Density of states for an electron in a correlated Gaussian random potential: Theory of the Urbach tail

Sajeev John
Department of Physics, Princeton University, Princeton, New Jersey 08544
and Corporate Research Science Laboratories, Exxon Research and Engineering Company, Annandale, New Jersey 08801

M. Y. Chou, M. H. Cohen, and C. M. Soukoulis*
Corporate Research Science Laboratories, Exxon Research and Engineering Company, Annandale, New Jersey 08801
(Received 27 July 1987)

A detailed study of the density of states (DOS) $\rho(E)$ in the tail for an electron in a spatially correlated Gaussian random potential $V(x)$ is presented. For disordered solids characterized by short-range order extending a distance L, of the order of the interatomic spacing, we consider autocorrelation functions $B(x)\equiv\langle V(x)V(0)\rangle$ of the form (i) $V_{\rm rms}^2\exp[-(|x|/L)^m]$ for $1\le m<\infty$. For short-range disorder characterized by two correlation lengths L_1 and L_2, we consider the model (ii) $B(x)=V_{\rm rms}^2[\alpha\exp(-x^2/L_1^2)+(1-\alpha)\exp(-x^2/L_2^2)]$. Finally, we consider the case of longer-range correlations (iii) $B(x)=V_{\rm rms}^2[1+(x/L)^2]^{-m_1/2}$, which may be relevant to systems with topological disorder or disordered polar materials in which the randomness may be modeled by frozen-in longitudinal-optical phonons. We find that for reasonable choices of the correlation lengths and rms potential fluctuation that the entire experimentally observable part of the DOS in three dimensions lies in the crossover regime between the shallow energy Halperin-Lax tail $\rho(E)\sim|E|^{3/2}\exp(-\text{const}\times|E|^{1/2})$ and the deep Gaussian tail $\rho(E)\sim|E|^3\exp(-|E|^2/2V_{\rm rms}^2)$ where the energy E is measured relative to the shifted continuum band edge. For systems with rapidly decaying short-range correlations ($m\ge2$) the crossover regime exhibits a linear exponential (Urbach) tail which easily spans five decades in the DOS. The extent of linearity is highly sensitive to the range of the correlation function $B(x)$. The screened-Coulomb impurity model ($m=1$) requires a screening length considerably smaller than the interatomic spacing to give an Urbach tail. These results are obtained numerically by saddle-point (instanton) evaluation of a replica-functional-integral representation of the one-electron propagator. The instanton method provides an asymptotic expansion for the band-tail DOS, which is nearly exact for all energies below the shifted continuum edge. Comparison is made to the Feynman path-integral method and to a simple physical argument which yields to a high degree of accuracy the results of the instanton method. Our results provide a basis for understanding the extent, precision, and universality of Urbach tails in disordered materials.

I. INTRODUCTION

A long-standing fundamental problem in the understanding of optical properties of disordered solids has been a clear physical explanation of the nearly universally observed Urbach optical absorption edge. The empirical rule first proposed by Urbach[1] in 1953 states that the optical absorption coefficient $\alpha(\omega)$ associated with electronic transitions from the valence- to conduction-band tail takes the form

$$\alpha(\omega)\sim\exp[(\hbar\omega-\hbar\omega_0)/W]\ , \tag{1.1}$$

where $\hbar\omega$ is the photon energy and $\hbar\omega_0$ and W are fitting parameters. $\hbar\omega_0$ is comparable to the band gap, and $W=A+B(k_BT)$ has a temperature-dependent part associated with static disorder as well as a part arising from the thermal excitation of phonons. In some recent papers[2,3] it was suggested that the linear (Urbach) exponential behavior evident in the absorption spectrum could be attributed to a corresponding energy dependence in the one-electron density of states $\rho(E)$ in a static Gaussian random potential $V(x)$ provided that a careful treatment of short-range order on the scale of the interatomic spacing was performed, this being the scale of the localized wave function in the Urbach region. For an autocorrelation function $B(x)=\langle V(x)V(0)\rangle$ chosen to be of the Gaussian form $V_{\rm rms}^2e^{-x^2/L^2}$ with L the interatomic spacing, it was shown[3] by means of a simple physical argument that $\rho(E)$ exhibits a linear exponential tail which easily spans five decades. Although we do not attempt to give a microscopic derivation of the Gaussian autocorrelation, we observe that both the extent and accuracy of linear behavior are sensitive functions of the range of correlations. For example in the model $B(x)=V_{\rm rms}^2e^{-|x|/L}$ studied by Halperin and Lax[4] the extent of linearity is considerably reduced. A true Urbach tail in this model would require a correlation length L considerably smaller than the interatomic spacing. In this paper we present an extensive study of the sensitivity of Urbach tails to the nature of short-range correlations.

In particular we consider correlation functions of the form $B(x)=V_{\rm rms}^2\exp[-(|x|/L)^m]$ for $1\le m<\infty$.

It is shown that accurate linear behavior in the DOS is characteristic of random potentials which are strongly correlated for lengths up to the interatomic distance and then rapidly lose correlation on longer scales. The case $m=1$ is relevant to heavily doped semiconductors with charged impurities.[4] It is plausible that in amorphous semiconductors that correlations decay even more rapidly ($m\ge 2$). This is supplemented by a study of correlations having both a short-range and long-range component. As a simple model we consider

$$B(x)=V_{\rm rms}^2[\alpha e^{-x^2/L_1^2}+(1-\alpha)e^{-x^2/L_2^2}] .$$

Finally we consider a class of power law decaying correlations $B(x)=V_{\rm rms}^2(1+x^2/L^2)^{-m_1/2}$ as may be realized in systems with topological disorder or polar semiconductors where it has been suggested that frozen-in longitudinal-optical phonons may produce long-range correlated random electric fields. Our results provide a basis for understanding the universality of Urbach tails in the one-electron DOS and how small deviations from linearity may represent characteristic signatures of particular forms of short-range or long-range order. A complete test of this hypothesis, however, would require an independent probe of microscopic correlations in the solid.

Theoretical efforts to describe band tailing and the associated optical absorption edge have a long and rich history. For heavily-doped semiconductors with screened-Coulomb impurities, Kane,[5] Bonch-Bruevich,[6] and others developed semiclassical treatments for the density of states which focused primarily on the probability distribution of the potential fluctuations. Taking advantage of the long-range nature of the impurity potential, it was shown that for a Gaussian probability distribution, the deep tail forms a Gaussian density of states. This result remains true in an asymptotic limit for any correlation function $B(x)$ having a finite correlation length L. For band tail states of energy $-|E|$ measured relative to the conduction band continuum edge, it has been shown[7] that the density of states in d dimensions takes the form

$$\rho(E)\sim |E|^d\exp(-|E|^2/2V_{\rm rms}^2),\quad \lambda\ll L \tag{1.2}$$

provided that the electron deBroglie wavelength $\lambda\equiv h/(2m^*|E|)^{1/2}$ is considerably smaller than the correlation length L. Here m^* is the conduction-electron effective mass. However, if L is close to the interatomic spacing as in the case of an amorphous semiconductor the asymptotic limit (1.2) is only realized at energies so low that the presence of the valence band cannot be ignored. Typically, the criterion for asymptotic validity of (1.2) requires that the energy be more than one quarter of the way into the band gap and that the localization length be much smaller than the interatomic spacing. Here, the effective-mass approximation to the true band structure of the solid is inadequate. On the other hand, Halperin and Lax[4] (HL) recognized that for shallow-energy-tail states, for which the effective-mass approximation is valid, the underlying physics is changed completely. Here the kinetic energy of localization plays a dominant role in determining the scale of the most probable potential fluctuation. In the asymptotic limit $\lambda\gg L$, the density of states takes the form[8]

$$\rho(E)\sim |E|^{d(5-d)/4}\exp(-\text{const}\times |E|^{2-d/2}),\quad \lambda\gg L . \tag{1.3}$$

Despite the firm mathematical foundation of the asymptotic forms (1.2) and (1.3), neither of these energy dependences can account for the universally observed Urbach tail in three-dimensional systems. This discrepancy may be resolved by a careful consideration of the relevant energy and length scales. The required interpolation between the asymptotic limits (1.2) and (1.3) exhibiting the observed Urbach spectral dependence was first reported by Sritrakool, Sa'Yakanit, and Glyde.[2] The underlying physics of a broad range of linear exponential behavior is the strong influence of the correlation length L on band-tail states in three dimensions and the consequent pinching of the Halperin-Lax behavior (1.3) into an unobservably narrow energy window near the shifted continuum band edge. In fact the breakdown of HL scaling occurs for deBroglie wavelength λ substantially longer than L as the spatial dimension $d\to 4$. This may be seen by the following simple physical argument.[3,7] For length scales long compared to the correlation length L, the probability distribution for the potential may be taken to be of the form

$$P\{V(x)\}\propto\exp\left[-\frac{1}{\gamma^2}S\{V\}\right], \tag{1.4a}$$

where

$$S\{V\}=\int d^dx\, V^2(x) . \tag{1.4b}$$

The averaged density of localized states at a particular energy E is simply the number of potential fluctuations $V(x)$ capable of supporting a bound state at precisely this energy and weighted by its probability of occurrence (1.4a). Consider, for instance, an elementary class of such fluctuations taking the form of square wells of depth $-v$ and linear extent $l>L$. Using the Heisenberg uncertainty principle, the ground-state energy $-|E|$ may be approximated by $-v+\hbar^2/(2m^*l^2)$. In this case

$$S\{V\}\equiv S(l)=l^d\left[|E|+\frac{\hbar^2}{2m^*l^2}\right]^2 . \tag{1.5}$$

For $d<4$, the density of states in the tail is dominated by those potential wells of size

$$l_{\rm min}^2=\left[\frac{4-d}{d}\right]\left[\frac{\hbar^2}{2m^*|E|}\right] \tag{1.6}$$

which minimize $S(l)$. The Halperin-Lax tail is then simply obtained as

$$\rho(E)\sim\exp\left[-\frac{1}{\gamma^2}S(l_{\rm min})\right] .$$

Thus result is valid provided $l_{\rm min}>L$. For $l_{\rm min}<L$ the lo-

cal minimum (1.6) is not accessible and the density of states approaches the value

$$\rho(E)\sim\exp\left[-\frac{1}{\gamma^2}S(L)\right]$$

leading to the Gaussian form (1.2). It is apparent from (1.6) that in high dimensions ($d\to 4$), $l_{\min}$ becomes smaller than L even for shallow band-tail states described by small $|E|$. That is to say, in higher spatial dimensions, the influence of the correlation length L is felt even at energies for which the deBroglie wavelength $\lambda \gg L$ and the HL tail is in fact pinched into a smaller and smaller energy window as $d\to 4$. This in turn leads to a very broad crossover regime between the two limiting forms (1.2) and (1.3) as well as a high degree of sensitivity to the range of the autocorrelation function of both the extent and linearity of the crossover regime.

