Coulomb gap in a model with finite charge-transfer energy

S. A. Basylko,[1] P. J. Kundrotas,[2,3] V. A. Onischouk,[1,2] E. E. Tornau,[2,4] and A. Rosengren[2]
[1]*Joint Institute of Chemical Physics of Russian Academy of Sciences, 117 977 Kosygin Street 4, Moscow, Russia*
[2]*Department of Physics/Theoretical Physics, Royal Institute of Technology, SE–100 44 Stockholm, Sweden*
[3]*Faculty of Physics, Vilnius University, Sauletekio al. 9, LT-2040, Vilnius, Lithuania*
[4]*Semiconductor Physics Institute, Goštauto 11, LT-2600 Vilnius, Lithuania*
(Received 17 March 2000; revised manuscript received 29 June 2000; published 12 December 2000)

The Coulomb gap in a donor-acceptor model with finite charge-transfer energy Δ describing the electronic system on the dielectric side of the metal-insulator transition is investigated by means of computer simulations on two- and three-dimensional finite samples with a random distribution of equal amounts of donor and acceptor sites. Rigorous relations reflecting the symmetry of the model presented with respect to the exchange of donors and acceptors are derived. In the immediate neighborhood of the Fermi energy μ, the single-particle density of states $g(\varepsilon)$ is determined solely by finite size effects, and $g(\varepsilon)$ further away from μ is described by an asymmetric power law with a nonuniversal exponent, depending on the parameter Δ.

DOI: 10.1103/PhysRevB.63.024201 PACS number(s): 71.23.−k, 71.30.+h, 71.45.Gm

I. INTRODUCTION

Doping of solids might lead to drastic qualitative changes in their properties. The metal-insulator transition (MIT) is a spectacular manifestation of this. The understanding of the driving forces of the MIT is a long-standing problem. In the early seventies, the prediction[1] was made that on the dielectric side of the MIT the long-range Coulomb interactions deplete the single-particle density of states (SPDOS) $g(\varepsilon)$ near the Fermi energy μ. Further, analytical calculations of $g(\varepsilon)$ with Coulomb correlation taken into consideration have been performed on the metallic side of the MIT. Altshuler and Aronov[2] showed that for the metallic case $g(\varepsilon)$ in three dimensions has a cusplike dependence $g(\varepsilon)\sim|\varepsilon-\mu|^{1/2}$ near μ. This was later confirmed in electron tunneling experiments for amorphous alloys[3] and granular metals.[4]

On the insulating side of the MIT, charge transport occurs via inelastic electron tunneling hopping between states localized on the impurity sites with one-electron energies close to μ. Mott[5] demonstrated that at low temperatures electrons seek accessible energy states by hopping distances beyond the localization length, leading to a temperature (T) dependence of the hopping conductivity $\sigma(T)\sim\exp(-T_0/T)^{\nu}$ with T_0 being a characteristic T depending on localization length and with the hopping exponent $\nu=1/4$ for the noninteracting case in three dimensions. Efros and Shklovskii[6] (ES) argued that in the ground state of a system with long-range Coulomb interactions, $g(\varepsilon)$ (when $\varepsilon\to\mu$) has the symmetric shape

$$g(\varepsilon)\sim|\varepsilon-\mu|^{D-1} \tag{1}$$

with the *universal* exponent $D-1$ (D is the dimensionality of the system). Because $g(\varepsilon)$ vanishes only at $\varepsilon=\mu$, this is called a ''soft'' Coulomb correlation gap with a width[6] $\Delta\varepsilon\sim e^3(N_0/\chi^3)^{1/2}$ (N_0 is the SPDOS far away from μ, χ the dielectric constant and e the electron charge). Equation (1) gives[7] for $D=3$ a hopping exponent $\nu=1/2$ at low T and, indeed, the transition $(\nu=1/4)\to(\nu=1/2)$ with lowering of temperature was observed in experiments.[8]

The intriguing hypothesis about the universality of the exponent in Eq. (1) [henceforth referred to as the universality hypothesis (UH)] has stimulated further theoretical research, both analytical[9] and numerical.[10–15] To establish the UH, Efros[16] used the ground-state stability conditions for localized electrons with respect to charge transfer

$$\varepsilon_j-\varepsilon_i-\frac{e^2}{\chi r_{ij}}>0, \tag{2}$$

where ε_i and ε_j are the single-particle energies (SPE) of a neutral donor (d^0) on a site i and of a charged donor (d^+) on a site j, respectively, and r_{ij} is the distance between the sites i and j. The conditions (2) were used to heuristically derive a nonlinear integral equation for $g(\varepsilon)$,[16–19] and then asymptotic analysis of this equation leads[16] to Eq. (1).

