Coulomb gap in two- and three-dimensional systems: Simulation results for large samples

A. Möbius and M. Richter
Institut für Festkörperforschung, Institut für Festkörper- und Werkstofforschung, Dresden, e.V., D-O-8027 Dresden, Germany

B. Drittler*
Institut für Festkörperforschung, Forschungszentrum Jülich G.m.b.H., D-W-5170 Jülich, Germany
(Received 13 May 1991; revised manuscript received 10 October 1991)

Computer experiments have been performed to study the single-particle density of states, $g(E)$, in the Coulomb gap. This gap occurs in disordered insulating systems as an effect of the long-range tail of the Coulomb interaction. In order to compare these numerical results with the analytical theory on a broader energy scale than previous authors, very large samples have been considered, up to 40 000 and 125 000 sites for dimensions $d=2$ and 3, respectively. Special algorithms have been developed for this aim. Our numerical results contradict the analytical theory from the literature in two main points: As $E\rightarrow\mu$, the numerical values for $g(E)$ are considerably smaller than the analytical predictions, and universality with respect to disorder is not present. It could not be decided finally whether or not $g(E\rightarrow\mu)$ follows a power law, $g(E)\sim|E-\mu|^{\nu}$. Provided it does, the simulation results, $\nu=1.2\pm0.1$ and 2.6 ± 0.2 for $d=2$ and 3, respectively, deviate from the analytical prediction, $\nu=d-1$. Moreover, as a first approach to the polaronic transport problem, the influence of relaxation down to stability with respect to all simultaneous two-electron hops is studied. In this case, $g(E)$ is diminished, but not changed qualitatively. In particular, the exponential behavior analytically predicted is not observed.

I. INTRODUCTION

The Coulomb gap is a long-range correlation effect occurring in disordered insulating systems, such as lightly doped crystalline semiconductors or amorphous semiconductors, at sufficiently low temperatures; for reviews see Refs. 1–3. Its significance for granular metals is discussed controversially.[4] The Coulomb gap can directly be detected by tunneling[5] and possibly also by photoemission measurements.[6] Moreover, it has a pronounced influence on the temperature dependence of the conductivity in the variable-range-hopping region.[7,8]

The first theoretical papers on the influence of the Coulomb interaction on the density of states near the Fermi energy appeared two decades ago. Pollak and Srinivasan showed by means of analytical studies that the density of states exhibits a minimum.[9] This result was corroborated numerically by Kurosawa and Sugimoto.[10] Efros and Shklovskii[7] pointed out that the long-range tail of the unscreened Coulomb interaction induces a pseudogap in the single-particle density of states $g(E)$ at the Fermi energy μ. This means $g(E)$ tends to zero as E approaches μ, but $g=0$ holds only for $E=\mu$. They considered a semiclassical model of a dilute impurity band of a semiconductor, viz., a disordered system of localized states without quantum interference,

$$H=\sum_i \varphi_i n_i+\tfrac{1}{2}\sum_{i,j;i\neq j} e^2 n_i n_j/(\kappa|\mathbf{x}_i-\mathbf{x}_j|)\ , \tag{1}$$

where n_i denotes the occupation number, $n_i\in\{0,1\}$, for an elementary charge e, located at $\mathbf{x}_i$, and where κ is the dielectric constant. (In this description, for n-type semiconductors, $n_i=1$ corresponds to an empty, i.e., ionized donor.) The random potential φ_i arises from fixed charges at random positions (e.g., acceptors in n-type semiconductors) and from short-range disorder (in particular, in amorphous semiconductors). This model is applicable, if (i) the decay length of the localized states is small compared with the nearest-neighbor distance of the impurity atoms, and if (ii) the temperature is sufficiently low so that screening, e.g., by electrons excited to the conduction band, can be neglected.

The model (1) should permit the study of static properties of the low-lying metastable states. But it does not describe any dynamics; hopping rates have to be phenomenologically included in transport studies. In the following, we shall restrict ourselves to the investigation of zero-temperature static properties.

The single-particle energy E_i, which is needed to bring a charge e from infinity to $\mathbf{x}_i$ without allowing for a simultaneous rearrangement of the charges on the other sites, results from the relation

$$E_i\{n_k\}=\frac{\delta H\{n_k\}}{\delta n_i}=\varphi_i+\sum_{j;j\neq i} e^2 n_j/(\kappa|\mathbf{x}_i-\mathbf{x}_j|)\ . \tag{2}$$

It should be noted that the values of E_i depend on the full set of occupation numbers, $\{n_k\}$. The single-particle density of states $g(E)$ is given by the probability distribution of the values of E_i.

As in spin glasses,[11] the macroscopic properties of real physical systems, described by (1), are determined by low-lying metastable states. These states are highly degenerate. For an isolated system, the excitations of lowest complexity are one-electron hops. By a qualitative analytical consideration, Efros and Shklovskii showed for

two- and three-dimensional systems that the stability with respect to one-electron hops implies the following asymptotic behavior of $g(E)$ as $E \to \mu$:

$$g(E) = \alpha_d (\kappa/e^2)^d |E-\mu|^{d-1} , \tag{3}$$

where d denotes the dimensionality and α_d a dimensionless constant. To be more precise, Efros and Shklovskii obtained this expression as an upper bound on g. By use of a self-consistency argument, they concluded that g should have just this value.

Efros[12] confirmed the above result by means of a more elaborate analytical theory. He obtained a so-called self-consistent equation,

$$g(\tilde{E}) = g_0 \exp\left[-C_d \int_0^\infty \frac{g(\tilde{E}')d\tilde{E}'}{(|\tilde{E}|+\tilde{E}')^d} \right] , \tag{4}$$

where $\tilde{E} = E-\mu$, $C_2 = \pi e^4/(2\kappa^2)$, and $C_3 = 2\pi e^6/(3\kappa^3)$. Roughly speaking, the prefactor g_0 denotes the density of states with the Coulomb interaction switched off. The solution of this integral equation obeys the relation (3) as $E \to \mu$, where the constants α_d are given by

$$\alpha_d = d/\pi , \tag{5}$$

cf. Ref. 13. It should be stressed that the asymptotic behavior [(3) and (5)] is universal, i.e., independent of g_0.

Bounds on $g(E)$ arising from stability with respect to (simultaneous) two-electron hops were analytically studied by Efros[12] and Baranovskii, Shklovskii, and Efros.[13] The latter authors obtained exponential behavior for $d=3$,

$$g(E) \sim \exp[-\gamma y/(\ln y)^{7/4}], \quad y = \Delta/|E-\mu| , \tag{6}$$

where γ is a dimensionless constant and Δ denotes the width of the Coulomb gap. The question of the influence of two-electron hops and even more complex excitations is closely related to the polaron transport problem. The role of dressed charged excitations has been discussed controversially in the literature.[3,14-17]

The analytical investigations mentioned above have been tested by a series of computer experiments.[14-16,18-21] Part of this work studies a simplified model, where the sites form a regular lattice and the disorder arises from stochastic energy shifts of the individual sites. All computer studies yield the same result qualitatively: Stability with respect to one-electron hops causes the occurrence of a Coulomb gap, and $g(E)$ seems to be universal as $E \to \mu$. However, considerable quantitative differences between the $g(E)$ dependences calculated are present: For $d=3$, Baranovskii, Efros, Gelmont, and Shklovskii[14] obtained a quadratic E dependence, coinciding with Eq. (3), whereas Davies, Lee, and Rice[15] obtained a steeper $g(E)$, possibly a clue to exponential behavior. Therefore, it is not clear if the self-consistent equation (4) can be justified by numerical simulation.

It seems to be useful to compare the two- and three-dimensional cases with the one-dimensional problem at this point. Both the numerical simulation[22] and the analytical self-consistent equation method[23] yield a logarithmic dependence,

$$g(E) = \frac{g_1}{\ln(E_c/|E-\mu|)} , \tag{7}$$

where g_1 and E_c are nonuniversal parameters. Nevertheless, $g(E)$ is considerably overestimated by the analytical study.

Considering the above points of view, one cannot regard as satisfactory the usual knowledge on the density of states within the Coulomb gap. The aim of the present numerical simulations is threefold.

(a) To perform a comparison with analytical theory on a broader energy scale than previous authors.

(b) To elucidate the contradiction between the investigations by Baranovskii, Efros, Gelmont, and Shklovskii[14] and Davies, Lee, and Rice[15]—in the following referred to as BEGS and DLR—.

(c) To study the influence of additional stability with respect to two-electron hops on the density of states.

We developed relaxation algorithms, based on the branch-and-bound method known from combinatorial optimization.[24-26] These algorithms allow the study of larger systems and, in this way, an extension of the energy region where reliable results on $g(E)$ can be obtained to smaller values of $|E-\mu|$. For early versions of the present investigation see Ref. 21.