The discrepancy between the limiting forms (1.2) and (1.3) and the observed Urbach absorption edge has led to extensive studies of alternative mechanisms.[9–11] The actual optical absorption coefficient $\alpha(\omega)$ is determined not only by the convolution of the valence- and conduction-band tail density of states but also by the oscillator strength of the associated transition. Although the universality of Urbach tails in our analysis is apparent from the one-electron DOS, there are many materials for which the transition matrix element plays an important role in determining the shape of the absorption edge. The most notable of these systems have been described by exciton models. In narrow-band molecular crystals the optically excited electron-hole pair interact strongly to form a tightly bound Frenkel exciton band below the conduction band. Sumi and Toyozawa[10] have described the associated absorption edge in terms of the interplay between the narrow free exciton band and excitons momentarily trapped by interaction with acoustic phonons. The exciton-phonon interaction is written as

$$H=\sum_i \Delta_i a_i^\dagger a_i \ , \tag{1.7a}$$

$$\Delta_j=\sum_k V_k Q_k e^{i\mathbf{k}\cdot\mathbf{r}_j} \ , \tag{1.7b}$$

where $a_i^\dagger$ is the creation operator for a Frenkel exciton at lattice site i, Q_k is an acoustic phonon normal coordinate and for a deformation potential interaction $V_k \propto k$. On the time scale of an optical absorption event, the modes Q_k are essentially frozen and satisfy the statistics of a harmonic oscillator in thermal equilibrium. In particular the "site energies" Δ_n satisfy

$$\langle \Delta_n \Delta_m \rangle_T=\sum_k e^{-i\mathbf{k}\cdot(\mathbf{r}_n-\mathbf{r}_m)}\frac{|V_k|^2}{2\omega_k}\coth\left[\frac{\hbar\omega_k}{2k_BT}\right] , \tag{1.8}$$

where $\langle\ \rangle_T$ denotes thermal expectation value and summation is over the crystalline Brillouin zone. For an acoustic phonon dispersion of the form $\omega_k \propto k$, this reduces in the high-temperature limit to $\langle \Delta_n \Delta_m \rangle_T = D^2\delta_{nm}$, where

$$D^2=k_BT\sum_k |V_k|^2/\omega_k^2 \ .$$

The $k=0$ part of the spectral density $A(k,E)=\langle k|(E-H)^{-1}|k\rangle$ describing an exciton with zero center-of-mass momentum exhibits a Lorentzian peak as a function of E at the center of the narrow-exciton band and a tail of lower energy localized exciton levels arising from disorder. The underlying physics of the localized exciton levels is the same as in our model of a single particle in a static Gaussian random potential. However, the existence of a sharp exciton band modifies the nature of the overall optical absorption edge. Abe and Toyozawa[12] and more recently Schreiber and Toyozawa[13] have performed extensive numerical studies on the effects of Frenkel exciton line shape on the Urbach edge. It was found that the density of states is the dominant cause of the exponential tail in the optical absorption spectra in the systems they considered.

In materials with high dielectric constant, the electron-hole interaction is considerably weaker. For example, in GaAs with a dielectric constant $\epsilon\simeq 12$, the radius of the lowest exciton orbit is almost 100 Å. The final-state interaction between the electron and hole again appears as a large enhancement of the optical absorption over the square-root continuum for the one-electron density of states and the effect of the enhancement is felt in the absorption edge. Dow and Redfield[9] have studied the influence of the internal degrees of freedom of such large-radius (Wannier) excitons on the absorption edge. In particular, for ionic crystals and polar semiconductors they argued that the dominant interaction is that of the Wannier exciton with frozen-in longitudinal-optical phonons. In this case we may use (1.7a) and (1.7b) except with $V_k \propto 1/k$ and $\omega_k=$const which in the high-temperature limit leads to long-range correlations of the site energies. From Eq. (1.8) we obtain $\langle \Delta_n \Delta_m \rangle_T \sim 1/|\mathbf{r}_m-\mathbf{r}_n|$. The resulting long-range-correlated random electric fields cause field ionization of the exciton. The final-state interaction determined by the overlap of the electron and hole wave functions decreases exponentially as one goes deeper in the tail. The energy dependence of this oscillator strength plays an important role in determining the overall shape of the optical absorption edge. The importance of matrix-element effects for long-range correlation random potentials is also evident in our calculation. We find that for an electron in a static random potential for which the autocorrelation function $B(x-y)\sim 1/(x-y)$ at large separations the one-electron DOS by itself does *not* exhibit an Urbach tail (see Fig. 5 and later discussion). This is a consequence of the absence of a Halperin-Lax region for autocorrelation functions for which $\int d^dx\, B(x)$ is infinite. For integrable correlation functions, however, the band-tail density of states by itself exhibits a linear exponential energy dependence, and we take this as our starting point from which additional effects such as those described above may be incorporated.

II. FIELD THEORY: GAUSSIAN WAVE-FUNCTION APPROXIMATION

A single electron near the conduction-band edge may be described in the effective mass approximation by a Schrödinger equation

$$\left[-\frac{\hbar^2}{2m^*}\nabla^2+V(x)\right]\psi(x)=-|E|\psi(x)\ . \tag{2.1}$$

Here E is measured relative to the unperturbed band edge. The averaged one-electron Green's function for a Gaussian random potential $V(x)$ characterized by a zero mean value and autocorrelation function $\langle V(x)V(y)\rangle = B(x-y)$ has a replica field theory representation:[7,8]

$$\langle G(x,x';E_+)\rangle=\lim_{n\to 0}\left[\frac{1}{n}\int D\phi\,\phi^\alpha(x)\phi^\alpha(x')e^{-S}\right], \tag{2.2a}$$

where

$$S=\tfrac{1}{2}\int d^dx\,\phi^\alpha(x)\left[-\frac{\hbar^2\nabla^2}{2m^*}-E_+\right]\phi^\alpha(x)$$
$$-\tfrac{1}{8}\int\int d^dx\,d^dy\,\phi^\alpha(x)\phi^\alpha(x)B(x-y)$$
$$\times\phi^\beta(y)\phi^\beta(y)\ . \tag{2.2b}$$

Here (2.2a) is a functional integral over the set of replica fields ϕ^α, $\alpha=1,2,\ldots,n$ in which the contour of integration for each $\phi^\alpha(x)$ runs from $-\exp[(i\pi/4)\infty]$ to $+\exp[(i\pi/4)\infty]$. For sufficiently large negative energies this integral is dominated by nontrivial saddle points of the action S. These instantons determine the most probable shape of the localized wave functions for a given energy $-|E|$ and are solutions of the classical equation

$$\left[-\frac{\hbar^2}{2m^*}\nabla^2+|E|\right]\phi^\alpha(x)=\tfrac{1}{2}\int d^dy\,B(x-y)\phi^\beta(y)$$
$$\times\phi^\beta(y)\phi^\alpha(x)\ . \tag{2.3}$$

Before proceeding with a detailed numerical analysis of this saddle point equation, it is instructive to obtain an approximation to the density of states by means of an approximate analytical minimization of the action (2.2b). This may be done by assuming that the localized wave function for a particular energy takes the form of a Gaussian:

$$\phi^\alpha(x)=A\psi(x)\hat{n}^\alpha\ , \tag{2.4a}$$

where

$$\psi(x)\equiv\frac{1}{(\pi R^2)^{d/4}}\exp(-x^2/2R^2) \tag{2.4b}$$

is a normalized harmonic oscillator ground-state wave function of spatial extent R. Here $\hat{n}^\alpha$ is the αth component of the unit vector in replica space and A is the instanton amplitude. Assuming an autocorrelation function of the form $B(x)=V_{\rm rms}^2\exp(-x^2/L^2)$, the best Gaussian wave function approximation to the density of states follows from inserting (2.4) into (2.2b) and minimizing the resulting action

$$S(A,R)=\frac{A^2}{2}\left[\frac{d\hbar^2}{4m^*R^2}+|E|\right]$$
$$-\frac{A^4}{8}V_{\rm rms}^2\frac{1}{(1+2R^2/L^2)^{d/2}} \tag{2.5}$$

with respect to the two variational parameters A and R. Introducing the energy scale,

$$\varepsilon_L\equiv\frac{\hbar^2}{2m^*L^2} \tag{2.6}$$

the required local minimum occurs at

$$A^2=\frac{2}{V_{\rm rms}^2}\left[\frac{d}{2}\varepsilon_R+|E|\right]\left[1+\frac{2\varepsilon_L}{\varepsilon_R}\right]^{d/2} \tag{2.7}$$

for the instanton amplitude, and the localization length R is determined by

$$\varepsilon_R\equiv\frac{\hbar^2}{2m^*R^2}=\frac{4|E|/(4-d)}{1+\left[1+\dfrac{16|E|}{(4-d)^2\varepsilon_L}\right]^{1/2}}\quad(d<4)\ . \tag{2.8}$$