Localized electrons have been studied using the so-called classical donor-acceptor (d-a) model (see, e.g., Ref. 17). Within this model, the system considered is modeled by a continuous sample with randomly distributed $k\times N$ ($k\leqslant 1$) acceptor (a) and N donor (d) sites. Each a site is negatively charged whereas out of N only $k\times N$ donors are d^+ which leads to a large number of configurations, each of which must obey not only conditions (2) but also more complicated conditions related to many-particles excitations (e.g., charge transfer involving four, six, etc. sites). The conjecture of Efros about the *universality* implies that $g(\varepsilon)$ does not depend on peculiarities of the particular model and, as a consequence, further theoretical studies of localized electrons[10,12–14] were confined to a *lattice* d-a model proposed in Ref. 16. In this model, N donors are localized on all the sites of a D-dimensional lattice and the negative charge of $k\times N$ acceptors is uniformly smeared over the lattice sites so that each site i has a charge $e(n_i-k)$, where $n_i=1(0)$ for d^+ (d^0). Disorder in this model is ensured by introducing randomly distributed on-site potentials. Monte Carlo (MC) simulations[13] on very large specimens of the lattice d-a model at $T=0$, however, have given rise to doubts about the universality of the $g(\varepsilon)$ behavior. Further, studies[20] of the Coulomb gap at $T\neq 0$ have revealed significant deviations of the exponent in Eq. (1) from the predicted universal value.

Another hint about possible nonuniversal behavior of $g(\varepsilon)$ has come from the intriguing and still not completely

decided problem of whether the so-called spin-glass phase does exist in the classical d-a model (see, e.g., Refs. 21–23). Grannan and Yu[21] studied the classical $D=3$ d-a model with $k=0.5$ but with the total acceptor charge uniformly distributed over d sites as in the lattice d-a model. In this case, the classical d-a model is equivalent to a model of Ising spins, localized on randomly distributed sites, with pairwise Coulomb interactions, a model in which a transition into the spin-glass state was found[21] to occur at $T\neq 0$. It was then concluded that such a transition should exist in *all* d-a models (with or without smearing of negative charge, defined on a lattice or on a continuous sample) as well because of the UH. Vojta and Schreiber,[23] however, have shown that the spin glass transition does not exist in the lattice d-a model.[16] Besides, in recent work by one of us,[22] it was unequivocally demonstrated that the ground state of the classical d-a model and that of the model studied in Ref. 21 are qualitatively different. An analysis of histograms $\mathcal{H}[Q_{\alpha\beta}]$ for overlaps $Q_{\alpha\beta}=(1/N)\Sigma_i \delta(n_i^{\alpha},n_i^{\beta})$ (here α and β refer to different lowest-energy states) has revealed that, indeed, for the model studied in Ref. 21, $\mathcal{H}[Q_{\alpha\beta}]$ has a maximum at $Q_{\alpha\beta}=0$, i.e., a large number of microscopically different lowest-energy states does exist in the model and according to Parisi's theory[24] this implies the existence of a spin-glass state at low temperatures. Further MC simulations[22] at $T\neq 0$ revealed the typical finite-size scaling of the spin-glass susceptibility. In the classical d-a model, however, $\mathcal{H}[Q_{\alpha\beta}]$ has its maximum at $Q_{\alpha\beta}=1$ which means that all lowest-energy states are the same from a microscopical point of view.

Therefore, it is highly desirable to study the properties of not only the classical d-a model, but of its various modifications as well. In the present work we consider a modified classical d-a model (MCDAM) in which acceptors can be neutral, so the energy Δ of the charge transfer from a donor to an acceptor, $d^0+a^0\rightarrow d^+ + a^-$, has to be finite. The classical d-a model might then be viewed as the limit of the MCDAM as $\Delta\rightarrow\infty$. We have investigated the shape of the Coulomb gap in two- and three-dimensional MCDAMs at $T=0$ and found that the behavior of $g(\varepsilon)$ is in strong contradiction to the UH of Efros. The paper is organized as follows. In Sec. II we introduce the MCDAM and arrive at some rigorous results which follow from a symmetry of the MCDAM with respect to the exchange of donor and acceptor sites. Further, the algorithm of energy minimization for the MCDAM, including a discussion about inherent finite size effects, is presented in Sec. III. Section IV is devoted to a description of the main results obtained. In Sec. V we discuss possible causes of universality violation in the MCDAM, analyze experimental data available in the literature, and predict possible experimental situations in which the nonuniversal behavior of $g(\varepsilon)$ might be observed. Finally, a summary is presented in Sec. VI.

