_s\xi)^{3/2}] f(n) \{ -e^{-R(1/a + 1/\xi)} / 1 - (a/r_s)^2 + [2a/R[1 - (a/r_s)^2]^2][e^{-R/rs} - e^{-R(1/a + 1/\xi)}] \}, \tag{A5}$$

TABLE III. Typical wave function parameters and overlap.

$1-n/n_c$	R_m(Å)[a]	ξ (Å)[a]	$(a/\xi)^{3/2}$	c	$S(q=0)$	$S(q)$[a]
0.118	214	129	0.041	0.947	0.0416	<0.024
0.081	146	191	0.0229	0.961	0.059	<0.055
0.06	128	249	0.0154	0.968	0.0622	<0.057
0.012	192	1104	0.00165	0.982	0.0348	<0.01

[a]Based on Si:As values from SKC, $a = 15.45$ Å, $r_s \sim 49$ Å, $T = 1$ K.

where small terms $(a/\xi)^2$ and $a^2/r_s\xi$ are again neglected. J_{ct} contains no term in $e^{-R/\xi}$ and is always smaller than L_{ct} since $\xi > R > r_s > a$. Even for the smallest value of R_m one finds $L_{ct} \sim 16 J_{ct}$ for Si:As. It is also important to note that $J(R)$ differs from the dilute MA result because a term $e^2/\epsilon_h a$ is missing. This in turn means there is no cancellation in the expression $L - SJ$ and that the leading term in W comes only from L since the SJ terms involve either $e^{-3\rho}$ or $\exp[-(\rho + 2R/2r_s)]$. In the present case for $\epsilon(n) \gg \epsilon_h$ the leading term in W is L_{ct}.

The overlap integral for the wave function in Eq. (4) is readily calculated using elliptic coordinates yielding

$$S = c^2[1 + R/a + 1/3(R/a)^2]e^{-R/a} + (a/r_s)^3[1 + R/\xi + 1/3(R/\xi)^2]e^{-R/\xi} + 16c(a/r_s)^{3/2}(a/\xi)^{3/2} \times [(1 - 4a^2/R\xi)e^{-R/\xi} + (1 + 4a^2/R\xi)e^{-R/a}]. \quad \text{(A6)}$$

In Eq. (A6) we have used $\alpha\beta = R^2(1/2a + 1/2\xi)(1/2a - 1/2\xi) = (R/2a)^2(1 - a^2/\xi^2) \approx (R/2a)^2$ since $\xi \gg a$ in the critical regime. In this regime the first term in Eq. (A6) is always negligible. Numerical calculations indicate the second term is always important and the third cross term is comparable for $0.88 n_c$ and $0.94 n_c$, but is negligible for $0.99 n_c$. Values of R_m, ξ, $(a/\xi)^{3/2}$, c, and $S(R_m, q=0)$ are given Table III for four values of n/n_c in the critical regime for Si:As. The largest value of S occurs at the smallest value of R_m at $n/n_c \sim 0.94$. S is small in the critical regime because most of the wave function density is in the core region since c is substantially larger than $(a/r_s)^{3/2}$. The significant result is that $S(R)$ is very small in the critical regime and even if $J(R)$ were comparable to $L(R)$ the leading term for $W = L - JS$ comes from L. The q-dependent overlap integrals $S(q)$ can also be calculated and are given by

$$S_o(q) = (a/r_s)^3 e^{i\mathbf{q}\cdot\mathbf{Ra}} e^{iqR/2} e^{-R/\xi} \left\{ \cos\phi \left\{ \frac{1}{[f(q\xi)]^2} + \frac{R/2\xi}{f(q\xi)} \right\} + \sin\phi \left\{ \frac{[1 - (q\xi/2)^2]}{[f(q\xi)]^2} + \frac{R/\xi}{f(q\xi)} \right\} - (q\xi/2)^{-3}\{\sin\phi(R/\xi - 1) - i(R/\xi)^2 e^{R/\xi} \times [Ei(-R/\xi - i\phi) - Ei(-R/\xi + i\phi)]\right\} \quad \text{(A7a)}$$