Our report is organized as follows. The model is introduced in detail in Sec. II. The relaxation procedure is described in Sec. III. Our numerical results are presented in Sec. IV, and conclusions, in particular those concerning analytical theories and transport, are discussed in Sec. V.

II. THE MODEL

We consider a d-dimensional simple cubic lattice of L^d sites, localized at $\mathbf{x}_i$,[14,15] where the lattice constant is denoted by a. A dimensionless description is used, i.e., $a=1$, $e=1$, and $\kappa=1$. The occupation numbers of the sites n_i take the values 0 and 1 only. The mean filling factor of the sites, n_i/L^d, is denoted by K. The disorder is simulated by a stochastic potential φ_i. The values of the φ_i are uniformly distributed between $-B/2$ and $B/2$; they are uncorrelated. Let the Coulomb interaction energy of charges e at sites i and j be f_{ij}. We restrict ourselves to the case of low impurity concentration with the only exception of Sec. IV E. That means we assume the nearest-neighbor distance is large compared with the radius of the localized states, so that $f_{ij} = 1/|\mathbf{x}_i - \mathbf{x}_j|$. In order to establish electroneutrality for the sample considered, we attach a charge $-K$ to each site. Thus we have

$$H = \sum_i \varphi_i n_i + \tfrac{1}{2} \sum_{i,j;i\neq j} f_{ij}(n_i-K)(n_j-K) . \tag{8}$$

Within this approximation, the states of H are characterized by the full set of occupation numbers, $\{n_i\}$. Note that Eq. (8) is equivalent to Eq. (2) of BEGS for $K=\frac{1}{2}$. In this case, the Hamiltonian is invariant with respect to the transformation $n_i \to 1-n_i$ and the single-particle den-

sity of states is symmetric about μ.

The question of boundary conditions is very important for the convergence of simulation results as $L \to \infty$. BEGS (Ref. 14) used free boundary conditions, where the surroundings of the sample studied are fully neglected,

$$f_{ij} = 1/r_{ij} \quad \text{with } r_{ij}^2 = \sum_{\beta=1}^{d} (x_{\beta,i} - x_{\beta,j})^2 \ . \tag{9}$$

These free boundary conditions have the disadvantage that sites close to the surface "feel a lower dimensionality" than sites deep inside the sample. In order to suppress this surface contribution, we adopt the periodic boundary conditions introduced by DLR.[15] That means we consider our sample as an elementary cell of a repeated lattice and assume that the interaction energy of two sites is determined by the shortest distance between them within this repeated lattice,

$$r_{ij}^2 = \sum_{\beta=1}^{d} (\min\{|x_{\beta,i} - x_{\beta,j}|, L - |x_{\beta,i} - x_{\beta,j}|\})^2 \ . \tag{10}$$

In this way, all sites are equivalent; they seem to be positioned in the middle of the finite sample.

The fluctuation of the Fermi level of isolated samples is a further finite-size effect, which limits the accuracy of computer simulations. In particular, sufficient statistics can only be reached by averaging a certain number N_s of samples, since the computational effort increases very fast with L. This problem was dealt with in the literature in a rather artificial way, by so-called simple averaging or by μ averaging.[14] In the first case, the density of states is calculated by simply adding the spectra of the individual samples, whereas the second method includes a shift of each individual spectrum to adjust the Fermi energy before adding the spectra. DLR took μ as mean of the random φ_i,[15] but this idea is only applicable if $K = \frac{1}{2}$. We approach this difficulty by considering the grand canonical Hamiltonian,

$$h = H - \mu \sum_i n_i \ , \tag{11}$$

instead of H, which significantly reduces the corresponding numerical errors. That means, we use relaxation procedures allowing not only for energy exchange with a heat bath but also for particle exchange with a reservoir at infinity, which fixes the chemical potential μ. However, the parameters μ and K are not independent of each other in this scheme. The electroneutrality condition,

$$\left\langle \sum_i n_i \right\rangle_{\mu,K} = KL^d \ , \tag{12}$$

has to be fulfilled, where the angular brackets denote ensemble averaging.

The single-particle energies e_i of the grand canonical Hamiltonian h are

$$e_i = E_i - \mu = \varphi_i + \sum_{j;j\neq i} f_{ij}(n_j - K) - \mu \ , \tag{13}$$

where E_i denotes the single-particle energies of H. The energy e_i is needed to bring a particle from the reservoir at infinity to an empty site i where the occupation of all other sites remains unchanged; it is gained when a particle is removed from an occupied site i. Note that during this process the single-particle energies of all other sites j are changed by f_{ij} or $-f_{ij}$, respectively. The aim of our simulations is the calculation of the single-particle density of states $g(e)$.

Now we turn to the hierarchy of conditions determining the ground state of h according to the demand that every rearrangement, charged or uncharged, must be accompanied by an increase of h. Thus the ground state must be stable with respect to the following rearrangements.

(i) Single-particle exchange with the reservoir, i.e.,

$$n_i = 1 \implies \Delta h = -e_i > 0$$

and

$$n_i = 0 \implies \Delta h = e_i > 0 \ . \tag{14}$$

(ii) Single-particle hops within the system, i.e.,

$$n_i = 1 \ , \ n_j = 0 \implies \Delta h = e_j - e_i - f_{ij} > 0 \ . \tag{15}$$

[This condition can easily be understood by decomposing the single-particle hop into two steps: (a) moving a particle from site i to infinity ($\Delta h_{(a)} = -e_i$), where e_j is changed to $e_j - f_{ij}$, and (b) moving it from infinity to site j ($\Delta h_{(b)} = e_j - f_{ij}$).]

(iii) Excitations composed of one single-particle exchange with the reservoir and one single-particle hop within the system, i.e.,

$$n_i = n_j = 1 \ , \ n_k = 0 \quad (i \neq j)$$
$$\implies \Delta h = e_k - e_i - e_j + f_{ij} - f_{ik} - f_{jk} > 0 \ , \tag{16a}$$

and

$$n_i = 1 \ , \ n_j = n_k = 0 \quad (j \neq k)$$
$$\implies \Delta h = e_j + e_k - e_i + f_{jk} - f_{ij} - f_{ik} > 0 \ . \tag{16b}$$

(iv) Two-particle hops within the system, i.e.,

$$n_i = n_j = 1 \ , \ n_k = n_l = 0 \quad (i \neq j, k \neq l)$$
$$\implies \Delta h = e_k + e_l - e_i - e_j + f_{ij} + f_{kl} - f_{ik} - f_{il} - f_{jk} - f_{jl} > 0 \ . \tag{17}$$

(v) Etc.

This hierarchy also permits a natural classification of the metastable states with respect to their decay rates. In fact, the transition probability to another state of lower energy decreases exponentially with the number of particles involved. This is the reason why the physical properties of disordered systems at low temperatures are determined by highly degenerate but "frozen" metastable states rather than by the ground state. Moreover, these states are macroscopically identical (the number of "exotic," highly symmetric states has zero measure). Hence, it should be sufficient to consider only the first levels of the hierarchy above. The highest degree of metastability taken into account in our simulations corresponds to condition (iv).

The conditions (iii), (iv), etc. seem to be related to the

polaron problem.[3] In our opinion, however, this difficult dynamical problem lies beyond the scope of the model studied. In order to simulate polarons, one has to relax a certain surrounding of the site considered. The delicate question is the choice of the permissible rearrangements. To take, e.g., the full set for an isolated system, viz., (ii), (iv), etc. would mean the relaxation to the ground state of this isolated system, where $\Delta h = \mu$ is independent of the site considered, which would be unphysical. DLR take into account only an additional relaxation via single-particle hops after a particle has been added to or taken away from the system. But compact two-particle hops should not be less probable than two successive long-range one-particle hops. The situation concerning more complex excitations is similar. Therefore, we believe that it is not justified to define "polarons" within the poor models (8) and (11), and we restrict ourselves to the investigation of classes of low-lying metastable states fulfilling certain of the conditions given above. Nevertheless, some useful information concerning polaronic transport should be obtained in this way.

Condition (ii) implies a further finite-size effect: Since all f_{ij} exceed a certain bound, $\delta_{fs} = 2/(Ld^{1/2})$, the lowest empty and the highest occupied state of any finite system, stable with respect to single-particle hops, are separated at least by an artificial hard gap of width δ_{fs}. It is obvious that δ_{fs} is a lower bound for the energy region of reliability of the simulation results for $g(e)$.

III. RELAXATION ALGORITHM

The determination of the ground state of the Hamiltonian (11) belongs to the class of the so-called NP-complete optimization problems.[27] For such problems, no algorithm is known which yields the optimum with an effort proportional to any power of the system size. For large systems, one has to look for some kind of approximate solutions. Usually, relaxation procedures are applied, which generally yield only some low-lying metastable state. Nevertheless, these metastable states should be the physically interesting ones, cf. above.