II. BACKGROUND

A. Model

We consider a D-dimensional system of volume L^D, in which an equal number N of a and d sites are allocated according to the Poisson distribution with a density $n=N\times L^{-D}$. Hereafter, all expressions will be written in dimensionless units, $n^{-1/D}$ for length, and $E_0=e^2n^{1/D}/\chi$ for energy. A microscopic state of a particular spatial arrangement of the d and a sites (henceforth referred to as the sample **R**) is determined by a set of occupation numbers $(n_a,n_d)\equiv\{n_a(i),n_d(k),\ i,k=1,2,\ldots,N\}$ with $n_a(i)=1(0)$ for $a^-(a^0)$ and $n_d(k)=1(0)$ for $d^0(d^+)$. We investigate the case of strictly localized electrons, when $\alpha_B \ll 1$ (α_B is the localization length of the electron on donor) and the energy of the sample then is

$$E(n_a,n_d)=\frac{1}{2}\sum_{i\neq j}\frac{n_a(i)n_a(j)}{r_{ij}^{a-a}}+\frac{1}{2}\sum_{k\neq l}\frac{(1-n_d(k))(1-n_d(l))}{r_{kl}^{d-d}}-\sum_{i,k}\frac{(1-n_d(k))n_a(i)}{r_{ik}^{a-d}}-\Delta\sum_i n_a(i), \quad (3)$$

where indices i,j and k,l number a and d sites, respectively, r_{ij}^{a-a}, r_{kl}^{d-d}, and r_{ik}^{a-d} are the distances between the acceptors, between the donors, and between the acceptor and the donor, correspondingly. When charge transfer occurs, $E(n_a,n_d)$ changes by

$$\delta E(n_a,n_d)=\sum_i \varepsilon_a(i)\,\delta n_a(i)+\sum_k \varepsilon_d(k)\,\delta n_d(k)+\sum_{i,k}\frac{\delta n_a(i)\,\delta n_d(k)}{r_{ik}^{a-d}}+\frac{1}{2}\sum_{i\neq j}\frac{\delta n_a(i)\,\delta n_a(j)}{r_{ij}^{a-a}}+\frac{1}{2}\sum_{k\neq l}\frac{\delta n_d(k)\,\delta n_d(l)}{r_{kl}^{d-d}}, \quad (4)$$

where $\varepsilon_a(i)$ is the SPE for the acceptors

$$\varepsilon_a(i)\equiv\frac{\delta E(n_a,n_d)}{\delta n_a(i)}=\sum_{j\neq i}\frac{n_a(j)}{r_{ij}^{a-a}}-\sum_k\frac{1-n_d(k)}{r_{ik}^{a-d}}-\Delta, \quad (5)$$

$\varepsilon_d(i)$ is the corresponding SPE for the donors and $\delta n_a(i)[\delta n_d(k)]$ denotes the change of the occupation number on the a (d) site. If a microscopic state (n_a^0,n_d^0) is the ground state of this sample, then for any excitation the relation

$$\delta E(n_a^0,n_d^0)\geqslant 0 \quad (6)$$

holds. The specific appearance of the conditions (6) depends on what excitations are allowed in the model system considered.

We investigate the simplest case with only pairs of sites involved in the charge transfer which can occur in four different ways, each of them denoted by a unique set of $\{\delta n_a(i),\delta n_a(j),\delta n_d(k),\delta n_d(l)\}$: (i) via electron hops between acceptors $\{-1,1,0,0\}$; (ii) via electron hops between donors $\{0,0,-1,1\}$; (iii) via ionization $\{1,0,-1,0\}$ and (iv)

via recombination $\{-1,0,1,0\}$. From Eq. (4) one obtains the ground-state stability relation for (i):

$$\varepsilon_a^0(j)-\varepsilon_a^1(i)-\frac{1}{r_{ij}^{a-a}}\geqslant 0, \tag{7}$$

where $\varepsilon_a^{1(0)}(i)$ denotes $\varepsilon_a(i)$ if $n(i)=1(0)$. The stability conditions for (ii)–(iv) are obtainable in a similar manner.

The pair $(a^0,a^-)[(d^0,d^+)]$ might be located on any distance and therefore in the thermodynamic limit μ for the acceptors [donors] (i.e., an energy level which separates $\varepsilon_{a[d]}^0$ and $\varepsilon_{a[d]}^1$) is determined as

$$\mu_{a[d]}=\min\{\varepsilon_{a[d]}^0(i)\}=\max\{\varepsilon_{a[d]}^1(i)\}, \tag{8}$$

and the stability relations with respect to the ionization and recombination lead to

$$\mu_a=\mu_d=\mu. \tag{9}$$

Despite the finite size of samples we investigated, the relation (9) with $\mu_{a[d]}$ calculated from Eq. (8) is valid within the limits of accuracy of our calculations.

A macroscopic state of the sample **R** is characterized by the degree of acceptor ionization

$$C_a(\mathbf{R})=\frac{1}{N}\sum_i n_a(i), \tag{10}$$

by the SPDOS for acceptors

$$g_a(\varepsilon_a,\mathbf{R})=\frac{1}{N}\sum_i \delta(\varepsilon-\varepsilon_a(i)), \tag{11}$$