$$\frac{S_{ct}(\mathbf{q})}{S_{ct}(0)} \sim \frac{e^{i\mathbf{q}\cdot\mathbf{R}} e^{iqR/2}\{\cos\phi[(1 - y^2)^4 + 2(qa)^2(1 - y^2)(1 - 3y^2) + (qa)^4] + \sin\phi[2qay + 1 + (qa)^2]\}}{g_+(qa,y)\, g_-(qa,y)}, \quad \text{(A7b)}$$

where $\phi = qR/2$, $f(q\xi) = [1 + (q\xi/2)^2]$, $g_\pm(qa,y) = [(1 \pm y)^2 + (qa)^2]^2$ with $y = a/\xi$. $S_i(q)$ is the same as in Eq. (A7a), but with ξ replaced by a and the prefactor $(a/r_s)^3$ replaced by c^2. Only the largest term in $S_{ct}(q)$ has been given (the $e^{-R/\xi}$ term). Terms in $S_{ct}(q) \propto e^{-R/\xi}$ that are smaller by the ratio a/R and $(a/R)^2$ have been neglected. It should be stressed that $S_{ct}(q)$ is much larger than $S_i(q)$ and $S_o(q)$, $S_o(q)/S_o(0) < 1$ since $f(q\xi) > 1$ and the ratio can be small when $q\xi \gg 1$, as is the case for $T = 1$ K and higher. At lower temperatures $[T \leq 0.1\,\text{K}]\, q\xi < 1$ and $f(q\xi) \to 1$ as $q\xi \to 0$, however in this regime R_m/ξ is increasing and the exponential term reduces both $S_{ct}(q)$ and $S_{ct}(0)$.

[1] N. F. Mott, J. Non-Cryst. Solids **1**, 1 (1968).
[2] A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).
[3] K. L. Chopra and S. K. Bahl, Phys. Rev. B **1**, 2545 (1970).
[4] F. R. Allen and C. J. Adkins, Philos. Mag. **26**, 1027 (1972).
[5] H. Hess, K. Deconde, T. F. Rosenbaum, and G. A. Thomas, Phys. Rev. B **25**, 5578 (1982).
[6] S. Yoshizume *et al.*, in *Localization and Metal-Insulator Transitions*, edited by H. Fritzsche and D. Adler (Plenum, New York, 1985), p. 77.
[7] J. Delahaye, J. P. Brison, and C. R. Berger, Phys. Rev. Lett. **81**, 4204 (1998).
[8] B. Ellman, H. M. Jaeger, D. P. Katz, T. F. Rosenbaum, A. S. Cooper, and G. P. Espinosa, Phys. Rev. B **39**, 9012 (1989).
[9] A. Aharony, Y. Zhang, and M. P. Sarachik, Phys. Rev. Lett. **68**, 3900 (1992).
[10] V. Ambegaokar, B. I. Halperin, and J. S. Langer, Phys. Rev. B **4**, 2612 (1971).
[11] M. Pollak, J. Non-Cryst. Solids **8-10**, 486 (1972); *ibid.* **11**, 1 (1972).
[12] A. Miller and E. Abrahams, Phys. Rev. **120**, 745 (1960).
[13] T. G. Castner, Phys. Rev. B **21**, 3523 (1980).
[14] A. L. Efros and B. I. Shklovskii, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Berlin, 1984), Chap. 9.
[15] T. G. Castner, in *Hopping Transport in Solids*, edited by M. Pollak and B. Shklovskii (North Holland, Amsterdam, 1991).
[16] M. A. Paalanen, T. F. Rosenbaum, G. A. Thomas, and R. N. Bhatt, Phys. Rev. Lett. **48**, 1284 (1982); G. A. Thomas, M. A. Paalanen, and T. F. Rosenbaum, Phys. Rev. B **27**, 3897 (1983).
[17] H. Stupp, M. Hornung, M. Läkner, O. Madel, and H. v. Löhneysen, Phys. Rev. Lett. **71**, 2634 (1993).
[18] W. N. Shafarman, D. W. Koon, and T. G. Castner, Phys. Rev. B **40**, 1216 (1989).
[19] M. Hornung and H. v. Löhneysen, in Proceedings of the 21st International Conference on Low Temperature Physics, Prague, 1996 [Czech J. Phys. **46**, 2437 (1996)].
[20] T. G. Castner, Phys. Rev. Lett. **73**, 3600 (1994).
[21] M. Watanabe, Y. Ootuka, K. M. Itoh, and E. E. Haller, Phys. Rev. B **58**, 9851 (1998).
[22] Y. Zhang, P. Dai, M. Levy, and M. P. Sarachik, Phys. Rev. Lett. **64**, 2687 (1990); Y. Zhang and M. P. Sarachik, Phys. Rev. B **51**, 2580 (1995).
[23] H. F. Hess, K. DeConde, T. F. Rosenbaum, and G. A. Thomas, Phys. Rev. B **25**, 5578 (1982).
[24] G. A. Thomas, Y. Ootuka, S. Kobayashi, and W. Sasaki, Phys. Rev. B **24**, 4886 (1981).
[25] J. G. Massey and M. Lee, Phys. Rev. Lett. **77**, 3399 (1996).