A natural relaxation procedure consists of three parts: Start from a randomly occupied system, thermalization according to the Metropolis algorithm[28] at a temperature large compared to Coulomb and disorder energies, and cooling down to zero with the temperature being diminished according to some schedule. This procedure is known as simulated annealing.[29] It was applied to the Coulomb glass problem by Summerfield, McInnes, and Butcher.[20] It should be stressed that the final state depends on the cooling rate and on the "complexity" of the rearrangements taken into account. Its energy expectation value decreases with increasing relaxation time, as well as with increasing complexity.

From this point of view, our algorithm can be described as follows: Start at $T = \infty$ (random occupation) and relaxation by contact with a heat bath at $T = 0$ (highest possible rate). The final state depends on the transformations considered within the relaxation steps, cf. the hierarchy given in Sec. II. It turned out that the relaxation path is of minor influence.

First, we consider only the relaxation down to stability with respect to the conditions (i) and (ii), as introduced in Sec. II. Numerical experience shows that the number of relaxation steps is roughly proportional to the sample size L^d. One relaxation step consists of two parts: (a) the search for a transformation diminishing h and (b) the recalculation of all e_i and change of n_i. The crucial part with respect to the performance of a program is (a). Using the BEGS algorithm, viz., relaxation via the first energy diminishing pair found, the total CPU time needed for this search is proportional to L^{2d}, up to some logarithmic corrections. Nearly the same holds for part (b). However, the prefactor for (a) is comparatively large. Moreover, it seems to be impossible to parallelize effectively (a) contrary to (b).

DLR used in each relaxation step the rearrangement with maximal energy gain. The simplest such programs require a total effort proportional to L^{3d} (L^{2d} per step). DLR avoided unnecessary tests to a large extent. Possibly they used ideas similar to those described below. In that case, the total CPU time for (a) would be proportional to L^{2d}, up to logarithmic corrections. But the prefactor should considerably exceed that of the BEGS algorithm.

We followed BEGS in performing the first h diminishing rearrangement which was found in each relaxation step. The search for it, i.e., (a), was greatly speeded up by constructing a branch-and-bound-type algorithm. This algorithm needs a total CPU time roughly proportional to $L^d \ln L$ for (a). However, the corresponding prefactor is comparatively large. Nevertheless, the recalculation of the e_i, which needs a total effort proportional to L^{2d}, is expected to be the most expensive part for sufficiently big samples.

Now, let us consider relaxation down to stability with respect to conditions (iii), etc. We performed the relaxation "levelwise": First, the sample is relaxed with respect to (i) and (ii). After this, an h diminishing 3-site rearrangement is searched for. As soon as one is found, it is performed, and, moreover, the sample is relaxed concerning (i) and (ii). Then, an h diminishing 3-site rearrangement is searched for, and so on.

For the search within the set of p-site excitations, the simplest programs need a total effort proportional to L^{pd} (up to logarithmic corrections). Thus, at first glance, it seems hopeless to investigate samples of reasonable size with respect to (i)–(iv). However, as in the previous case, the search could enormously be speeded up by use of branch-and-bound-type algorithms. Again, the effort for searching in the highest level of the hierarchy considered turned out to be roughly proportional to $L^d \ln L$, but the prefactor increases fast with the complexity of the rearrangements.

The basic idea of branch-and-bound algorithms, developed originally for the exact solution of combinatorial optimization problems, can be visualized in the following way. Consider a transformation changing the occupation of p sites of a state $s^{(0)}$. The number of all such rearrangements is of the order L^{pd}. The formation of them can be visualized as the construction of a tree. This tree roots at $s^{(0)}$ and includes p subsequent branching lev-

els, where $O(L^d)$ new branches arise from every branching point, $s_{i_k}^{(k)}$ with $k=0,\ldots,p-1$, so that $\text{card}\{s_{i_k}^{(k)}\}=O(L^{kd})$, cf. Fig. 1(a). Appropriate inequalities on the values of h for all those states $s_i^{(p)}$, which can be reached by starting at a certain branch, permit mostly the search for "better" states to be terminated at an early stage of the evaluation of the corresponding subtree. A further improvement can be reached by an appropriate ordering scheme for the branches starting from the same point, cf. Fig. 1(b).

Kobe and Hartwig[25] used a branch-and-bound algorithm for the exact determination of the ground state and of all low-lying excitations of a spin-glass cluster. Unfortunately, such investigations can only be performed for fairly small samples, at present less than roughly 100 sites. Nevertheless, the branch-and-bound idea proves to be useful also for the consideration of very large systems, where it can be utilized for complex relaxation steps in approximation procedures.[26]

We turn now to the practical realization of this idea for the relaxation down to states which are stable with respect to conditions (i) and (ii). Let M_1 be the set of all single-particle exchanges with the reservoir, M_2 the set of all single-particle hops within the system, $M_{2,1}$ the set of all nearest-neighbor single-particle hops within the system, $M_{2,2}$ the set of all next-nearest-neighbor hops, etc. The values of the Coulomb interaction energy, related to nearest neighbors, next-nearest neighbors, and so on, are denoted by $f_1, f_2, \ldots$. The relaxation down to states being stable with respect to $M_1 \cup M_2$ is performed stepwise.

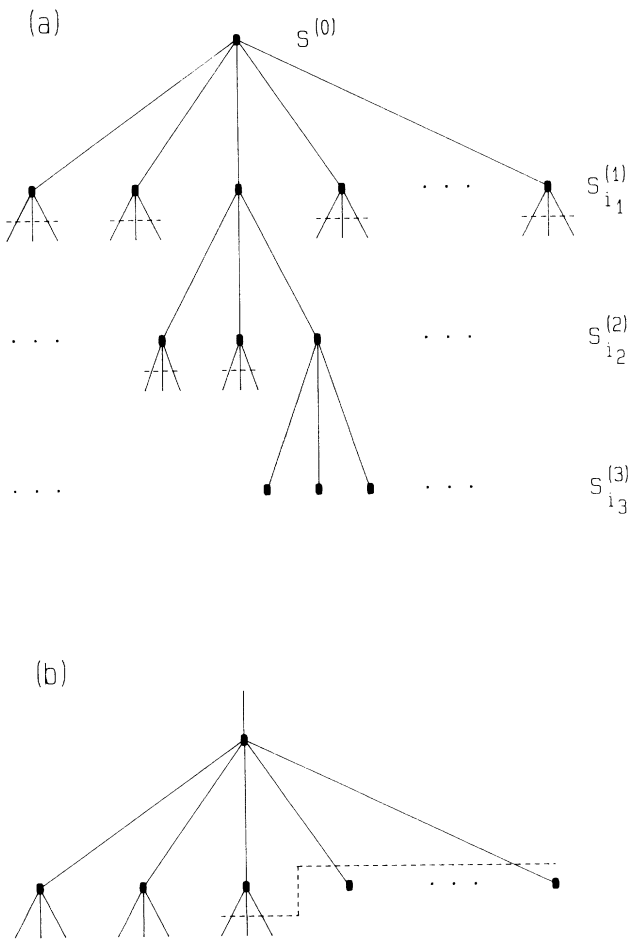

FIG. 1. Examples of branch-and-bound trees: (a) evaluation of all 3-site excitations, and (b) "horizontal" elimination by means of an ordering scheme.

(1) Only M_1 is considered at first. We start at a randomly chosen site and check for this and, if not successfully, for the other sites in succession, whether a particle exchange between that site and the reservoir diminishes h. If so, the occupation is changed and all e_i are recalculated. Then we randomly choose a new site and so on.

(2) After stability with respect to M_1 has been reached, we turn to $M_1 \cup M_{2,1}$ and relax the system analogously to step (1). At this stage, we check a considered site i at first with respect to M_1. After this, only if $n_i=0$ and $e_i \le f_1$, we take into consideration all nearest neighbors j and check whether $n_j=1$ and $e_i-e_j-f_1<0$. The idea is that the latter inequality cannot hold if $e_i>f_1$, which enables many of the "nearest-neighbor tests" to be skipped.

(3) The next step is the consideration of $M_1 \cup M_{2,1} \cup M_{2,2}$. After checking analogously to step (2) the site i with respect to $M_1 \cup M_{2,1}$, we ask whether $n_i=0$ and $e_i \le f_2$. A check with respect to $M_{2,2}$ is performed only if this condition is fulfilled. Note that the latter inequality presupposes $e_i \le f_1$. Thus, if a site can be eliminated from the consideration of $M_{2,1}$, a search within $M_{2,2}$ is useless and can be skipped.

(4) Etc.

The search with respect to $M_1 \cup M_{2,1} \cup \cdots \cup M_{2,k}$ is accompanied by tabulating all sites with $n_i=0$ and $e_i \le f_{k+1}$. Performing the transition to the next level of our relaxation procedure, we have to consider only the sites, which are contained in this table, up to the next rearrangement. Moreover, as soon as this table is empty after the full search within that set, the above procedure can be terminated.