The exponential part of the density of the states is then approximated by

$$\rho(E)\sim\exp(-S_{\rm min})\ , \tag{2.9a}$$

where

$$S_{\rm min}=\frac{1}{2V_{\rm rms}^2}\left[\frac{d}{2}\varepsilon_R+|E|\right]^2\left[1+\frac{2\varepsilon_L}{\varepsilon_R}\right]^{d/2}\ . \tag{2.9b}$$

It is straightforward to verify that this analytical form exhibits the three regimes of physical interest suggested earlier. For $|E|<(1-d/4)^2\varepsilon_L$, the localization length may be approximated by $\varepsilon_R\simeq 2|E|/(4-d)\ll\varepsilon_L$. It follows that (2.9b) may be approximated by

$$S_{\rm min}\simeq\frac{1}{2V_{\rm rms}^2}\left[\frac{4|E|}{4-d}\right]^2\left[\frac{(4-d)\varepsilon_L}{|E|}\right]^{d/2},$$
$$|E|<\left[1-\frac{d}{4}\right]^2\varepsilon_L \tag{2.10}$$

yielding the energy dependence of the Halperin-Lax density of states. If, on the other hand, $|E|\gg 4\varepsilon_L$, the term $2\varepsilon_L/\varepsilon_R$ appearing in (2.9b) may be neglected in comparison to unity since in this limit $\varepsilon_R\propto(\varepsilon_L|E|)^{1/2}$. It follows that

$$S_{\rm min}\simeq\frac{|E|^2}{2V_{\rm rms}^2},\quad |E|\gg 4\varepsilon_L \tag{2.11}$$

yielding the expected deep Gaussian tail. The crossover regime,

$$\left[1-\frac{d}{4}\right]^2 < \frac{|E|}{\varepsilon_L} < 4 \quad \text{(Urbach)} \tag{2.12}$$

for $d=3$ exhibits an essentially linear exponential behavior characteristic of the Urbach tail throughout most of the observable energy range. The Gaussian wave function approximation further exhibits the pinching of the Halperin-Lax regime into a narrow energy range between $(1-d/4)^2\varepsilon_L$ and a shifted (extrapolated) continuum-band edge. In three dimensions, for weak disorder, the square-root continuum extrapolates to a band edge at $E \simeq -V_{\text{rms}}^2/2\varepsilon_L$ (see Appendix).

Our use of the Gaussian wave function approximation in this section was primarily to make the physics of the replica technique transparent. It is in fact an exact formalism from which a systematic, asymptotically exact result for the band-tail DOS may be obtained as will be discussed in Secs. IV and V. For the particular case of Gaussian autocorrelations, the Gaussian wave function ansatz reproduces to within a 3% error the results of the more exact analysis.

III. RELATIONSHIP OF FIELD THEORY TO THE FEYNMAN PATH INTEGRAL

An alternative representation of the averaged one-electron Green's function in a random potential follows from the Feynman path integral. Recently Sritrakool *et al.* and Sa'Yakanit[2] have used this method to compute band tails in disordered systems. In this section we explicitly demonstrate the equivalence of the Gaussian wave function approximation to the introduction of a harmonic oscillator trial action in the Feynman path integral. We define a time-dependent Green's function as the matrix element of the evolution operator for the electron initially at position x and at a later time t at position x':

$$G(x,x';t) \equiv \langle x' \mid e^{-iHt/\hbar} \mid x \rangle \ . \tag{3.1}$$

The ensemble average of this propagator for the random potential described in Sec. II has a path-integral representation

$$\langle G(x,x';t) \rangle_{\text{ensemble}} = \int_{\substack{X(0)=x,\\ X(t)=x'}} D\mathbf{X}(\tau) \exp(iS/\hbar) \tag{3.2a}$$

in which the effective action

$$S = \frac{m^*}{2} \int_0^t d\tau \dot{\mathbf{X}}^2(\tau) + \frac{i}{2\hbar} \int_0^t d\tau \int_0^t d\tau' B(\mathbf{X}(\tau)-\mathbf{X}(\tau')) \ . \tag{3.2b}$$

As discussed by Sa'Yakanit,[14] an approximation to the Green's function (3.2a) may be obtained by introducing a trial harmonic-oscillator action

$$S_{\text{trial}} = \frac{m^*}{2} \int_0^t d\tau \dot{\mathbf{X}}^2(\tau) - \frac{m^* \nu^2}{4t} \int_0^t dt \int_0^t d\tau' \mid \mathbf{X}(\tau)-\mathbf{X}(\tau') \mid^2 \tag{3.3}$$

which corresponds to an electron coupled harmonically with oscillator frequency ν to an infinite mass. Expanding the true action S about this trial action yields the first cumulant approximation to the one-electron Green's function which for $x=x'=0$ is simply

$$\langle G(0,0;t) \rangle_{\text{ensemble}} = J_{\text{trial}} \exp(i\hbar^{-1} \langle S - S_{\text{trial}} \rangle) \tag{3.4a}$$

where

$$J_{\text{trial}} \equiv \int_{X(0)=X(t)=0} D\mathbf{X}(\tau) e^{i\hbar^{-1} S_{\text{trial}}} = \left[\frac{m^*}{2\pi i\hbar t}\right]^{d/2} \left[\frac{\nu t/2}{\sin(\nu t/2)}\right]^d \tag{3.4b}$$

and the angular brackets denote an average with respect to S_{trial}:

$$\langle A \rangle \equiv \int D\mathbf{X}(\tau) \, A \exp(i\hbar^{-1} S_{\text{trial}})/J_{\text{trial}} \ . \tag{3.4c}$$

As discussed by Feynman,[15] the first cumulant may be evaluated by knowledge of the associated forced harmonic-oscillator amplitude. A straightforward but tedious calculation yields

$$\frac{i}{\hbar} \langle S - S_{\text{trial}} \rangle = I_0(t) + I_{\text{int}}(t) \ , \tag{3.5a}$$

where

$$I_0(t) = \frac{d}{2} \left[\frac{\nu t}{2} \cot\left[\frac{\nu t}{2}\right] - 1 \right] \ , \tag{3.5b}$$

$$I_{\text{int}}(t) = \frac{-V_{\text{rms}}^2}{\hbar^2} \int_0^t d\tau \frac{t-\tau}{[1-(8i\varepsilon_L/\hbar\nu)Q(\tau,t)]^{d/2}} \ , \tag{3.5c}$$

and

$$Q(\tau,t) = \frac{\sin(\nu\tau/2)\sin\{\nu[(\tau-t)/2]\}}{\sin(\nu t/2)} \ . \tag{3.5d}$$

The averaged one-electron density of states is then given by

$$\rho(E) \simeq \frac{1}{2\pi\hbar} \oint dt \, e^{iEt/\hbar} \langle G(0,0;t) \rangle_{\text{ensemble}} \ , \tag{3.6}$$

where the contour of integration is from $-\infty$ to $+\infty$ in the lower half complex t plane avoiding the singularities along the real axis of the harmonic oscillator prefactor J_{trial}. It is convenient to distort this contour deep into the lower complex t plane. For $t=-iT$ and $\nu T \gg 1$ we may replace $Q(\tau,t)$ by $-1/2i$ and the density of states becomes

$$\rho(E) \simeq \frac{1}{2\pi\hbar} \oint dt \left[\frac{m^*}{2\pi i\hbar t}\right]^{d/2} (i\nu t)^d \exp[I(t)] \ , \tag{3.7a}$$

where

$$I(t) = -i \left[\frac{|E|}{\hbar} + \frac{d\nu}{4} \right] t - \frac{V_{\text{rms}}^2}{2\hbar^2} \frac{t^2}{(1+4\varepsilon_L/\hbar\nu)^{d/2}} \ . \tag{3.7b}$$

This integral has a saddle point t_0 along the negative

imaginary-time axis for which

$$I(t_0)=-\frac{(|E|+\frac{1}{4}d\hbar\nu)}{2V_{\rm rms}^2}\left[1+\frac{4\varepsilon_L}{\hbar\nu}\right]^{d/2}. \tag{3.8}$$

If we now identify the harmonic-oscillator frequency ν with the spatial extent R of the ground-state harmonic-oscillator wave function:

$$\frac{\hbar\nu}{2}=\varepsilon_R\equiv\frac{\hbar^2}{2m^*R^2} \tag{3.9}$$

it follows that $I(t_0)$ is identical to $S_{\rm min}$ [Eq. (2.9b)] obtained from the Gaussian wave function approximation to the field theory of the previous section. The variational parameter ν may be chosen according to the Lloyd-Best principle[16] which for the band-tail states corresponds approximately to minimizing the exponential part of the density of states $I(t_0)$. This minimization yields precisely the result (2.8). Keeping up to quadratic fluctuations about the saddle point t_0 yields a density of states

$$\rho(E)\sim\frac{1}{2\pi\hbar}\left[\frac{m^*}{2\pi i\hbar t_0}\right]^{d/2}(i\nu t_0)^d \times\exp[I(t_0)]\left[\frac{2\pi}{|I''(t_0)|}\right]^{1/2} \tag{3.10}$$

identical to that of Sec. II. The Gaussian wave function approximation to the field theory is therefore equivalent to the first cumulant approximation to the Feynman path integral expanded about a harmonic-oscillator trial action.