and by the corresponding SPDOS $g_d(\varepsilon_d,\mathbf{R})$ for the donors. Note, that for the finite samples (especially for the relative small systems we were able to investigate) $C_a(\mathbf{R}),g_a(\varepsilon_a,\mathbf{R})$, and $g_d(\varepsilon_d,\mathbf{R})$ depend to a large extent on **R** (if a sample would be big enough all quantities would be self-averaging). Therefore, in order to obtain reliable results, one has to work with the quantities $C_a\equiv\langle C_a(\mathbf{R})\rangle, g_a(\varepsilon)\equiv\langle g_a(\varepsilon_a,\mathbf{R})\rangle$ and $g_d(\varepsilon)\equiv\langle g_d(\varepsilon_d,\mathbf{R})\rangle$, where $\langle\ldots\rangle$ denotes the average over a number of **R**'s. Note, that the values $g_{a(d)}(\varepsilon_{a(d)},\mathbf{R})d\varepsilon$ obtained for independent **R**'s are scattered according to the Gaussian distribution with mean $g_{a(d)}(\varepsilon)d\varepsilon$ and standard deviation $\sqrt{g_{a(d)}(\varepsilon)d\varepsilon}$. In the region of the Coulomb gap, $g_{a(d)}(\varepsilon)d\varepsilon\sim 10^{-4}$ and the dispersion is several orders of magnitude larger than the mean. Therefore, in order to reduce the statistical noise in the final $g_{a(d)}(\varepsilon)$ dependences, an average is needed over a sufficient large amount of independent samples (we performed calculations with up to 10^4 samples).

B. Acceptor-donor symmetry

The system investigated is electrically neutral, i.e., for any sample

$$\sum_i n_a(i)=\sum_k (1-n_d(k)). \tag{12}$$

Then, the energies of any pair of microscopic states (n_a,n_d) and (n_a^*,n_d^*) for a sample **R** and its ''mirror'' reflection **R*** (when the donor and acceptor sites exchange places keeping the spatial arrangement of sites unchanged) are equal under the following conditions:

$$\varepsilon_a(i)+\varepsilon_d^*(i)=\varepsilon_d(k)+\varepsilon_a^*(k)=-\Delta \tag{13}$$

and

$$n_a^*(i)=[1-n_d(i)]\quad n_d^*(k)=[1-n_a(k)]. \tag{14}$$

The stability relations (7) for the ground state (n_a^0,n_d^0) of the sample **R** transform into stability relations for the ground state (n_a^{0*},n_d^{0*}) of the sample **R*** through the relations (13), (14) as well.

Since averaging over samples includes all possible pairs **R** and **R***, it follows from the symmetry relations (13) and (14) along with the definition (11) that $g_a(\varepsilon)$ can be mapped to $g_d(\varepsilon)$ using the relation

$$g_d(\varepsilon)=g_a(-\varepsilon-\Delta). \tag{15}$$

The symmetry of the model imposes also a relation between the Fermi energy μ [Eqs. (8), (9)] and the parameter Δ of the model. Expressing $n_{a[d]}(i[k])$ in terms of the Heaviside's step functions $n_{a[d]}=\theta(\mu-\varepsilon_{a[d]}(i[k]))$, the quantity C_a can be written in the form

$$C_a=\int_{-\infty}^{\mu} g_a(\varepsilon)d\varepsilon=\int_{\mu}^{\infty} g_d(\varepsilon)d\varepsilon. \tag{16}$$

The symmetry relation (15) transforms Eq. (16) into an integral relation

$$\int_{-\mu-\Delta}^{\infty} g_d(\varepsilon)d\varepsilon=\int_{\mu}^{\infty} g_d(\varepsilon)d\varepsilon, \tag{17}$$

which has a meaning only if

$$\mu=-\frac{\Delta}{2}. \tag{18}$$

Thus, the Fermi energy of our model system in the thermodynamic limit is a fundamental quantity depending only on the energy of charge transfer from an acceptor to a donor.

III. METHOD

A. Algorithm of energy minimization

We start from a random allocation of N d and N a sites in the continuous D-dimensional system with the density $n=1$, and then generate an initial microscopic state (IMS) (n_a,n_d) of the sample **R** by charging randomly chosen $C_a\times N$ both donors and acceptors (usually we take $C_a=0.7$). Further, we search for such a microscopic state (n_a^0,n_d^0) which obeys the stability conditions (7). We used an algorithm which is an extension of the algorithm proposed in Ref. 10 to the case $\Delta\neq\infty$ and which consists of the three main steps.

In order to save computer time, first we look for pairs $a^0-a^-(d^0-d^+)$ for which the ''crude'' stability relation $\Delta\varepsilon\equiv\varepsilon^0_{a(d)}-\varepsilon^1_{a(d)}>0$ is violated. Then, the energy of the system is decreased by transferring an electron between such a pair of sites for which $\Delta\varepsilon$ has its minimal nonpositive value. This process is repeated until a state is reached, in which $\Delta\varepsilon>0$ for all possible a^0-a^- and d^0-d^+ pairs (step I). In a similar manner, we further minimize the energy of the system with respect to the ''true'' stability relations (7) for the charge transfer between the a^0-a^- and d^0-d^+ pairs (step II). And, finally, in step III we diminish the energy of the system with respect to the stability relations for ionization and recombination processes. Since ionization and recombination processes change the degree C_a of the system ionization, each time after one of these processes takes place during calculations, we go back to step II. Repeating steps II and III, we finally arrive at a microscopic state (n^0_a,n^0_d) for which all four stability conditions are fulfilled. We name this procedure $(n_a,n_d)\rightarrow(n^0_a,n^0_d)$ ''a single descent.''