According to our experience, the largest part of the nearest-neighbor hops is performed at first. This seems very reasonable from the physical point of view.

The relaxation down to states which are stable with respect to the conditions (i)–(iv) of Sec. II is performed in a similar manner. The most compact excitations are very important in this case, so they are treated separately. For the consideration of the other excitations, we constructed branch-and-bound algorithms analogous to the procedure described above, where the inequalities used are based on stability with respect to less complex rearrangements. An ordering of the distances between the sites involved is very useful.

The final computer experiments, described below, were performed on a Cray X-MP, where the calculation of the initial e_i as well as the recalculation of the e_i are vectorized to a large extent. We started from stochastic initial occupations and made the following practical observations.

(a) Stability with respect to single-particle hops [conditions (i) and (ii)]: For $d=2$, $L=200$, the total relaxation (search and recalculation of e_i) was only by a factor

2.0–3.0 more expensive than the calculation of the initial e_i, whereas for $L=50$, this factor is 3.4–5.2. This behavior can be well understood in terms of the above-mentioned exponents. For $d=3$, $L=50$, the corresponding factor amounts to 1.1–1.7.

(b) Stability with respect to two-particle hops [conditions (i)–(iv)]: The total relaxation needed by nearly a factor of 40 more CPU time than the calculation of the initial e_i, both for $d=2$, $L=200$ and for $d=3$, $L=50$. Thus the construction of the states fulfilling the conditions (i)–(iv) is more expensive than that of the states, fulfilling solely (i) and (ii), by only one order of magnitude.

IV. NUMERICAL RESULTS

A. General behavior

Overall views of the single-particle density of states $g(e)$ are given for different degrees of disorder B and filling of the band K in Figs. 2(a)–2(d) for two- and three-dimensional systems, respectively. Relaxation has been performed down to metastable states fulfilling the conditions (i) and (ii). The results presented agree well with the Figs. 3(b) and 3(a) by DLR.

Figures 2(a)–2(d) show the universal existence of a Coulomb gap. According to Figs. 2(a) and 2(b) it seems likely that $g(e)$ decreases faster than linear as $e\to 0$ for $d=2$. But, since the asymptotic regions are fairly small, it is not possible to draw definite conclusions as to appropriate quantitative descriptions. Moreover, it cannot be checked whether $g(e)$ is indeed universal with respect to B and K. A detailed analysis presupposes logarithmic representations, as they are considered in the following sections.

B. Region of reliability

Before investigating the asymptotic behavior in detail, we have to exclude artifacts arising from the finite sample

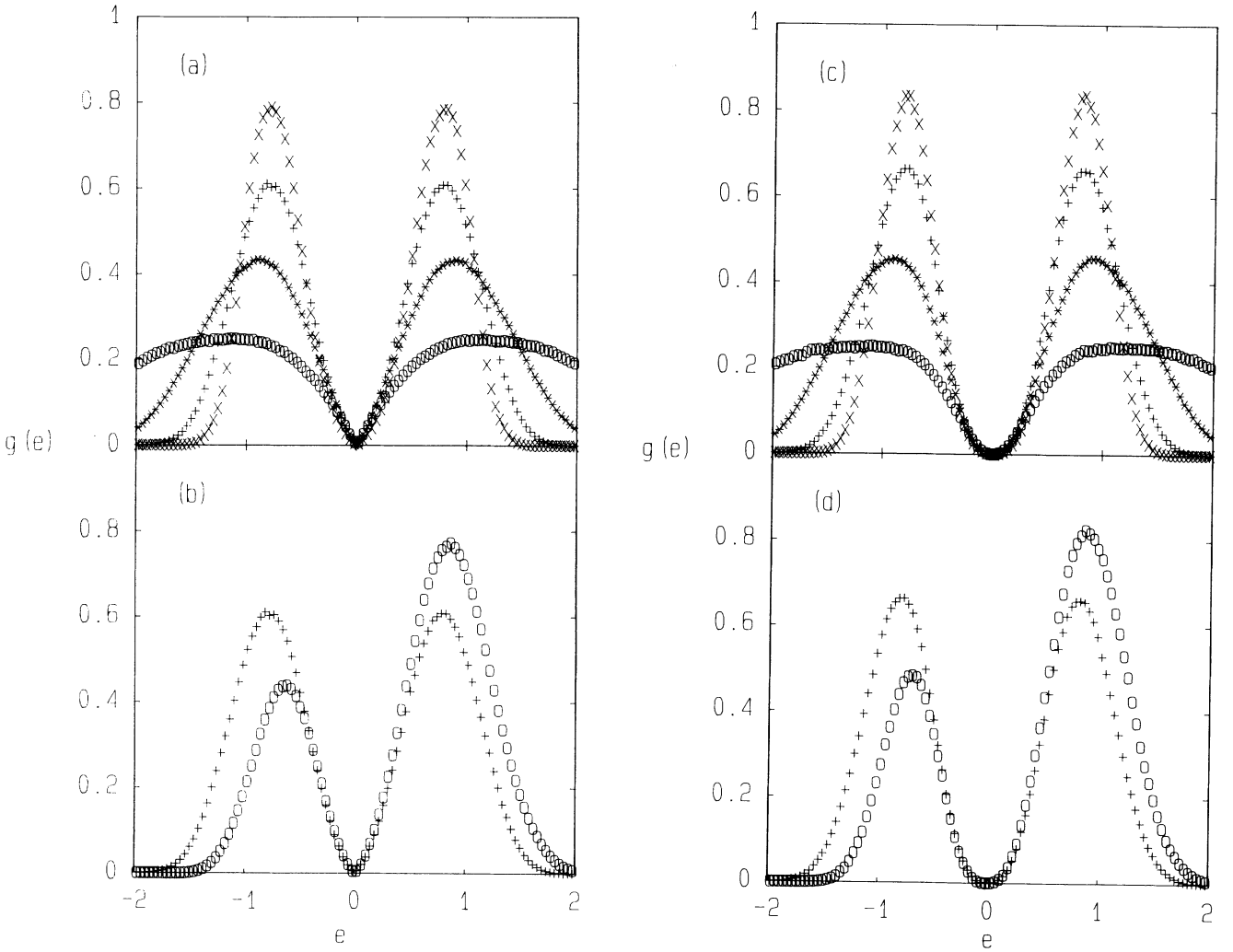

FIG. 2. Overall view of the single-particle density of states $g(e)$ for two-dimensional [(a) and (b)] and three-dimensional systems [(c) and (d)] in linear representations. Different degrees of disorder are compared in (a) and (c) for $K=0.5$: $B=0.5$(×), $B=1$ (+), $B=2$ (*), and $B=4$ (○). The influence of band filling is demonstrated in (b) and (d) for $B=1$: $K=0.5$ (+) and $K=0.3$ (○). For $K=0.3$, the self-consistent solution of the electroneutrality condition (12) yields $\mu=-0.4245$ and -0.420 for the two- and three-dimensional cases, respectively. Symmetry implies $\mu=0$ for all parameter sets with $K=0.5$. Data are obtained by averaging N_s samples of size L^d, which are relaxed down to metastable states satisfying the conditions (i) and (ii) of Sec. II. We consider $L=200$, $N_s=100$ for $d=2$, except for $B=1$, $K=0.3$, where $N_s=200$, and we use $L=50$, $N_s=16$ for $d=3$.

size L^d. For this aim, Figs. 3(a) and 3(b) compare the results for different L in $\log_{10}$-$\log_{10}$ representations, where bin widths proportional to $|e|^{1/3}$ are used, for $d=2$ as well as for $d=3$. It should be mentioned that the L values considered here are larger than those considered by DLR by factors of 10 and 5 for $d=2$ and 3, respectively.