The harmonic-oscillator approximation yields an excellent qualitative picture of the one-electron density of states exhibiting the three energy regimes of physical interest. The actual numerical coefficients which determine the slope of the exponential part of the density of states as well as set the absolute magnitude of the energy-dependent prefactor are, however, sensitive to the actual shape of the most probable localized wave function. We now proceed to relax the restriction to a Gaussian wave function as well as consider the effect of fluctuations about the most probable state.

IV. FIELD THEORY: INSTANTONS

The form of the most probable wave function for a given energy $-|E|$ in the band tail is determined by solving the saddle point (instanton) equation (2.3) associated with the replica-functional-integral representation (2.2) of the averaged one-electron Green's function. Writing $\phi^\alpha(x)=\phi_{\rm cl}(x)\hat{n}^\alpha$, these instantons are solutions to the Hartree equation

$$(-\nabla^2+|E|)\phi_{\rm cl}(x)-\tfrac{1}{2}\int d^dy\,B(x-y)\phi_{\rm cl}^2(y)\phi_{\rm cl}(x)=0\,. \tag{4.1}$$

Sufficiently deep in the tail the solution $\phi_{\rm cl}$ is identical to the Gaussian wave function determined previously. More generally, Eq. (4.1) must be solved numerically. A detailed discussion of numerical results is presented in Sec. VI. The density of states in the single instanton approximation is given by

$$\rho(E)=A(E)\exp[-S(\phi_{\rm cl})]\,, \tag{4.2a}$$

where $A(E)$ is an energy-dependent prefactor and

$$S(\phi_{\rm cl})=\tfrac{1}{8}\int d^dx\int d^dy\,\phi_{\rm cl}^2(x)B(x-y)\phi_{\rm cl}^2(y)\,. \tag{4.2b}$$

As shown previously by one of us,[7] $A(E)$ may be obtained by evaluating fluctuations about the saddle point. There is a degeneracy in the set of solutions to the saddle point equation (2.3) since the instanton $\phi_{\rm cl}(x)\hat{n}^\alpha$ breaks both the translational symmetry in coordinate space as well as the rotational symmetry in replica space of the action S. The method of collective coordinates may be used to include contributions to the functional integral (2.2a) from such degenerate saddle points obtained by global translational or rotation of the instanton. Evaluating all other fluctuations associated with distortions in the shape of the instanton to quadratic order yields

$$A(E)=\frac{1}{2\pi}\|\phi_{\rm cl}\|J_{\rm trans}|\Delta|\,, \tag{4.3}$$

where $\|\ \|$ denotes the L^2 Hilbert space norm:

$$\|\phi_{\rm cl}\|\equiv\left[\int d^dx\,|\phi_{\rm cl}(x)|^2\right]^{1/2} \tag{4.4}$$

and the Jacobian associated with integration over uniform instanton translations is

$$J_{\rm trans}=\left[\frac{1}{d}\right]^{d/2}\|\nabla\phi_{\rm cl}\|^d\,. \tag{4.5}$$

Fluctuations in the instanton shape are incorporated in the factor Δ which is a ratio of Fredholm determinants:

$$\Delta^2=\det'[M_T(1)/2\pi]/\det'[M_L(1)/2\pi]\,. \tag{4.6}$$

Here, the operator

$$M_T(\lambda)=\delta^d(x-y)\left[-\frac{\hbar^2\nabla^2}{2m^*}+|E|-\frac{\lambda}{2}\int d^dz\,B(x-z)\phi_{\rm cl}^2(z)\right] \tag{4.7}$$

describes fluctuations transverse to the direction $\hat{n}^\alpha$ of broken symmetry in replica space and the operator

$$M_L(\lambda)=M_T(\lambda)-\lambda M_{\rm exch}\,, \tag{4.8a}$$

where

$$M_{\rm exch}\equiv\phi_{\rm cl}(x)B(x-y)\phi_{\rm cl}(y) \tag{4.8b}$$

describes the corresponding longitudinal fluctuations. The prime on both determinants denotes the omission of the zero eigenvalues of these operators since these fluctuations have been calculated exactly by the method of collective coordinates. M_T has one zero eigenvalue with eigenfunction $\phi_{\rm cl}(x)$ corresponding to rotations of the instanton in replica space and M_L has d zero eigenvalues with eigenfunctions $(\partial/\partial x_i)\phi_{\rm cl}(x)$, $i=1,2,\ldots,d$ corre-

sponding to translations of the instanton in each of the d independent Cartesian directions.

The evaluation of the ratio of determinants may be simplified[17] by eliminating the continuous spectra of both M_T and M_L. Multiplying both operators by $(-\hbar^2\nabla^2/2m^*+|E|)^{-1}$ leaves Δ invariant and reduces the calculation to that of two discrete spectra. In particular,

$$\Delta^2=\lim_{\lambda\to 1}\left[\frac{\det\tilde{M}_T(\lambda)}{\det\tilde{M}_L(\lambda)}\frac{[e_0^L(\lambda)]^d}{e_0^T(\lambda)}\right], \tag{4.9}$$

where

$$\tilde{M}_T(\lambda)\equiv 1+\lambda(-\hbar^2\nabla^2/2m^*+|E|)^{-1}v_{\rm cl}(x) \tag{4.10a}$$

and

$$\tilde{M}_L(\lambda)\equiv\tilde{M}_T(\lambda)-\lambda(-\hbar^2\nabla^2/2m^*+|E|)^{-1}M_{\rm exch}\ . \tag{4.10b}$$

Here, we have introduced the classical potential

$$v_{\rm cl}(x)\equiv-\tfrac{1}{2}\int d^dz\,B(x-z)\phi_{\rm cl}^2(z)\ . \tag{4.11}$$

Also, $e_0^L(\lambda)$ and $e_0^T(\lambda)$ are the eigenvalues of $M_L(\lambda)/2\pi$ and $M_T(\lambda)/2\pi$ which vanish as $\lambda\to 1$. It follows that

$$\det\tilde{M}_T(\lambda)=\prod_i\left[1-\frac{\lambda}{\zeta_i^T}\right], \tag{4.12a}$$

where ζ_i^T are eigenvalues defined by the equation

$$\left[-\frac{\hbar^2\nabla^2}{2m^*}+\zeta_i^T v_{\rm cl}(x)\right]\psi_i^T=-|E|\,\psi_i^T\ . \tag{4.12b}$$

For instance $\zeta_0^T=1$ and the corresponding eigenfunction $\psi_0^T=\phi_{\rm cl}$. The computation of the determinant (4.12a) may be simplified by knowledge of the asymptotic behavior of ζ_i^T for very large i. The density of bound states in the vicinity of $-|E|$ for a very deep potential (large ζ_i^T) may be estimated using the Thomas-Fermi approximation. If, for instance, we consider filling the potential well $\zeta v_{\rm cl}(x)$ with electrons (neglecting spin degeneracy) up to a Fermi level $-|E|$, the local electron number density $n(x)$ at position x has an asymptotic behavior

$$n(x)\underset{\zeta\to\infty}{\sim}\frac{[\zeta v_{\rm cl}(x)]^{3/2}}{6\pi^2}\quad(d=3)\ . \tag{4.13}$$

The required density of eigenvalues $\rho_T(\zeta)$ is simply the derivative of total number of states with respect to ζ:

$$\rho_T(\zeta)=\frac{d}{d\zeta}\int d^3x\,n(x)\ . \tag{4.14}$$

The Fredholm determinant (4.12a) may then be written as

$$\det\tilde{M}_T(\lambda)=\prod_{\zeta_i^T<\bar{\zeta}}\left[1-\frac{\lambda}{\zeta_i^T}\right]\exp\int_{\bar{\zeta}}^{\infty}d\zeta\,\rho_T(\zeta)\times\ln\left[1-\frac{\lambda}{\zeta}\right]. \tag{4.15}$$

A similar analysis holds for the longitudinal operator:

$$\det\tilde{M}_L(\lambda)=\prod_{\zeta_i^L<\bar{\zeta}}\left[1-\frac{\lambda}{\zeta_i^L}\right]\exp\int_{\bar{\zeta}}^{\infty}d\zeta\,\rho_L(\zeta)\times\ln\left[1-\frac{\lambda}{\zeta}\right], \tag{4.16}$$

where the eigenvalues ζ_i^L are obtained from

$$\{-\hbar^2\nabla^2/2m^*+\zeta_i^L[v_{\rm cl}(x)-M_{\rm exch}]\}\psi_i^L=-|E|\,\psi_i^L\ . \tag{4.17}$$

Unlike the delta function correlated disorder considered by Brezin and Parisi,[17] the asymptotic behavior of the density of eigenvalues $\rho_T(\zeta)$ and $\rho_L(\zeta)$ are identical as a result of the nonlocality of the operator $M_{\rm exch}$: The effect of $M_{\rm exch}$ on eigenfunctions ψ_i^L whose scale of oscillation is very short compared to the correlation length of the disorder and the spatial extent of $\phi_{\rm cl}$ is negligible compared to the direct term. Although the individual determinants (4.15) and (4.16) are divergent for $d\geq 2$, the ratio of these two determinants is well defined. In the white-noise disorder model of Brezin and Parisi[17] this ratio was divergent for $d\geq 2$, and it was necessary to perform a one-loop renormalization of these determinants. The existence of a finite correlation length to the disorder removes this divergence and the density of states in the zero-loop approximation is finite.