It should be noted, however, that the state (n^0_a,n^0_d) is not necessarily the ground state of sample **R**, since for the ground state, in general, not only the simplest relations (7) with only pairs of sites included, but more complicated relations involving quadruplets, sextets, etc. of sites have to be fulfilled. Therefore, the state (n^0_a,n^0_d) (after Ref. 10) hereafter will be referred to as the pseudo–ground state (PGS) of the sample **R**. In order to check how the ''erroneousness'' of PGS influences the outcome of our calculations, we also performed an analysis of ground states obtained by means of *m*-rank descent which comprises a sequence of the single descents on the same sample with different IMS when calculations are stopped after the lowest observed PGS energy repeats *m* times. We find that $g_a(\varepsilon)$ and $g_d(\varepsilon)$ obtained for the PGS's generated by the single descents and by descents with $m=10$, say, do not differ within the limits of statistical errors. Note, that in the *d-a* models studied before,[13,15] $g(\varepsilon)$ decreases (leaving qualitative behavior unchanged, though) when many-particle excitations are taken into account. This is not the case in MCDAM where even the simplest stability conditions give good-quality PGS. We think this is due to the ionization and recombination processes included in our calculations. So, we conclude, that reliable results can be obtained by means of single descents already, thereby saving a lot of computer time and resources.

B. Finite-size effects

Due to constraints in computer resources, the largest samples we were able to deal with, comprise up to $N=2000$ donor and $N=2000$ acceptor sites ($L\sim45$ for $D=2$ and $L\sim12$ for $D=3$). Such relatively small sizes of the samples investigated might influence the outcome of calculations. Detailed analysis of finite size effects on the results obtained will be presented in Sec. IV, and here we want to make two remarks about inherent finite size effects in the model system considered.

First, as follows from Eq. (7), for finite samples at $T=0$ $g_a(\varepsilon)=0$ within the interval

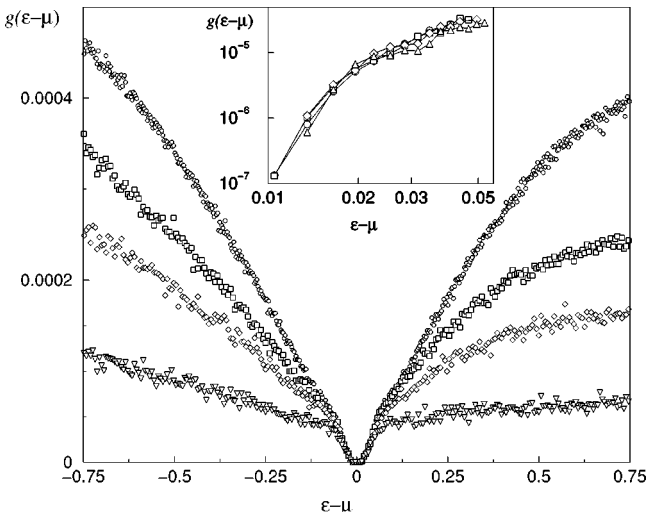

FIG. 1. $g(\varepsilon-\mu)$ for $D=2$ with $N=1500$ at $\Delta=0$ (circles), 2 (squares), 4 (diamonds), and 10 (triangles) averaged over 10 000 ($\Delta=0$), 5100 ($\Delta=2$), 3700 ($\Delta=4$), and 2200 ($\Delta=10$) samples. Insert shows double logarithmic plot of $g(\varepsilon-\mu)$ for $\varepsilon>\mu$ in the region $\varepsilon-\mu\lesssim0.05$.

$$|\varepsilon-\mu|<(2L\times\sqrt{D})^{-1}. \tag{19}$$

The relation (19) gives the estimation of how close to μ data on the energy spectrum are, in principle, obtainable from the calculations on finite samples.

Secondly, μ for finite samples does differ, in general, from sample to sample. A straightforward averaging of $g(\varepsilon)$ [due to the symmetry relation (15) hereafter we consider $g(\varepsilon)\equiv g_a(\varepsilon)$ only] over different samples might thus lead to a distortion of the $g(\varepsilon)$ shape especially in the region where the Coulomb gap is observed. In order to avoid this undesired effect, we added together $g(\varepsilon)$ for the same values of $\varepsilon-\mu(\mathbf{R})$ rather than for the same values of ε with $\mu(\mathbf{R})$ calculated as

$$\mu(\mathbf{R})=\frac{1}{2}(\min\{\varepsilon^0_a(i)\}+\max\{\varepsilon^1_a(i)\}). \tag{20}$$

Finally, we remark that all the data presented below were obtained for the open boundary condition. In order to ensure that results obtained are not determined by the type of the boundary conditions used in calculations, we performed calculations of the $D=2$ MCDAM at $\Delta=0$ with different boundary conditions and found that behavior of $g(\varepsilon)$ is not governed by the boundary conditions used. For example, $g(\varepsilon)$'s for $D=2$ system with $N=500$ and open boundary conditions and with $N=2000$ and periodic boundary conditions do not differ within the limits of statistical errors for any ε calculated.