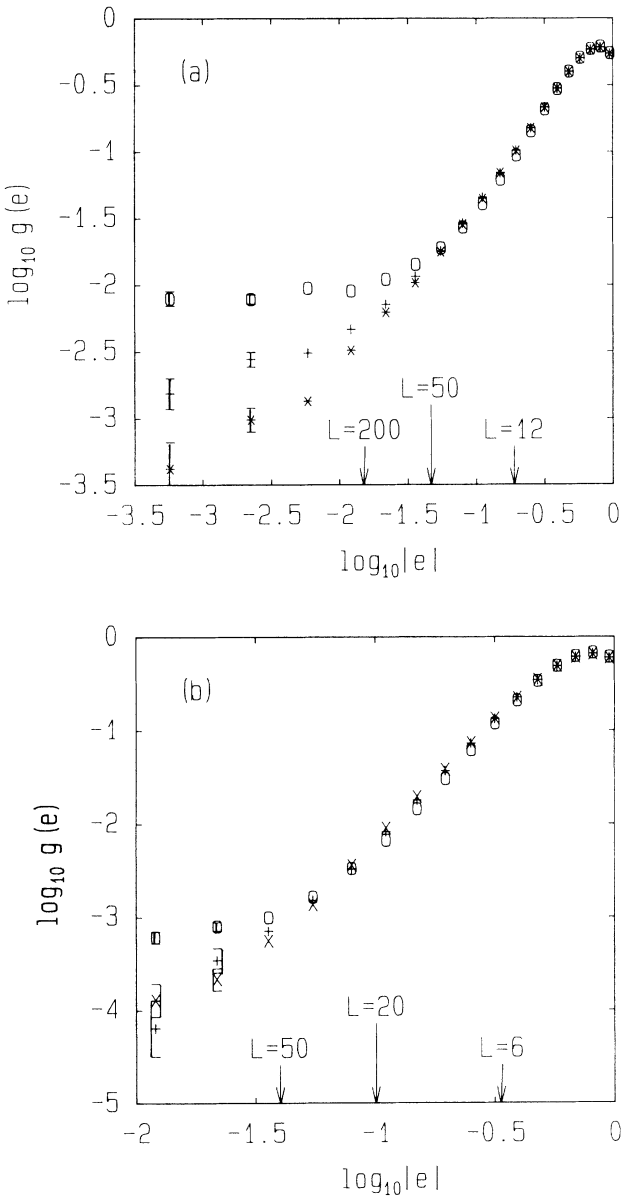

FIG. 3. $\log_{10}$-$\log_{10}$ representations of $g(e)$ for $K=0.5$, $B=1$, where the metastable states satisfy the conditions (i) and (ii). The following simulation results are compared. (a) Two-dimensional systems: $L=12$, $N_s=28\,000$ (○), $L=50$, $N_s=1600$ (+), and $L=200$, $N_s=100$ (×). (b) Three-dimensional systems: $L=6$, $N_s=20\,000$ (○), $L=20$, $N_s=125$ (+), and $L=50$, $N_s=16$ (×). The error bars correspond to a Poisson distribution, i.e., to $g(1\pm 1/m^{1/2})$, where m denotes the number of counts for the related histogram bin. The values of the most restrictive size-dependent bounds of the region of reliability, $e_0(L)$, cf. text, are marked by arrows.

The results presented in Figs. 3(a) and 3(b) are nearly analogous to the one-dimensional case:[22] There is an energy $e_0(L)$, above which the values of $g(|e|)$ can be considered as "exact," and below which "saturation," viz., convergence towards a finite value as $e\to 0$, occurs. The bound $e_0(L)$ decreases roughly as $1/L$ in the region considered. Moreover, for the lowest L values, small deviations not changing the qualitative behavior are always present at higher $|e|$.

The "low-energy deviations" arise from three effects.

(a) The distribution of interaction energies f_{ij} is size dependent below $2/L$. Thus reliable results can only be expected for $|e|>2/L$.

(b) The distribution of the f_{ij} vanishes even below $2/(d^{1/2}L)$, so that $g(e)$ has a hard gap of this width around 0. If the results of several samples are averaged, this hard gap implies a saturation of $g(e\to 0)$ due to its fluctuating position, cf. Ref. 14.

(c) We investigate a correlation effect. Therefore, studying the influence of f_{ij} on g for values of $|e|$ below a certain e_a is only possible if the mean number of sites with $-e_a\le e_i\le e_a$ per sample is large compared with unity, cf. the analytical consideration in Ref. 7.

The most restrictive limitations are the effects (c) and (a) for the two- and three-dimensional cases, respectively. The corresponding values of e_0 are marked in Figs. 3(a) and 3(b). These values, together with the graphical comparison of $g(e)$ for different L, lead to the conclusion that the numerical results can be regarded as reliable if $\log_{10}|e|\gtrsim -1.8$ for $d=2$, $L=200$, and if $\log_{10}|e|\gtrsim -1.4$ for $d=3$, $L=50$. Only these regions are considered in the following sections.

C. Asymptotic behavior

The dependence of g on disorder and band filling is shown in $\log_{10}$-$\log_{10}$ representations in Figs. 4(a)–4(d) for $d=2$ and 3, respectively. The analytical prediction [(3) and (5)] is included as a dashed line for comparison. Figures 4(a)–4(d) allow the following conclusions.

(a) Universality with respect to disorder, as predicted for $e\to 0$ analytically, is not found, neither for $d=2$ nor for $d=3$.

(b) For $K\neq 0.5$, symmetry as $e\to 0$ seems to be present for $d=2$ but not for $d=3$.

(c) The analytical prediction overestimates $g(e)$ considerably, both for $d=2$ and $d=3$.

(d) Nearly parallel, roughly linear relations are present in the regions $\log_{10}|e|\lesssim -0.9$ and $\log_{10}|e|\lesssim -0.7$ for $d=2$ and 3, respectively, but the slope exceeds the analytical prediction, $d-1$, significantly.

The difference between $d=2$ and 3 in (b) is surprising. Note that asymmetry seems to be present for $d=1$, too.[22] However, at the present stage, we cannot exclude that the asymmetries for $d=1$ and 3 arose from numerical uncertainties in the determination of μ by solving Eq. (12).

Now, the question arises whether one can find an empirical description of the asymptotic behavior as in the one-dimensional case.[22] Two hypotheses are suggested:

logarithmic correction to the analytical prediction as generalization of the one-dimensional result (7),

$$g(e) \stackrel{?}{=} g_1 |e|^{d-1} / \ln(E_c/|e|) \tag{18}$$

with nonuniversal g_1 and E_c, or a power law with an exponent differing from $d-1$, i.e.,

$$g(e) \stackrel{?}{=} g_1 |e|^{\nu} , \tag{19}$$

where ν is universal contrary to g_1.

The "logarithmic correction hypothesis" is tested in Figs. 5(a) and 5(b), where $|e|^{d-1}/g$ is represented versus $\log_{10}|e|$. The analytical prediction is included for comparison. If our assumption were justified, linear relations should be present for $e \to 0$. This seems not to be the case in the energy region considered; see in particular the "low-disorder" data.

In order to test the "power-law hypothesis," we shift a window, which includes 12 neighboring bins, along the energy scale and perform a corresponding fit for each window position. (The data points shown in Figs. 3–5 and 7–9 are mean values of four neighboring bins.) The values of ν are represented in Figs. 6(a) and 6(b) for $d=2$ and 3, respectively. In these figures, the $\log_{10}|e|$ interval considered is smaller than in Figs. 3–5 due to the finite width of the "fit window."

According to Figs. 6(a) and 6(b), it is difficult to decide whether a power law with universal exponent holds as $e \to 0$. Provided it does, we obtain $\nu = 1.2 \pm 0.1$ in the two-dimensional case and, only for $B \geq 1$, $\nu = 2.6 \pm 0.2$ in the three-dimensional case. Figures 6(a) and 6(b) provide an additional counterargument to the "logarithmic correction hypothesis." If this asymptotic behavior were present, the adjusted exponents would tend to $d-1$ as $e \to 0$. This is not the case in the energy interval considered.