V. BEYOND THE SINGLE-INSTANTON APPROXIMATION

The single-instanton approximation to the density of states which we have described is the leading term in an asymptotic expansion which becomes exact in the deep tail and which is extremely accurate throughout the entire band tail. The important corrections to this approximation occur only in the very shallow energy $|E|\to 0$ limit of the Halperin-Lax region for weak disorder. In this limit, the scale of the localized wave function becomes large compared to the correlation length L of the disorder, and it is convenient to approximate the autocorrelation function as $B(x)\simeq\gamma_0\delta^d(x)$. The action S appearing in (2.2b) may be simplified by rescaling all lengths with respect to the deBroglie wavelength $\lambda\equiv\hbar/(2m^*|E|)^{1/2}$ and rewriting the field variables $\phi^\alpha(x)=A\psi^\alpha(x/\lambda)$. If the amplitude A is chosen to be $(|E|/\gamma_0)^{1/2}$ then the action S may be rewritten as $S=1/g\tilde{S}$ where the dimensionless coupling constant

$$1/g\equiv(|E|^2/\gamma_0)(2m^*|E|/\hbar^2)^{-d/2}$$

and $\tilde{S}$ is expressed completely in terms of dimensionless variables. $1/g$ plays the role of a large parameter which organizes the saddle point expansion. The single-instanton approximation to the DOS described in the previous section takes the form

$$\rho_0(E)=A(E)\exp\left[-\frac{1}{g}\tilde{S}(\psi_{\rm cl})\right]. \tag{5.1}$$

Here, the prefactor $A(E)$ arises from considering quadratic fluctuations about the single-instanton solution. The higher-order fluctuations may be written as

$$\Delta\tilde{S}=\tfrac{1}{2}\int d^dy\,\psi_{\rm cl}(\hat{n}^\alpha\hat{\psi}^\alpha)(\hat{\psi}^\beta\hat{\psi}^\beta)+\tfrac{1}{8}\int d^dy\,(\hat{\psi}^\alpha\hat{\psi}^\alpha)^2\ , \tag{5.2}$$

where we have let $\psi^\alpha=\psi_{\rm cl}\hat{n}^\alpha+\hat{\psi}^\alpha$. The exact density of states is given by

$$\rho(E)=\rho_0(E)\langle e^{-(1/g)\Delta\tilde{S}}\rangle_M\ , \tag{5.3a}$$

where the angular brackets denote an average with respect to the quadratic fluctuation matrix $M^{\alpha\beta}$

$$\langle\cdots\rangle_M\equiv\frac{\int D\hat{\psi}(\cdots)\exp\left[-\frac{1}{2g}\int d^dy\,\hat{\psi}^\alpha M^{\alpha\beta}\hat{\psi}^\beta\right]}{\int D\hat{\psi}\exp\left[-\frac{1}{2g}\int d^dy\,\hat{\psi}^\alpha M^{\alpha\beta}\hat{\psi}^\beta\right]}\ . \tag{5.3b}$$

The longitudinal and transverse components of the matrix $M^{\alpha\beta}$ were described in the previous section. An expansion in the small parameter g follows by evaluating (5.3a) in a cumulant expansion. If we define

$$f(g)\equiv\ln\langle\exp(-\Delta\tilde{S}/g)\rangle_M$$

then it follows that

$$f(g)\sim a_1g+a_2g^2+\cdots\ .$$

This constitutes an asymptotic expansion for the density of states. For small g, higher-order terms in $f(g)$ may be neglected. For weak disorder this occurs already in the Halperin-Lax region, and the single-instanton solution is extremely accurate. Although the coefficients $a_1,a_2,\ldots$ are cumbersome to evaluate, a straightforward combinatorial analysis reveals that $\lim_{m\to\infty}|a_{m+1}/a_m|=\infty$. The expansion for $f(g)$ therefore has a zero radius of convergence and is only asymptotic in nature. Physically this occurs because for shallow energy states the probability of occurrence of a potential well is relatively large and a localized state in any given potential fluctuation will interact with other nearby potential fluctuations capable of supporting a bound state at approximately the same energy. Mathematically this appears as an essential singularity in the function $f(g)$ which itself has contributions of order $e^{-1/g}$ as $g\to0$. This corresponds to the increasing importance of multiple instanton configurations in the shallow part of the band tail ($|E|\to0$). In fact the transition region between the band tail and the positive energy continuum is dominated by a strongly interacting gas of instantons. A treatment of the statistical mechanics of this interacting instanton gas is beyond the scope of this paper. However it may be important for a complete description for the DOS throughout the entire energy spectrum and in particular for the behavior near the electronic mobility edge.

VI. NUMERICAL RESULTS

In this section, we present numerical results of $S(\phi_{\rm cl})$ in Eq. (4.2b) by solving the Hartree equation (4.1) for various autocorrelation functions in $d=3$. The autocorrelation functions are assumed to be isotropic in space and of a general form

$$B(z)=B_0f\left[\frac{|z|}{L}\right]\ , \tag{6.1}$$

where L is the characteristic correlation length and f is a dimensionless function with $f(0)=1$. For $d=3$ we obtain a radial Hartree equation

$$\left[-\frac{\hbar^2}{2m^*}\frac{d^2}{dr^2}u(r)+|E|\right]u(r)-\frac{B_0}{2}\int dr'\int dr'f\left[\frac{|\mathbf{r}-\mathbf{r}'|}{L}\right]u^2(r')u(r)=0 \tag{6.2}$$

with $u(r)\equiv r\phi_{\rm cl}(r)$. It is convenient to express Eq. (6.2) in terms of the natural length unit L and energy unit $\varepsilon_L\equiv\hbar^2/2m^*L^2$. The resulting dimensionless equation becomes

$$-\frac{d^2}{dt^2}v(t)+\frac{|E|}{\varepsilon_L}v(t)-\frac{B_0}{2\varepsilon_L}4\pi\int dt'\,F(t,t')v^2(t')v(t)=0\ , \tag{6.3}$$

where $t\equiv r/L$, $v(t)\equiv L^{1/2}u(r)$, and

$$F(t,t')\equiv\frac{1}{4\pi}\int d\Omega_{t'}f(|\mathbf{t}-\mathbf{t}'|)\ . \tag{6.4}$$

The function $F(t,t')$ is listed in Table I for the various autocorrelation functions we studied. For any given $|E|/\varepsilon_L$, Eq. (6.3) is solved self-consistently for $v(t)$. The operational procedure is to write $v(t)=Aw(t)$, where $w(t)$ is a properly normalized function and A is the instanton amplitude. Equation (6.3) may then be written as

$$-\frac{d^2}{dt^2}w(t)+\frac{|E|}{\varepsilon_L}w(t)-g_L\int dt'\,F(t,t')w^2(t')w(t)=0\ , \tag{6.5}$$

where the parameter

$$g_L\equiv\frac{4\pi B_0A^2}{2\varepsilon_L}$$

is to be determined for every $|E|/\varepsilon_L$, and the normalized function $w(t)$ is subject to the boundary conditions $w(0)=0$ and $w(\infty)=0$. An initial guess, $w_{\rm in}(t)$, is first used to determine the "effective" potential $\int dt'F(t,t')w_{\rm in}(t')$ in Eq. (6.5). With this effective potential, we then obtain both g_L and an output $w_{\rm out}(t)$ satisfying Eq. (6.5). The output function $w_{\rm out}(t)$ is again used to construct the effective potential and the whole procedure

TABLE I. The angular average and the ratio between ε_L and E_0 (see text) for general autocorrelations functions f: $F(t,t')\equiv(1/4\pi)\int d\Omega_{t'} f(|\mathbf{t}-\mathbf{t}'|)$ and $\alpha\equiv\varepsilon_L/E_0$.

$f(\lvert z\rvert)$	$f(t,t')$	α
$\exp(-\lvert z\rvert^m)$		
$m=1$	$\frac{1}{tt'}e^{-t_>}[(t_>+1)\sinh t_< - t_<\cosh t_<]$	4
$m=2$	$e^{-(t^2+t'^2)}\frac{\sinh(2tt')}{2tt'}$	0.5
$m=4$	$\frac{1}{4tt'}\int_{(t-t')^2}^{(t+t')^2} e^{-x^2}dx$	0.2466
$(1+\lvert z\rvert^2)^{-n}$	$\frac{1}{(n-1)4tt'}\{[1+(t-t')^2]^{1-n}-[1-(t+t')^2]^{1-n}\}$ $(n\geq 3)\quad(n\neq 1)$	$(2n-5)^{-1}$

is repeated until the difference between the input and output $w(t)$ becomes sufficiently small.

With both g_L and $w(t)$ determined for a given $|E|/\varepsilon_L$ in Eq. (6.5), we are ready to evaluate $S(\phi_{cl})$, the exponent of the density of states. By using (6.1)–(6.5), Eq. (4.2b) can be rewritten as

$$S=\frac{\varepsilon_L^2}{B_0}\frac{1}{2}g_L^2\int dt\int dt' F(t,t')w^2(t)w^2(t') . \quad (6.6)$$

The energy dependence of S comes implicitly from the energy-dependent functions g_L and $w(t)$.