IV. RESULTS

Figure 1 shows $g(\varepsilon-\mu)$ in the vicinity of the Fermi energy μ obtained for the $D=2$ samples with $N=1500$ and various values of Δ. As is seen, $g(\varepsilon-\mu)$ depends considerably on Δ except for a narrow window $|\varepsilon-\mu|\lesssim0.05$, where

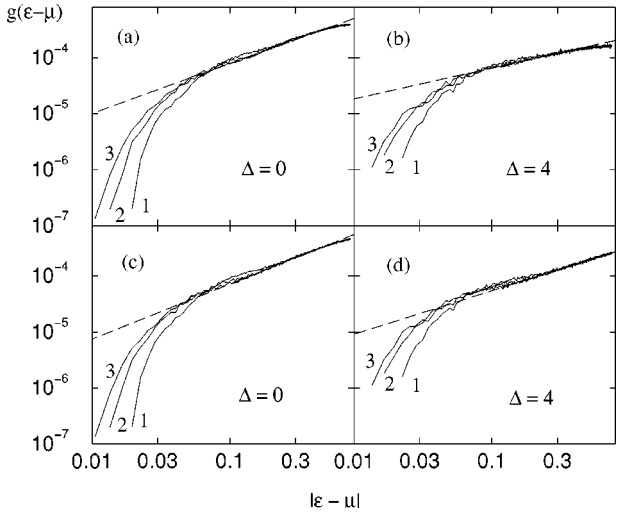

FIG. 2. $g(\varepsilon-\mu)$ for $\varepsilon>\mu$ (a), (b) and $\varepsilon<\mu$ (c), (d) for $D=2$ at $\Delta=0$ (a), (c) and 4 (b), (d), with $N=500$ (curves numbered 1), 1000 (2), and 1500 (3) averaged over 10 000 samples (except the case $N=1500$ and $\Delta=4$ with the average over 3700 samples). The dashed lines are least-squares power-law fits $g(\varepsilon-\mu)\sim|\varepsilon-\mu|^{\gamma}$ with $\gamma=0.9$ (a), 0.55 (b), 0.98 (c) and 0.78.

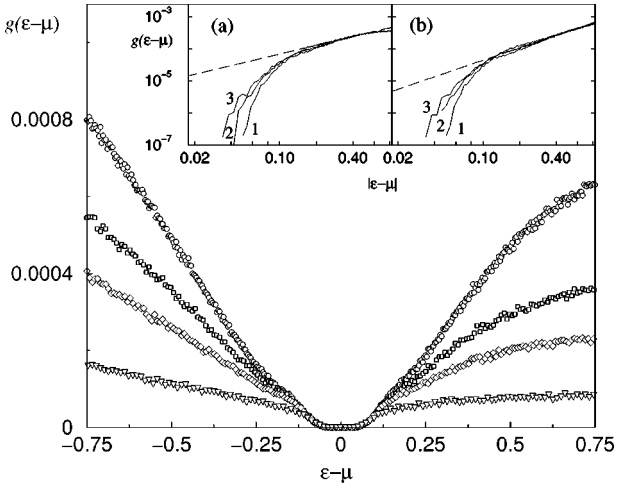

FIG. 3. $g(\varepsilon-\mu)$ for $D=3$ with $N=1000$ at $\Delta=0$ (circles), 2 (squares), 4 (diamonds), and 10 (triangles) averaged over 10 000 samples. Inserts show double-logarithmic plots of $g(\varepsilon-\mu)$ at $\Delta=2$, for $N=500$ (curves numbered 1), 1000 (2), and 2000 (3), in the regions $\varepsilon>\mu$ (a) and for $\varepsilon<\mu$ (b). The dashed lines in the inserts are least-squares power-law fits $g(\varepsilon-\mu)\sim|\varepsilon-\mu|^{\gamma}$ with $\gamma=1.16$ (a), 1.29 (b).

all data merge into some ''universal'' curve symmetric with respect to μ, the curve which can be anticipated to obey the Efros UH (1). However, a double-logarithmic plot of the ''universal'' $g(\varepsilon-\mu)$ (insert in Fig. 1), reveals that $g(\varepsilon-\mu)$ in the ''universality'' region is not even a power law. The width of this ''universality'' region is comparable to the width of the region (19) (for the data presented in Fig. 1 $|\varepsilon-\mu|<0.011$), so it is plausible to suggest that the ''universal'' behavior of $g(\varepsilon-\mu)$ is governed by the finite-size effects. This is clearly demonstrated in Fig. 2 where $g(\varepsilon-\mu)$ are shown for several N investigated.