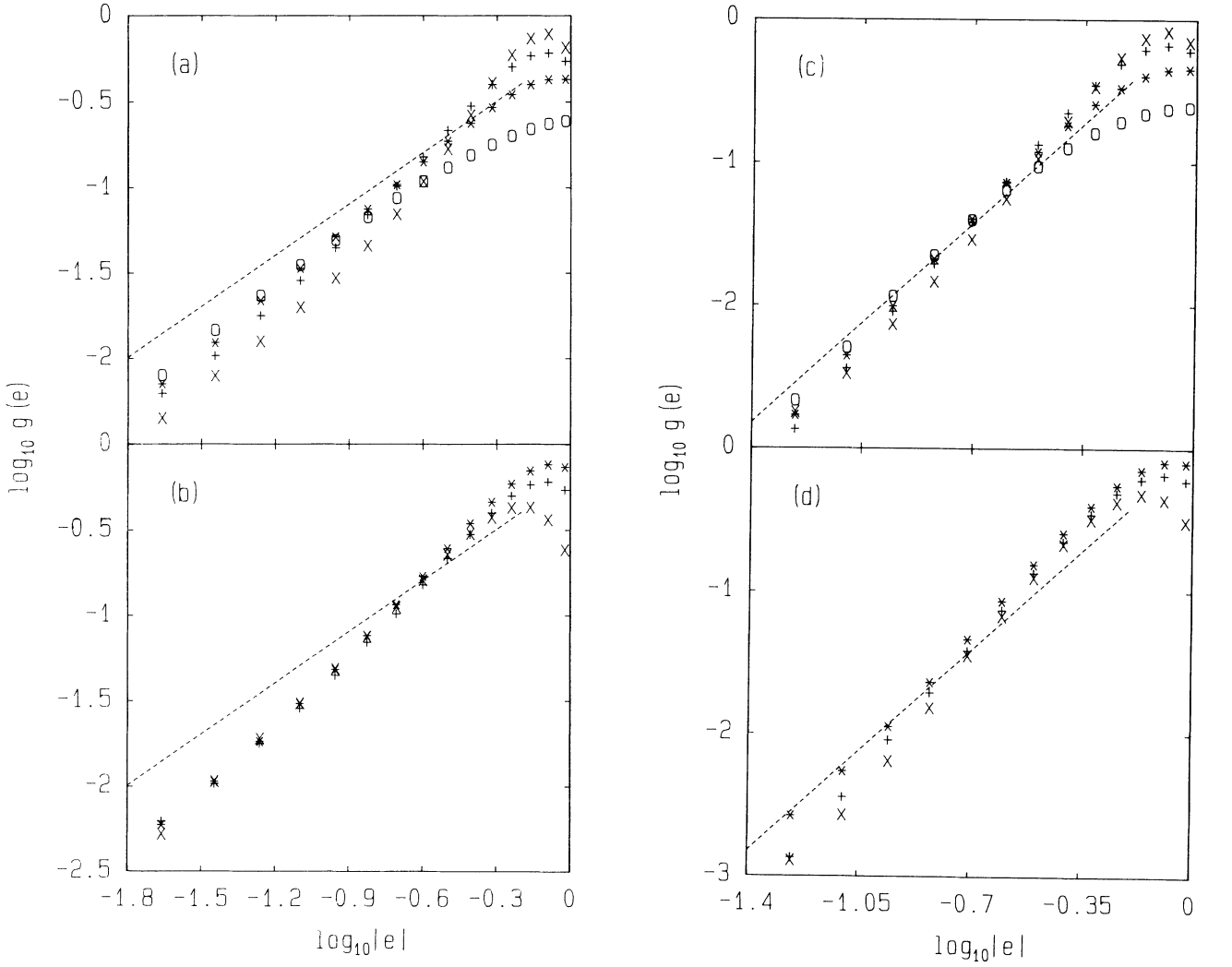

FIG. 4. Influence of the degree of disorder B and of the filling factor K on $g(e)$ as $e \to 0$ in $\log_{10}$-$\log_{10}$ representations. The data, presented in Fig. 2, are redrawn: (a) and (b) $d=2$ and (c) and (d) $d=3$. The symbols are the same as in Figs. 2(a)–2(d) except for $B=1$, $K=0.3$, for which $\times$ and $*$ denote lower and upper branches, i.e., the regions $e<0$ and $e>0$, respectively. The dashed lines give the analytical prediction [(3) and (5)] of the asymptotic behavior.

D. Stability with respect to more complex excitations

The construction of states, being stable not only with respect to single-particle hops [conditions (i) and (ii) of Sec. II], but also concerning two-particle hops [conditions (iii) and (iv)], is a fairly difficult task. The consideration of a reasonable system size seems to be possible only by means of branch-and-bound search algorithms. In the one-dimensional case, we found that additional stability with respect to higher-order excitations leads to a decrease of the density of states. But it does not destroy the logarithmic behavior, only prefactor and characteristic energy are shifted.

Figures 7(a) and 7(b) give $\log_{10}$-$\log_{10}$ representations of $g(e)$ vs $|e|$ for different degrees of metastability for $d=2$ and 3, respectively. As in the case $d=1$, $g(e)$ decreases with increasing degree of metastability monotonously. However, the qualitative behavior is not changed: A transition towards exponential $g(e)$ in the energy region considered can be excluded for $d=2$, and it seems unlikely for $d=3$. [If $g(e)$ would follow an exponential law as Eq. (6), negative $d^2\log_{10}g / d(\log_{10}|e|)^2$ were expected.]

It is interesting to study the influence of the initial conditions in this context, see Fig. 8 for $d=2$, $B=0.5$, $K=0.5$. A regular (NaCl-like) occupation is frozen in for this parameter set when only 1- and 2-site excitations are considered; there is a hard gap of width 1.115 in this case. States of lower total energy can only be reached if

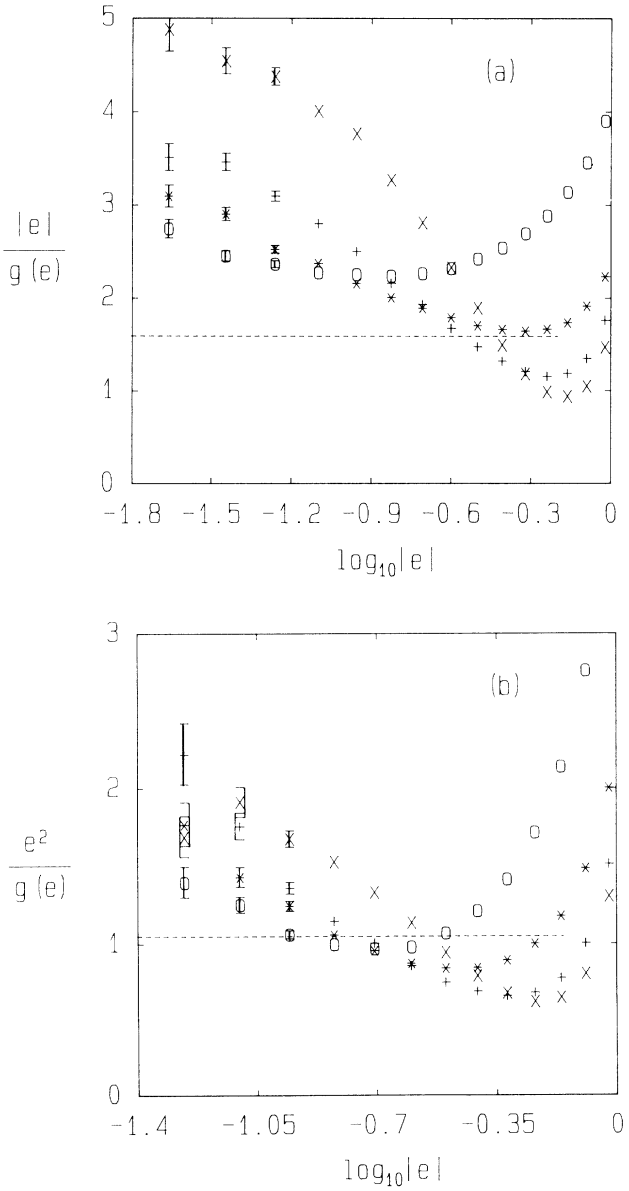

FIG. 5. Test with respect to hypothesis (18): influence of the degree of disorder B on $g(e)$ as $e\to 0$ in $|e|^{d-1}/g(e)$ vs $\log_{10}|e|$ representations, (a) $d=2$ and (b) $d=3$. For the meaning of the symbols see the captions of Figs. 2 and 4.

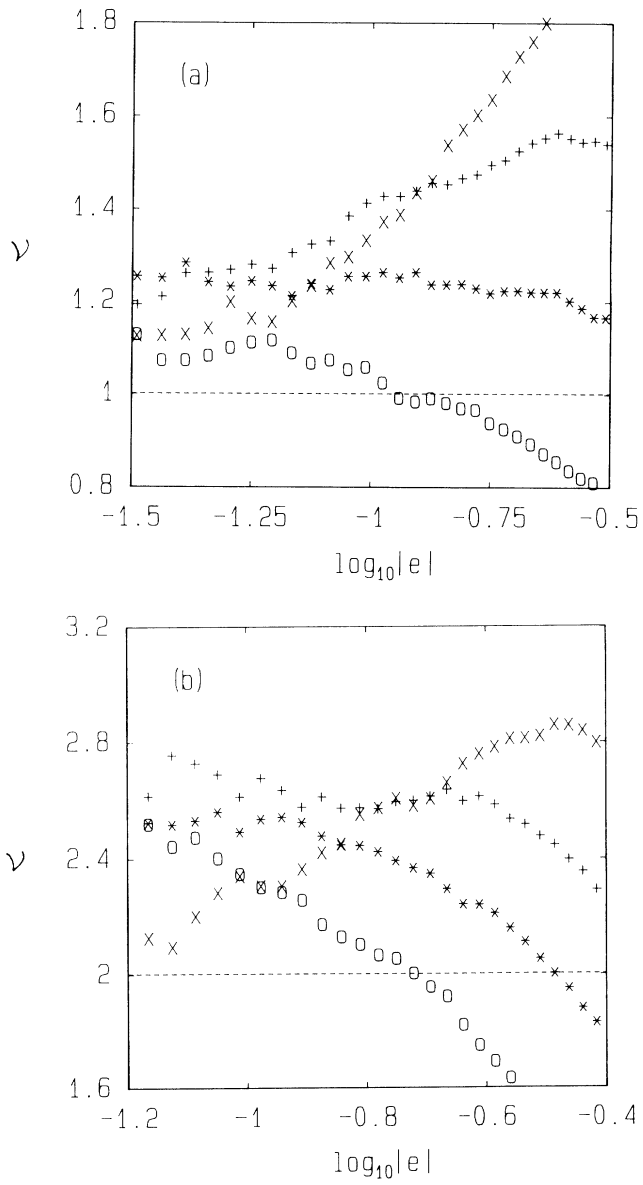

FIG. 6. Asymptotic exponent ν as determined by fitting $g(e)$ to Eq. (19), (a) $d=2$ and (b) $d=3$. Different degrees of disorder are compared. For the meaning of the symbols see the captions of Figs. 2 and 4.

more complex excitations are taken into account.