The first set of the autocorrelation functions we have examined takes the form

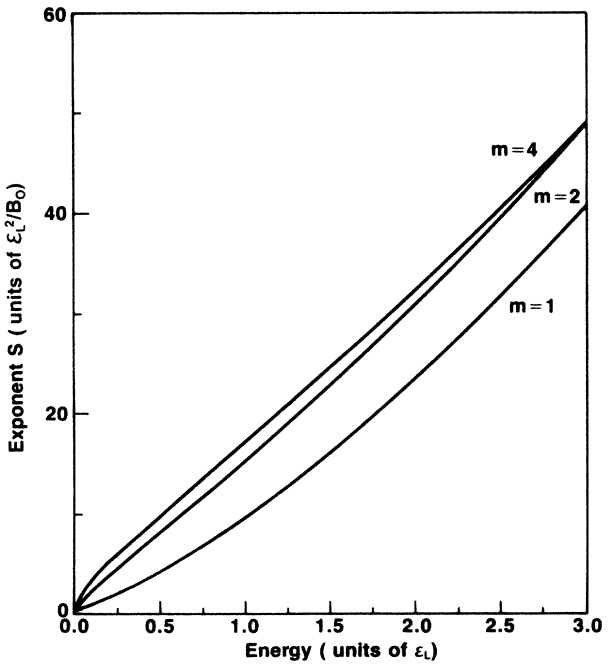

FIG. 1. The calculated S (exponent of the density of states) vs energy for the autocorrelation functions of the form $B(x)=B_0\exp[-(|X|/L)^m]$. The energy unit is $\varepsilon_L\equiv\hbar^2/2m^*L^2$ and S is in units of ε_L^2/B_0. Curves shown are for $m=1$, 2, and 4.

$$B(x)=B_0\exp\left[-\left[\frac{|x|}{L}\right]^m\right] . \quad (6.7)$$

In Fig. 1, the calculated S, in units of ε_L^2/B_0, is plotted as a function of the dimensionless energy $|E|/\varepsilon_L$ for $m=1$, 2, and 4. The case of $m=2$ corresponds to the single Gaussian autocorrelation function which has been used in previous studies,[3] while $m=1$ originates from the screened Coulomb potential with L being the screening length.[4] All of the three curves do exhibit $|E|^{1/2}$ behavior for sufficiently small $|E|$ and $|E|^2$ behavior for sufficiently large $|E|$ in their respective asymptotic limits. On the linear energy scale used in Fig. 1, however, the $|E|^{1/2}$ behavior is inconspicuous (almost invisible in the $m=1$ case) while the asymptotic $|E|^2$ behavior is approached only very slowly as $|E|$ increases. For the intermediate energy region S is well fitted by a straight line over a large energy range. The slope and range of the straight line together with the root-mean-square deviation of the fit are listed in Table II for $m=1$, 2, and 4. Here the linear fit is performed with an energy grid of 0.1 (in units of ε_L) for the $m=2$ and 4 curves in Fig. 1 and 0.02 for the $m=1$ curve. We have chosen the largest linear region possible with a root-mean-square deviation of the fit from linearity being approximately 0.2 (in units of ε_L^2/B_0) in all cases. Some estimates of the parameters ε_L and B_0 and comparisons with experiment will be discussed in the next section.

To analyze the data further, we have plotted in Fig. 2 the logarithmic derivative $n(E)\equiv d(\log S)/d(\log|E|)$ for the three curves of Fig. 1. It starts from 0.5 at $|E|=0$, passes 1.0 as $|E|$ increases, and eventually approaches 2.0 at large $|E|$. However since the experimentally observable energy range relevant to band tailing is located in the neighborhood of $n=1$, the exponent of the density of states appears as a straight line.

Comparison of the three curves in Fig. 1, also reveals the influence of the range of the autocorrelation function on the linear behavior. We can define a "range" L_0 of the autocorrelation function by

$$L_0^2\equiv\frac{1}{d}\frac{\int x^2B(x)d^dx}{\int B(x)d^dx} . \quad (6.8)$$

TABLE II. A list of the slopes, root-mean-square deviations, and range (E_{min} and E_{max}) of the fitted straight lines for curves in Figs. 1 and 4. The logarithmic derivatives $n(E)\equiv d[\ln S(E)]/d(\ln|E|)$ at E_{min} and E_{max} are also listed.

$f(\|z\|)$	Slope	rms deviation	E_{min}	E_{max}	$n(E_{min})$	$n(E_{max})$
$\exp(-\|z\|^m)$						
$m=1$	9.6	0.18	0.02	1.02	0.69	1.21
$m=2$	14.6	0.21	0.1	1.9	0.65	1.07
$m=4$	15.1	0.21	0.1	2.6	0.60	1.03
$(1+\|z\|^2)^{-n}$						
$n=3$	31.2	0.20	0.1	2.0	0.66	1.00
$n=4$	44.6	0.21	0.3	2.6	0.68	0.96
$n=5$	58.2	0.22	0.6	3.2	0.71	0.94

As m decreases, the autocorrelation becomes relatively long ranged. In the reduced energy scale $|E|/\varepsilon_L$, the effect of decreasing m is to push the $|E|^{1/2}$ behavior at low $|E|$ into a smaller and smaller energy region and make the $|E|^2$ behavior appear sooner. Consequently, the extent of the approximate linear behavior becomes the smallest for $m=1$ and also shifts toward smaller $|E|/\varepsilon_L$. In the opposite limit $m\to\infty$, the autocorrelation function becomes a step function. Here the extent of linear behavior appears greatest.

The nature of the deviation from exact linearity may also be adjusted by introducing two correlation lengths to the disorder. As an example, we studied the double Gaussian autocorrelation function

$$B(x)=\frac{B_0}{1+\beta^4}(e^{-x^2/\beta L^2}+\beta^4 e^{-\beta x^2/L^2}), \tag{6.9}$$

where β is the ratio of these two correlation lengths. The coefficients of the combinations are chosen in such a way that the resulting autocorrelation function in Eq. (6.9) has the same range [defined in Eq. (6.8)] as the autocorrelation function of one Gaussian with a correlation length L. In Fig. 3, the calculated S for $\beta=1$, 2, 4, 9, and 16 is plotted as a function of $|E|/\varepsilon_L$. As β increases, the curve changes from concave to convex and the linear behavior region moves toward higher energy.

The third class of autocorrelation functions we have studied has the form $B(x)=B_0(1+x^2/L^2)^{-n}$, describing polar materials and systems with topological disorder. In Fig. 4 we plot the corresponding exponent S versus the dimensionless energy $|E|/\varepsilon_L$ for $n=3$, 4, and 5. The results are similar to those of Fig. 1. Both of the high- and low-$|E|$ limiting cases exhibit the $|E|^2$ and $|E|^{1/2}$ asymptotic behavior, and an apparent straight line dominates in the crossover energy region. The linear behavior extends over a larger energy range going beyond $|E|/\varepsilon_L=3$ in the dimensionless scale.

The case $n=\frac{1}{2}$ actually corresponds to the model of band tailing considered by Dow and Redfield[9] (see Sec. II). Our calculated S for this case is plotted in Fig. 5. Compared with the results in Figs. 1, 3, and 4, the present curve does not exhibit the $|E|^{1/2}$ behavior at small $|E|$. Consequently, there is no crossover energy

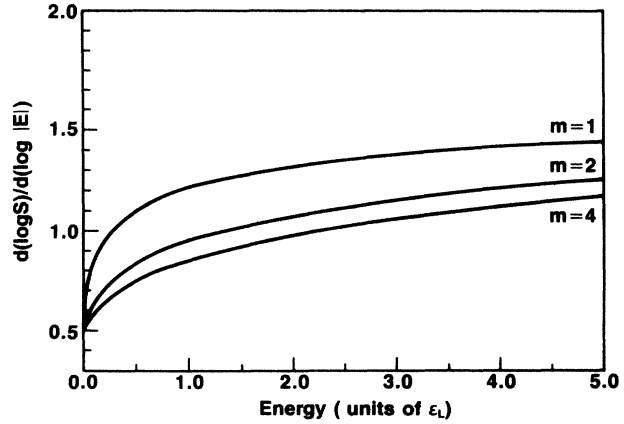

FIG. 2. The logarithmic derivatives $n(E)=d(\log S)/d(\log|E|)$ for the three curves in Fig. 1.

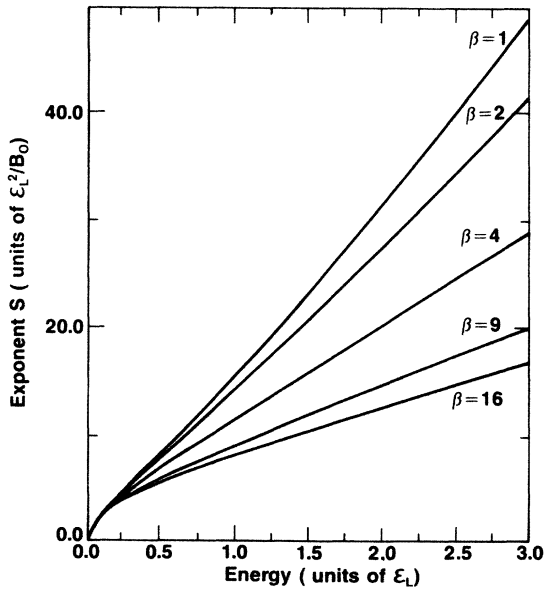

FIG. 3. The calculated S (exponent of the density of states) vs energy for the autocorrelation functions of the form $B(x)=[B_0/(1+\beta^4)](e^{-x^2/\beta L^2}+\beta^4 e^{-\beta x^2/L^2})$. The energy unit is $\varepsilon_L\equiv\hbar^2/2m^*L^2$. Curves shown are for $\beta=1$, 2, 4, 9, and 16.