The ε window where finite size effects are severe, shrinks considerably with increasing N for all values of Δ we investigated. For instance, $g(\varepsilon-\mu)$ for $N=500$ and $N=1000$ at $\Delta=0$ [see Figs. 2(a), 2(c)] merge when $|\varepsilon-\mu|\gtrsim0.2$ while corresponding curves for $N=1000$ and $N=1500$ are indistinguishable already at $|\varepsilon-\mu|\gtrsim0.1$. The statistical noise observed for the curves in Fig. 2 is quite small even close to μ and hence, the influence of insufficient large statistics on the results obtained is excluded. Note that the ''universal'' behavior of $g(\varepsilon)$ in the vicinity of μ obtained for the classical *d-a* model (see Fig. 3 in Ref. 12) is most likely due to the finite size effects as well.

In the region $|\varepsilon-\mu|\gtrsim0.2$, where the curves for all N collapse into a single curve (and where we believe the thermodynamic limit is reached), $g(\varepsilon-\mu)$ is described by a power law $g(\varepsilon-\mu)\sim|\varepsilon-\mu|^{\gamma}$. The deviation from the power law observed far away from $\mu(|\varepsilon-\mu|\gtrsim0.7)$ is due to the boundaries of the Coulomb gap which are ~1 in units of E_0. One can see from a comparison of the data shown in Fig. 2, that the exponent γ depends considerably on Δ. Furthermore, values of γ in the region $\varepsilon>\mu$ and those in the region $\varepsilon<\mu$ differ as well with this difference increasing with increasing Δ. The data for γ obtained for the $D=2$ MCDAM are summarized in Fig. 4 where a significant deviation of γ from the value $D-1$ predicted by the hypothesis (1) is observed at all values of Δ investigated except for the case $\Delta=0$ when $\gamma\approx1$ within the limits of statistical accuracy. Note that the deviation of γ from its predicted value grows monotonically with increasing Δ. At $\Delta=10$ where the features of the MCDAM are expected to be nearly the same as those of the classical *d-a* model with all the acceptors being ionized (indeed, the degree of the acceptor ionization $C_a\sim0.9$ for the two-dimensional MCDAM at $\Delta=10$, see Fig. 5 below) the deviation from the Efros exponent is very large.

The main results for $g(\varepsilon-\mu)$ obtained for the $D=3$ MCDAM are summarized in Figs. 3 and 4. It is seen that the behavior of $g(\varepsilon-\mu)$ in three dimensions does not differ qualitatively from the behavior of $g(\varepsilon-\mu)$ in two dimensions. Some quantitative differences observed arise from the fact that at given N (the parameter which determines the amount of computer memory needed for the calculations) the linear size of a $D=2$ sample with a given density of sites is larger than that of a $D=3$ sample with the same density of sites, and thereby, the finite size effects for $D=3$ samples are more pronounced. For example, the lower boundary of the region where $g_a(\varepsilon-\mu)$ can be described by the power law $|\varepsilon-\mu|^{\gamma}$ shifts towards larger $|\varepsilon-\mu|\gtrsim0.4$ values (see inserts in Fig. 3). Remarkably, the exponent γ does not reach the value $D-1$ predicted by the UH (1) even at $\Delta=0$ (see upper curves in Fig. 4).

Unlike $g(\varepsilon-\mu)$ in the vicinity of the Coulomb gap, C_a describes the state of the entire sample and therefore reaches the thermodynamic limit much faster than $g(\varepsilon-\mu)$. This allows us to obtain quite accurate results for C_a from data on a relatively small amount of samples with $N=500$ only. Figure 5 shows the variations of C_a with Δ both for $D=2$ and $D=3$. In $D=3$ almost all acceptors become ionized rather soon, while for two dimensions even for the largest Δ investigated, around 10% of the acceptors remain neutral. So, one

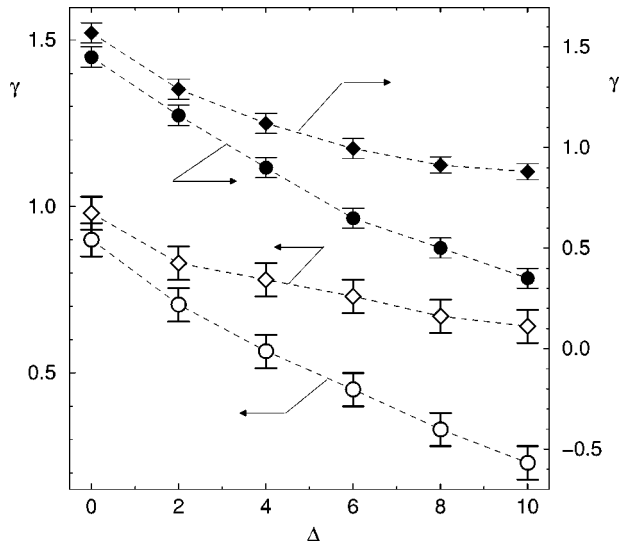

FIG. 4. The exponents γ as a function of Δ for $D=2$, $N=1500$ (open symbols, left scale) and $D=3$, $N=1000$ (filled symbols, right scale). Circles represent the positive values of $\varepsilon-\mu$ while diamonds the negative values of $\varepsilon-\mu$. Dashed lines are guides to the eye.