Figure 8 shows large differences between the results for regular and stochastic initial conditions. However, convergence towards a common limit, as the degree of metastability increases, is present. For stability with respect to (i)–(iv), the $g(e)$ values for the alternative initial occupations seem to differ only by an e independent factor at $\log_{10}|e| \lesssim -0.6$ Moreover, it is instructive to compare the results obtained for regular initial conditions and stability with respect to (i)–(iv) on the one hand and for stochastic initial conditions and stability concerning (i), (ii) on the other hand. Although the $g(e)$ curves differ considerably, the total energy has nearly the same value in both cases. Thus, the states constructed cannot be characterized only by its energy values.

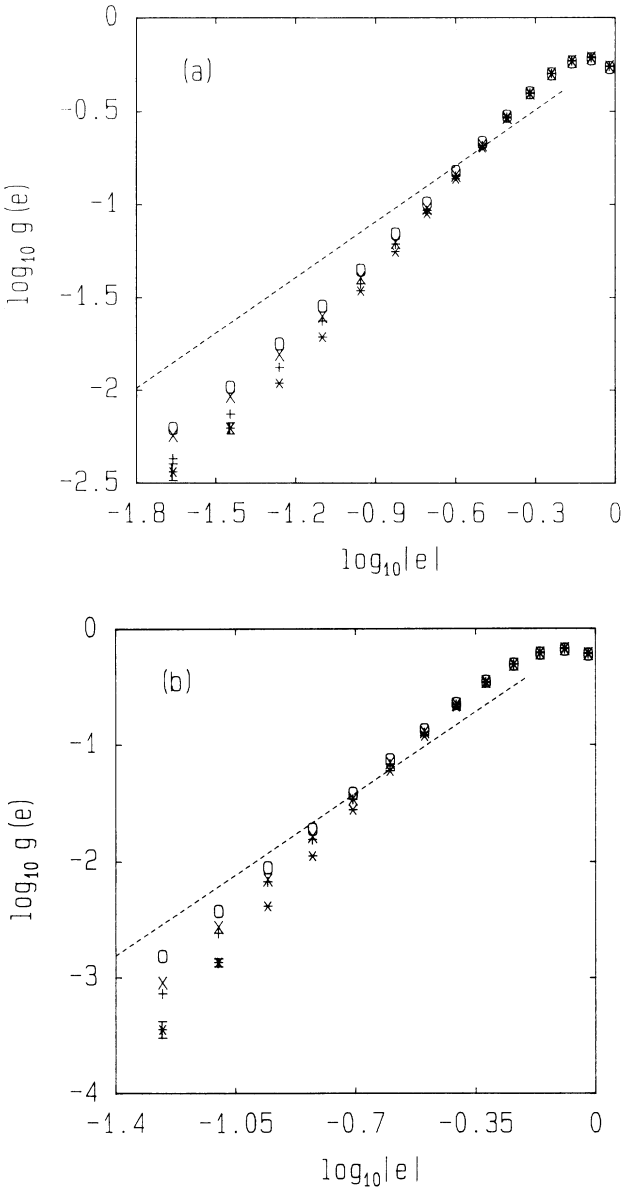

FIG. 7. Comparison of different degrees of metastability for $B=1$, $K=0.5$: (a) $d=2$ and (b) $d=3$. The meaning of the symbols is as follows: ○, stability with respect to (i) and (ii); ×, relaxation of the ○ via nearest neighbor 3- and 4-site rearrangements; +, relaxation of the × including all 3-site rearrangements; and *, full stability with respect to (i)–(iv). In the two-dimensional case, simulations were performed for $L=200$, $N_s=100$, except for + and *, where $N_s=25$. In the three-dimensional case, 16 samples with $L=50$ were simulated, except for ○ where $N_s=32$.

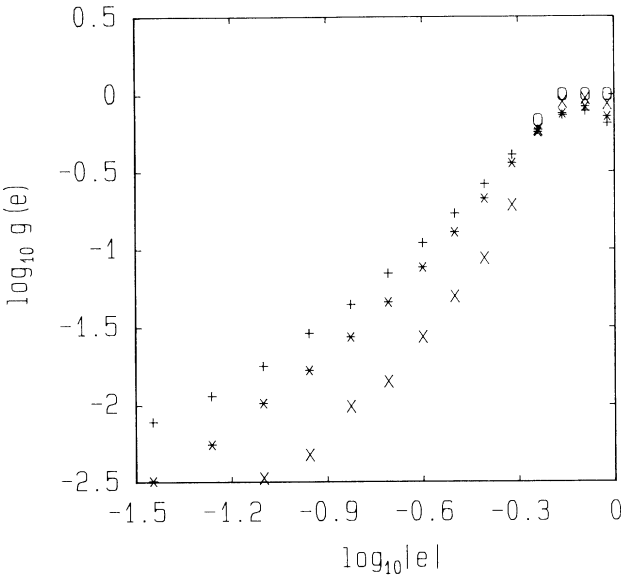

FIG. 8. Influence of initial conditions for $d=2$, $B=0.5$, $K=0.5$. + and *, stochastic initial occupation; ○ and ×, regular initial occupation. + and ○ are stable with respect to (i) and (ii), whereas × and * fulfil conditions (i)—(iv). In the case of ○, $g(e)=0$ for $|e| \leq 0.558$. For all parameter sets, we use $L=100$, $N_s=100$. The values of the mean total energy per site, $\langle h/L^d \rangle$, are as follows: (+), −0.2058, (○) −0.2020, (*) −0.2094, and (×) −0.2059.

E. Modification of the short-range part of the interaction

The Coulomb gap should survive when the states at neighboring sites begin to overlap, since the asymptotic properties are determined by the long-range interaction behavior. The same hypothesis has been derived from the observation of a universal temperature dependence of the conductivity near the metal-semiconductor transition in disordered systems.[8] However, DLR observe a clustering of "low-energy sites." Thus a stronger influence of the overlapping on $g(e)$ than in the case of a homogeneous spatial distribution is imaginable.

In order to get a qualitative impression what happens when the localized states begin to overlap, we replace now the point charges by exponentially smeared-out charge distributions. That means, the charge density $\rho(r)$ is given by

$$\rho(r)=\frac{1}{8\pi R^3}\exp\left[-\frac{r}{R}\right] . \tag{20}$$

Provided $R \ll L$, some lengthy integration[30] yields the following expression for the interaction energy of two

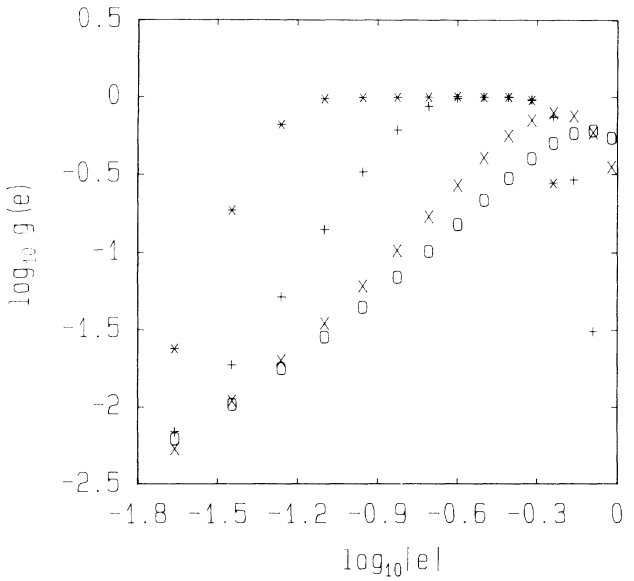

FIG. 9. Coulomb gap in the case of exponentially smeared charges for $d=2$, $B=1$, $K=0.5$: ($\circ$) $R=0$, ($\times$) $R=0.3$, ($+$) $R=1$, and ($*$) $R=3$. A hundred samples with $L=200$ are averaged in each case.

smeared-out charges localized at sites i and j with distance r_{ij}:

$$f_{ij}=\frac{1}{r_{ij}}[1-(\tfrac{1}{48}\zeta^3+\tfrac{3}{16}\zeta^2+\tfrac{11}{16}\zeta+1)e^{-\zeta}] , \tag{21}$$

where $\zeta=r_{ij}/R$, and where r_{ij} is determined by means of Eq. (10). Of course, only a part of the influence of finite-range charge distributions is incorporated in this way; hybridization and polarization are neglected.

Figure 9 shows the results of corresponding simulations for $d=2$, $R=0-3$. The following conclusions can be drawn from this figure.