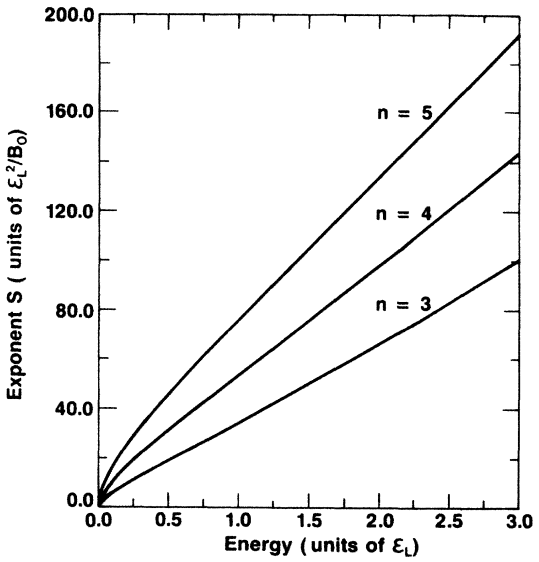

FIG. 4. Same as Fig. 1, except that the autocorrelation functions have the form of $B(x)=B_0(1+x^2/L^2)^{-n}$. Curves shown are for $n=3$, 4, and 5.

region with a linear behavior as found in previous cases. The values of $d(\log S)/d(\log|E|)$ range from 1.5 to 1.6 for the energies plotted in Fig. 5. The lack of the Halperin-Lax tail is attributed to the fact that the integral $\int B(x)d^dx$ does not exist. A similar result holds for any power-law correlation function which falls off more slowly than $1/r^3$ for large r in three dimensions.

Although the results of Figs. 1–3 are obtained from different autocorrelation functions, the overall behavior exhibits certain similarities. To understand the universality of these results, we express Eq. (4.2b) in terms of new variables $\mathbf{R}\equiv\frac{1}{2}(\mathbf{x}+\mathbf{y})$ and $\mathbf{u}\equiv\mathbf{x}-\mathbf{y}$:

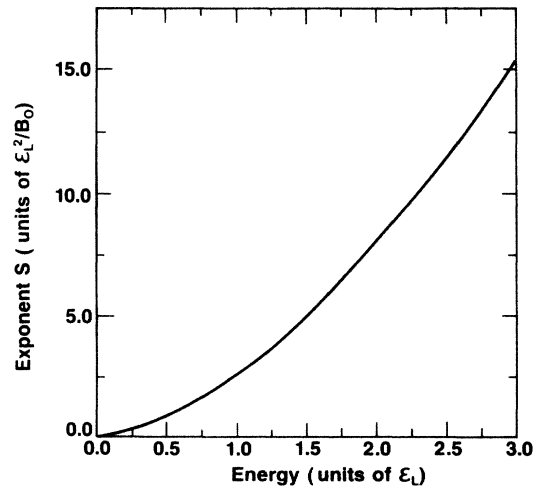

FIG. 5. The calculated S (exponent of the density of states) vs energy for the autocorrelation function of the form $B(x)=B_0(1+x^2/L^2)^{-1/2}$. The energy unit is $\varepsilon_L\equiv\hbar^2/2m^*L^2$.

$$S=\tfrac{1}{8}\int d^dR\int d^du\,\phi_{\rm cl}^2\left[\mathbf{R}+\frac{\mathbf{u}}{2}\right]B(u)\phi_{\rm cl}^2\left[\mathbf{R}-\frac{\mathbf{u}}{2}\right]. \tag{6.10}$$

Next we expand both $\phi_{\rm cl}^2(\mathbf{R}+\mathbf{u}/2)$ and $\phi_{\rm cl}^2(\mathbf{R}-\mathbf{u}/2)$ around $\mathbf{R}$ and arrange the terms in ascending order of u. With the assumption that $B(\mathbf{u})$ is isotropic, the first two terms in (6.10) become

$$S=\tfrac{1}{8}\left[\gamma_0\int\phi_{\rm cl}^4(R)d^dR+\tfrac{1}{2}\gamma_2\int|\nabla_R\phi_{\rm cl}^2(R)|^2d^dr+\cdots\right], \tag{6.11}$$

where

$$\gamma_0\equiv\int B(u)d^du \quad\text{and}\quad \gamma_2\equiv\frac{1}{d}\int u^2B(u)d^du\ .$$

In the low-$|E|$ limit, where the characteristic wave functions do not change appreciably over the correlation length of the potential, the autocorrelation function is usually approximated by a delta function and one has only the first term left in Eq. (6.11). In this case $\phi_{\rm cl}$ scales as $(|E|/\gamma_0)^{1/2}$ and consequently S scales as $(1/\gamma_0)|E|^{2-d/2}$. As $|E|$ increases the potential correlation length becomes non-negligible compared with the electron wavelength. In this case, the actual form of the autocorrelation function has to be considered and thus higher terms in Eq. (6.11) need to be included.

The next step is to choose the scaling factor for both the energy and the exponent S. It is found that the "range" defined in Eq. (6.8) is related to the ratio of γ_2 and γ_0:

$$L_0^2=\frac{\gamma_2}{\gamma_0}\ . \tag{6.12}$$

Therefore it is natural to define an energy scale E_0 by

$$E_0\equiv\frac{\hbar^2}{2m^*L_0^2}\ . \tag{6.13}$$

By evaluating Eqs. (6.12) and (6.13), we can find the ratio between ε_L and E_0 for the different autocorrelation functions discussed above. These values of $\alpha=\varepsilon_L/E_0$ are listed in Table I. With the energy scale E_0, the scale for S is then chosen as $S_0\equiv1/\gamma_0E_0^{2-d/2}$ in order to be consistent with results of the low-$|E|$ limit.

The normalized S versus $|E|$ plots are shown in Fig. 6 for all the autocorrelation functions considered above. In the low-$|E|$ limit, all the curves coincide as required in our scaling. When $|E|$ increases slightly ($\lesssim0.2E_0$) the first two terms in Eq. (6.11) are still sufficient to represent S and all the curves remain almost identical. As $|E|$ gets larger, higher-order terms in Eq. (6.11) become more significant and the curves depart from each other. However, all of them show a linear behavior in the energy range plotted. The slopes of these linear regimes differ by no more than a factor of 2 in these reduced scales. Since we have kept the ranges of all the autocorrelation functions to be the same, the differences shown in Fig. 6 arise from more detailed shape

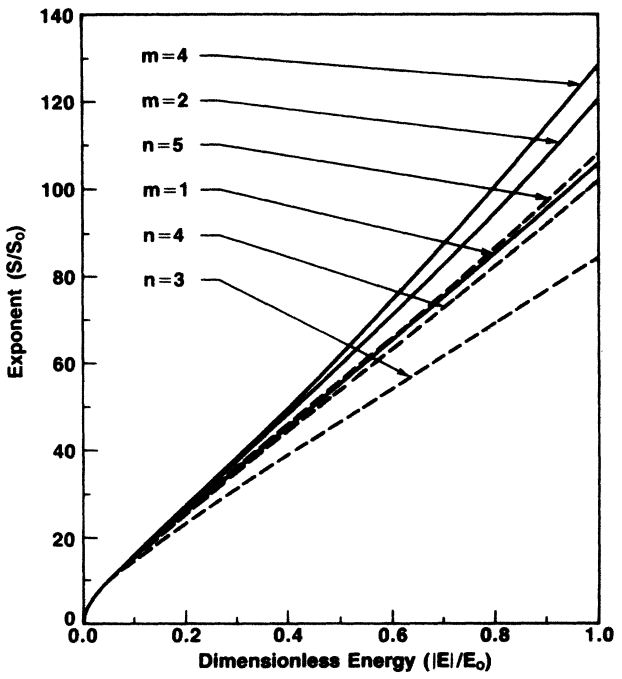

FIG. 6. Results in Figs. 1 and 4 are replotted in the new dimensionless scales E_0 and S_0 to examine the universality. E_0 and S_0 are defined as $\hbar^2/2m^*L_0^2$ and $(1/\gamma_0)E_0^{2-d/2}$, respectively, where L_0 is the range of the autocorrelation function (see text) and $\gamma_0 \equiv \int d^3x\, B(x)$.

differences. The universality of linear exponential behavior is manifest by this rescaling of energies whereas the relation between the actual experimentally measurable energy range and the scaling energies ε_L and E_0 depends on the material to be modeled.

Universality of the deep tail can also be shown for correlations functions $B(x)$ with a Taylor expansion about $x=0$ and following similar procedures as discussed above. Physically this is because the electron wavelength is much smaller than the correlation length of the potential as $|E| \to \infty$. Therefore, the autocorrelation function changes only slightly within the range of the wave function. In this case, the Hartree equation is reduced to a Schrödinger equation of a simple harmonic oscillator[7] and can be solved analytically. It has been shown that S goes to $|E|^2/2B_0$ as $|E|$ gets sufficiently large.[7] In our numerical study, $|E|/\varepsilon_L$ has to be as large as 10^4–10^5 in most cases to reach this limit to within a few percent. This lies far beyond the energy region measured experimentally. In fact the entire band gap of real materials is many orders of magnitude smaller, with the exception of certain molecular crystals or glasses. Real band-structure effects will invariably modify the nature of the deep tail. However, we expect that the results of the continuum effective-mass model which we have used are quite accurate for band-tail states one quarter of the way into the energy gap, since in this regime the localized wave function spans at least a couple of lattice sites.