[26] G. Deutscher, Y. Levy, and B. Souillard, Europhys. Lett. **4**, 577 (1987).
[27] S. Alexander and R. Orbach, J. Phys. (France) Lett. **43**, 625 (1982).
[28] R. Mansfield, in *Hopping Transport in Solids*, edited by M. Pollak and B. I. Shklovskii (North Holland, Amsterdam, 1991), p. 349.
[29] M. L. Knotek and M. Pollak, J. Non-Cryst. Solids **8-10**, 505 (1972); Phys. Rev. B **9**, 644 (1974); Philos. Mag. B **5**, 1183 (1977); M. Pollak and M. Ortuno, in *Electron-Electron Interactions in Disordered Systems* (North-Holland, Amsterdam, 1985).
[30] H. Haken, Z. Naturforsch. A **9A**, 228 (1954).
[31] S. Katsumoto, in *Anderson Localization*, edited by T. Ando and H. Fukuyama (Springer-Verlag, Berlin, 1988), p. 45.
[32] K. Slevin and T. Ohtsuki, Phys. Rev. Lett. **78**, 4083 (1997); **82**, 669 (1999).
[33] T. G. Castner, Phys. Rev. Lett. **84**, 1539 (2000).
[34] A. Blaschette, A. Ruzzu, S. Wagner, and H. v. Löhneysen, Europhys. Lett. **36**, 527 (1996).
[35] M. Capizzi, G. A. Thomas, F. DeRosa, R. N. Bhatt, and T. M. Rice, Phys. Rev. Lett. **44**, 1019 (1980); M. Paalanen, T. F. Rosenbaum, G. A. Thomas, and R. N. Bhatt, *ibid.* **51**, 1896 (1983).
[36] J. S. Brooks, O. G. Symko, and T. G. Castner, Proceedings of 18th Conference on Low Temperature Physics [Jpn. J. Appl. Phys. **26**, 721 (1980), Suppl. 26-3].
[37] R. A. Faulkner, Phys. Rev. **184**, 713 (1969).
[38] T. G. Castner, N. K. Lee, H. S. Tan, L. Moberly, and O. Symko, J. Low Temp. Phys. **38**, 447 (1980).
[39] A. G. Zabrodskii and K. N. Zinov'eva, Pis'ma Zh. Eksp. Teor. Fiz. **37**, 3699 (1983) [JETP Lett. **37**, 436 (1983)].
[40] T. G. Castner and W. N. Shafarman, Phys. Rev. B **60**, 14 209 (1999).