(a) The Coulomb gap persists if the charges are smeared out.

(b) The "width" of the gap decreases as R increases. For not too small R, the width is roughly proportional to $1/R$.

(c) At sufficiently low $|e|$, $g(e;R)$ might be universal with respect to R, cf. the results for $R=0$, and 0.3.

Therefore, the numerical simulations confirm the hypothesis that the Coulomb gap persists when the states of neighboring sites begin to overlap.

V. CONCLUSIONS

The results of the numerical investigations presented in Sec. IV can be summarized in the following five points.

(a) The numerical $g(e)$ data are considerably smaller than the analytical results, both for $d=2$ and 3.

(b) Universality with respect to disorder, as predicted analytically, is not present.

(c) It cannot be decided finally, whether or not $g(e\rightarrow 0)$ follows a power law. Provided it does, the numerical simulations yield the exponents 1.2 ± 0.1 and 2.6 ± 0.2 for $d=2$ and 3, respectively, in contradiction to the analytical result $d-1$.

(d) Additional relaxation down to stability with respect to 3- and 4-site excitations leads to a decrease of the single-particle density of states, but the qualitative behavior is not changed. Exponential behavior can be excluded for $d=2$, and it seems unlikely for $d=3$.

(e) The Coulomb gap persists if the point charges are smeared out exponentially. For not too small R, its "width" is roughly indirectly proportional to the localization radius R.

These findings involve several consequences. First, the basic assumptions of the analytical theory, in particular of the self-consistent equation method,[12] should be reanalyzed critically. This method considers the neighborhood of a certain site in a mean-field-like manner, which corresponds to assuming the sites with small $|e|$ to be homogeneously distributed. However, it was numerically shown by DLR (Ref. 15) that low-energy sites cluster. Thus it is likely that at least the pair correlation has explicitly to be taken into account in a better analytical theory.

Second, the single-particle density of states has an essential influence on the conductivity in the variable-range-hopping regime, in particular on its temperature dependence. Exponent optimization, as the simplest approximation, leads to $\sigma(T)\sim\exp[-(T_0/T)^{\alpha}]$, where $\alpha=(\nu+1)/(\nu+d+1)$. In this manner, Efros and Shklovskii[7] obtained $\alpha=\frac{1}{2}$ for both $d=2$ and 3, instead of the Mott result for a flat density of states, $\alpha=\frac{1}{3}$ and $\frac{1}{4}$ for $d=2$ and 3, respectively. However, these results are based on a series of assumptions. For a more elaborate analytical theory see Ref. 31 and for related computer simulations Ref. 32. The values obtained for ν above lead to $\alpha=0.52\pm0.01$ and 0.55 ± 0.01 for $d=2$ and 3, respectively. In the three-dimensional case, it should be possible to distinguish between $\frac{1}{2}$ and 0.55 in precision experiments.

ACKNOWLEDGMENTS

We are deeply indebted to the Forschungszentrum Jülich GmbH, in particular to G. Eilenberger, for generous help: Two of us (A.M. and M.R.) would like to thank this Institute for financial support, and A.M. appreciates the hospitality during a two week stay and the opportunity to use its computer facilities. Moreover, we are indebted to F. Bialas and F. Goedsche for collaboration in early stages of this investigation. Numerous discussions with A. L. Efros, B. I. Shklovskii, M. Pollak, H. Eschrig, and the late R. Richter are gratefully acknowledged.

*Present address: SAP A.G., Max-Planck-Strasse 8, D-W-6909 Walldorf, Germany.

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

[2]A. L. Efros and B. I. Shklovskii, in *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985), p. 409.

[3]M. Pollak and M. Ortuño, in Ref. 2, p. 287.

[4]O. Entin-Wohlmann, Y. Gefen, and Y. Shapira, J. Phys. C **16**, 1161 (1983); C. J. Adkins, J. Phys. Condensed Matter **1**, 1253 (1989).

[5]E. L. Wolf, R. H. Wallis, and C. J. Adkins, Phys. Rev. B **12**, 1603 (1975); W. L. McMillan and J. Mochel, Phys. Rev. Lett. **46**, 556 (1981); G. Hertel, D. J. Bishop, E. G. Spencer, J. M. Rowell, and R. C. Dynes, *ibid.* **50**, 743 (1983); A. E. White, R. C. Dynes, and J. P. Garno, *ibid.* **56**, 532 (1986); S. Schmitz, thesis, Reinisch-Westfälische Technische Hochschule Aachen, 1989.

[6]J. H. Davies and J. R. Franz, Phys. Rev. Lett. **57**, 475 (1986); G. Hollinger, P. Pertosa, J. P. Doumerc, F. J. Himpsel, and B. Reihl, Phys. Rev. B **32**, 1987 (1985).

[7]A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).

[8]"$\sigma \sim \exp[-(T_0/T)^{1/2}]$" behavior has been observed in many experiments; for a survey up to 1984, see A. Möbius, J. Phys. C **18**, 4639 (1985). For more recent related investigations cf. e.g., W. Schoepe, Z. Phys. B **71**,455 (1988); A. Audouard, A. Kazoun, N. Cherradi, G. Marchal, and M. Gerl, Philos. Mag. B **59**, 207 (1989).

[9]M. Pollak, Discuss. Faraday Soc. **50**, 13 (1970); G. Srinivasan, Phys. Rev. B **4**, 2581 (1971).

[10]T. Kurosowa and H. Sugimoto, Prog. Theor. Phys. Suppl. **57**, 217 (1975).

[11]S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978); G. Parisi, Phys. Rev. Lett. **50**, 1946 (1983).

[12]A. L. Efros, J. Phys. C **9**, 2021 (1976).

[13]S. D. Baranovskii, B. I. Shklovskii, and A. L. Efros, Zh. Eksp. Teor. Fiz. **78**, 395 (1980) [Sov. Phys.—JETP **51**, 199 (1980)].

[14]S. D. Baranovskii, A. L. Efros, B. L. Gelmont, and B. I. Shklovskii, J. Phys. C **12**, 1023 (1979).

[15]J. H. Davies, P. A. Lee, and T. M. Rice, Phys. Rev. Lett. **49**, 758 (1982); Phys. Rev. B **29**, 4260 (1984).

[16]B. D. Hadley, M. Pollak, and M. Ortuño, Phys. Rev. B **37**, 9006 (1988).

[17]R. Chicon, M. Ortuño, and M. Pollak, Phys. Rev. B **37**, 10 520 (1988).

[18]A. L. Efros, Nguyen Van Lien, and B. I. Shklovskii, J. Phys. C **12**, 1869 (1979).

[19]M. Grünewald, B. Pohlmann, L. Schweitzer, and D. Würtz, J. Phys. C **15**, L1153 (1982).

[20]S. Summerfield, J. A. McInnes, and P. N. Butcher, J. Phys. C **20**, 3647 (1987).

[21]A. Möbius, F. Goedsche, M. Richter, and F. Bialas, *Proceedings of the 15th Annual International Symposium on Electronic Structure of Metals and Alloys, Johnsbach, GDR, 1985,* edited by P. Ziesche (Technische Universität, Dresden, 1985), p. 198; A. Möbius and M. Richter, *Proceedings of the International Seminar on Localization in Disordered Systems,* Bad Schandau-Ostrau, 1986, edited by W. Weller and P. Ziesche (Teubner, Leipzig, 1988), p. 167; A. Möbius and M. Richter, J. Non-Cryst. Solids **97&98**, 483 (1987).

[22]A. Möbius and M. Richter, J. Phys. C **20**, 539 (1987).

[23]M. E. Raich and A. L. Efros, Pisma Zh. Eksp. Teor. Fiz. **45**, 225 (1987) [Sov. Phys.—JETP Lett. **45**, 280 (1987)].

[24]M. M. Syslo, N. Deo, and J. S. Kowalik, *Discrete Optimization Algorithms* (Prentice-Hall, Englewood Cliffs, NJ, 1983).

[25]S. Kobe and A. Hartwig, Comput. Phys. Commun. **16**, 1 (1978).

[26]A. Möbius and M. Richter, in *Evolution and Optimization '89. Selected Papers on Evolution Theory, Combinatorial Optimization, and Related Topics,* edited by H.-M. Voigt, H. Mühlenbein, and H.-P. Schwefel (Akademie-Verlag, Berlin, 1990), p. 199.

[27]M. R. Garey and D. S. Johnson, *Computers and Interactibility: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).

[28]N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[29]S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Science **220**, 671 (1983).

[30]H. Eschrig, Phys. Status Solidi B **96**, 329 (1979).

[31]F. Goedsche, Phys. Status Solidi (b) **140**, 225 (1987).

[32]E. I. Levin, V. L. Nguen, V. L. Shklovskii, and A. L. Efros, Zh. Eksp. Teor. Fiz. **92**, 1499 (1987) [Sov. Phys.—JETP **65**, 842 (1987)].