VII. DISCUSSION

The universality of linear exponential band tails is evident from experimental studies of a variety of disordered systems.[18–25] In the transient-photocurrent measurements[22,23] the accurate power-law relationship between the photocurrent and time delay requires an exponential density of tail states as the results are interpreted by a multiple-trapping model.[26,27] This linear exponential behavior has been observed for as many as five decades in the DOS.[22] As described in the Introduction additional indirect evidence comes from the optical absorption coefficient in the tail which varies exponentially with the photon energy (known as the Urbach rule). Urbach's rule has been observed in a large variety of glasses and amorphous semiconductors subsequent of the original studies in silver and alkali halides. In many of these latter systems oscillator strength effects and exciton mechanisms do not provide strong enough variations of the absorption coefficient with photon energy. The linear exponential behavior must therefore be a measure of the convoluted density of states between the valence and conduction band tail.

In order to compare our calculated results with experiment, we need to have an estimate of the parameters B_0 and ε_L in the model. The correlation length L is usually of the order of the interatomic distance. With L in the range of 2–5 Å and the effective mass m^* taken to be that of an electron in the corresponding crystal, $\varepsilon_L \equiv (\hbar^2/2m^*L^2)$ is then in the range 0.3–1 eV for many materials. Alternatively ε_L may be regarded as a fitting parameter which is of order 1 eV to be physically reasonable.[2] B_0 is the square of the variance of the potential fluctuations and is also expected to be of the order of ~ 1 (eV)^2 for the materials under consideration. Therefore, the linear energy regions in Figs. 1 and 4 are consistent with the experimentally observed Urbach-band tails. One exception is found for the case with an autocorrelation function of $\exp(-|x|/L)$, where the linear behavior only exists for a considerably small energy $|E|/\varepsilon_L$ (see Fig. 2). If this correlation function is used as a model of screened Coulomb impurities, it would require a large ε_L and therefore either an unusually small screening length L or effective mass m^* to produce an Urbach tail over many decades.

The magnitude of the variations in the density of states associated with the linear region can also be extracted from Figs. 1 and 4. Since $\rho(E) \sim e^{-S(E)}$, the number of decades of linear behavior in the density of states is equal to the change in S divided by $\log_e 10$ (=2.3). From the results in Figs. 1 and 4, it is clearly seen that even a small portion of the linear region would correspond to a change of many decades.

Comparison of the slope of the observed Urbach tail with those obtained by the models we have described provides insight into the actual form of disorder present in various materials. The slopes we calculated for various autocorrelation functions range from 9.6 to 58.2 in units of ε_L/B_0 (see Table II). With the estimates of $\varepsilon_L \sim$ eV and $B_0 \sim 1$ eV^2, these slopes are in the range of (0.1 eV)$^{-1}$ to (0.02 eV)$^{-1}$. The typical value in the optical absorp-

tion experiments for a-SiH_x is about $(0.05-0.08\ eV)^{-1}$.[25] Likewise in the photocurrent measurements the slope is about $(0.02\ eV)^{-1}$ for glassy As_2Se_3,[22] and $(0.03-0.04\ eV)^{-1}$ for a-SiH_x.[23] It is evident that the models which we have considered can provide accurate descriptions of both of these materials. Moreover, we anticipate that a very large class of materials may be modeled using physically reasonable choices of length and energy scale parameters entering the autocorrelation functions which we have presented.

The primary conclusion is that the shape of Urbach tails in disordered materials provide a sensitive measure of the microscopic spatial autocorrelations in the random potential. The observed linearity of Urbach tails in a variety of materials suggests strong short-range order on the scale of the interatomic distance but correlations which decay more rapidly than exponentially on longer length scales. An independent, microscopic investigation of shape of the autocorrelation function would be of considerable importance in testing our hypothesis, which we leave to future investigation.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation (NSF) Grant No. DMR-80-20263 at Princeton University and by a grant from Exxon Research and Engineering Company. C. M. Soukoulis wants to thank Professor K. M. Ho for discussions about numerical techniques.

APPENDIX: THE SHIFTED CONTINUUM BAND EDGE

For large positive energies E, the density of states in three dimensions approaches the well-known square-root continuum: $\rho(E)\sim\sqrt{E}$ $(E\rightarrow+\infty)$. This continuum extrapolates to a small negative energy-band edge which is now estimated for the weak-disorder limit.

The averaged one-electron Green's function may be calculated in perturbation theory for extended nearly plane-wave states. The field theory representation of Secs. II and IV yields a simple effective medium approximation:[28]

$$G^{\dagger}(k)\equiv\int d^dx\, e^{ik\cdot x}\langle G(x,0,E_+)\rangle_{\text{ensemble}} \tag{A1a}$$

$$\simeq\frac{1}{\hbar^2k^2/2m^*-E_++Q(k)}\ , \tag{A1b}$$

where the self-energy correction

$$Q(k)=\int\frac{d^3q}{(2\pi)^3}\tilde{B}(k-q)G^{\dagger}(q)\ , \tag{A2a}$$

and we have introduced the Fourier transform of the autocorrelation function

$$\tilde{B}(k)\equiv\int d^3x\, e^{ik\cdot x}B(x)$$

$$=V_{\text{rms}}^2(\pi L^2)^{3/2}e^{-k^2L^2/4}\ . \tag{A2b}$$

The location of the extrapolated continuum edge occurs at $E=\text{Re}Q(0)$ in this approximation. For $E<0$ and to leading order in the scattering strength V_{rms}^2,

$$\text{Re}Q(0)\simeq-\frac{V_{\text{rms}}^2}{2\sqrt{\pi}\,|E|}\left[\frac{L}{\lambda}\right]^3\int_0^{\infty}dq\frac{q^2e^{-L^2q^2/4\lambda^2}}{1+q^2}\ , \tag{A3}$$

where we have rescaled the integration variation in (A2a) by the deBroglie wavelength $\lambda\equiv\hbar^2/(2m\,|E|)$. For weak disorder, near the band edge λ is much larger than L so that the dominant contribution to the integral (A3) comes from the region $1\ll q<2\lambda/L$. Replacing $q^2/(1+q^2)$ in Eq. (A3) by unity yields

$$\text{Re}Q(0)\simeq-\frac{V_{\text{rms}}^2}{2\varepsilon_L},\quad\varepsilon_L\equiv\frac{\hbar^2}{2mL^2} \tag{A4}$$

for the location of the continuum edge. The actual density of states remains smooth at this energy. However, this energy represents a crossover point below which the density of states is dominated by nonperturbative contributions (instantons) and above which perturbative corrections in V_{rms}^2 give rise to the high-energy square-root continuum.

*Permanent address: Ames Laboratory and Department of Physics, Iowa State University, Ames, Iowa 50011.

[1]F. Urbach, Phys. Rev. **92**, 1324 (1953); W. Martienssen, J. Phys. Chem. Solids **2**, 257 (1957).
[2]W. Sritrakool, V. Sayakanit, and H. R. Glyde, Phys. Rev. B **33**, 1199 (1986); V. Sayakanit and H. R. Glyde, Comments Cond. Matter Phys. (to be published).
[3]S. John, C. Soukoulis, M. H. Cohen, and E. Economou, Phys. Rev. Lett. **57**, 1777 (1986).
[4]B. I. Halperin and M. Lax, Phys. Rev. **148**, 722 (1966).
[5]E. O. Kane, Phys. Rev. **131**, 79 (1963).
[6]V. L. Bonch-Bruevich and A. G. Mironov, Fiz. Tverd. Tela (Leningrad) **3**, 3009 (1962) [Sov. Phys.—Solid State **3**, 1219 (1962)].
[7]S. John and M. J. Stephen, J. Phys. C **17**, L559 (1984).
[8]J. L. Cardy, J. Phys. C **11**, L321 (1978).
[9]J. D. Dow and D. Redfield, Phys. Rev. B **5**, 594 (1971).
[10]H. Sumi and Y. Toyozawa, J. Phys. Soc. Jpn. **31**, 342 (1971).
[11]T. Skettrup, Phys. Rev. B **18**, 2622 (1978).
[12]S. Abe and Y. Toyozawa, J. Phys. Soc. Jpn. **50**, 2185 (1985).
[13]M. Schreiber and Y. Toyozawa, J. Phys. Soc. Jpn. **51**, 1544 (1982).
[14]V. Sayakanit, J. Phys. C **7**, 2849 (1974).
[15]R. P. Feynman and H. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[16]P. Lloyd and P. R. Best, J. Phys. C **8**, 3752 (1975).
[17]E. Brezin and G. Parisi, J. Stat. Phys. **19**, 269 (1978); J. Phys. C **13**, L307 (1980).
[18]J. D. Wiley, D. Thomas, E. Schönherr, and A. Breitschwerdt, J. Phys. Chem. Solids **41**, 801 (1980).

[19]E. Mohler and B. Thomas, Phys. Rev. Lett. **44**, 543 (1980).
[20]G. D. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. **47**, 1480 (1981).
[21]D. J. Dunstan, J. Phys. C **30**, L419 (1982).
[22]D. Monroe and M. A. Kastner, Phys. Rev. **33**, 8881 (1986).
[23]T. Tiedje, J. M. Cebulka, D. L. Morel, and B. Abeles, Phys. Rev. Lett. **46**, 1425 (1981).
[24]G. D. Cody, in *Hydrogenated Amorphous Silicon,* Vol. 21B of *Semiconductors and Semimetals,* edited by J. Pankove (Academic, New York, 1984), p. 11.
[25]B. Abeles, C. R. Wronski, T. Tiedje, and G. D. Cody, Solid State Commun. **36**, 537 (1980).
[26]G. Pfister and H. Scher, Adv. Phys. **27**, 747 (1978).
[27]J. Orenstein, M. A. Kastner, and V. Vaninov, Philos. Mag. B **46**, 23 (1982); D. Monroe and M. A. Kastner, *ibid.* **47**, 605 (1983).
[28]S. John and M. J. Stephen, Phys. Rev. B **28**, 6358 (1983).