can say, that the $D=3$ MCDAM at $\Delta \gtrsim 7$ reduces already to the classical *d-a* model. It is known that the classical $D=3$ *d-a* model exhibits the so-called Coulomb fluctuational catastrophe.[17] For calculations on finite samples it implies that statistical fluctuations of $\mu(\mathbf{R})$ grow dramatically with increasing Δ which is the case in our calculations (see Table I). Therefore, in order to reduce the statistical noise in $D=3$, the average of $g(\varepsilon-\mu)$ over a much larger (compared to $D=2$) number of samples is needed. Note, that $\mu(\mathbf{R})$ in both two and three dimensions are scattered according to the Gaussian distribution with the mean $\overline{\mu}$ obeying the relation (18).

V. DISCUSSION

Despite the fact that $g(\varepsilon-\mu)$ is indeed described by a power law in a wide range of ε inside the region of the

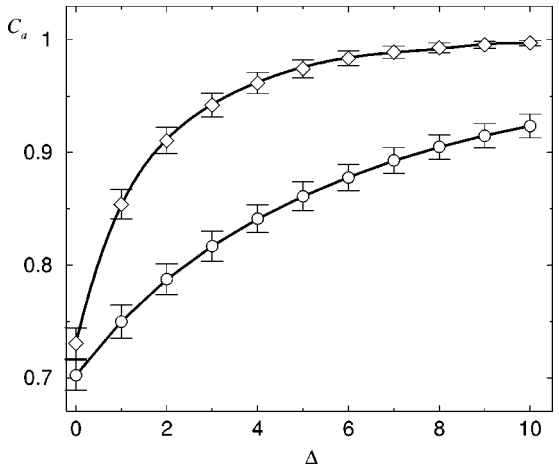

FIG. 5. C_a as a function of Δ for $D=2$ (circles) and $D=3$ (diamonds) with $N=500$ averaged over 1000 samples. The solid lines are third-degree polynomial fits.

TABLE I. The means $\overline{\mu}$ and standard deviations $\Delta\mu$ of the Fermi energy calculated for the three-dimensional model (3) with $N=1000$ and various Δ.

Δ	$\overline{\mu}$	$\Delta\mu$
0	−0.017	0.100
2	−1.016	0.187
4	−2.0149	0.287
6	−3.016	0.392
8	−4.011	0.502
10	−5.018	0.607

Coulomb gap, the exponent γ is considerably smaller than that predicted by the UH both for $D=2$ and $D=3$. Moreover, the exponent γ depends significantly on Δ and is different for the cases $\varepsilon>\mu$ and $\varepsilon<\mu$. It is believed that information about $g(\varepsilon)$ might be directly obtained from tunneling and photoemission experiments,[25] and recent experiments[26] on boron-doped silicon crystals have shown that $g(\varepsilon)$ at higher energies obeys a power law with an exponent slightly less than 0.5, which is in good agreement with our results for $D=3$ and $\Delta \gtrsim 8$. However, the nonmetallic samples show around the Fermi energy a nearly quadratic Coulomb gap, so the question arises whether our results could be related to the intermediate asymptotic behavior observed. Here we want to make three remarks concerning this question:

First, the power law is valid above a value $\varepsilon_0(N)$ below which the finite size effects take over (Figs. 2 and 3). It seems from our results, that $\varepsilon_0(N)\to\mu$ when $N\to\infty$. In $D=2$ we were able to obtain size-independent results down to $\varepsilon_0 \sim 0.1$, i.e., for $\sim 90\%$ of the whole Coulomb gap.

Secondly, as follows from Eq. (2), the distance r_{ij} between d^0, with $\varepsilon_i^1 \in [-\varepsilon,0]$ [ε here is the half-width of a narrow band around $\mu=0$ (Ref. 28)] and d^+ with $\varepsilon_j^0 \in [0,\varepsilon]$ should be not less than $1/2\varepsilon$. That is, sites with energies $\varepsilon_i^1 \in [-\varepsilon,0]$ cannot be inside a D-dimensional sphere of radius $R_{sp}=1/2\varepsilon$ and with the center in a site with the energy $\varepsilon_j^0 \in [0,\varepsilon]$. Assuming that *all* such spheres *do not intersect*, the total volume occupied by the spheres is

$$V_{sp}=N\times S(D)\left(\frac{1}{2\varepsilon}\right)^D \int_0^{\varepsilon} g(\varepsilon')d\varepsilon', \tag{21}$$

where $S(D)$ is the volume of a D-dimensional sphere with the radius equal to unity. Since V_{sp} cannot exceed the total volume V of a sample ($V=N$ at $n=1$) we arrive at the inequality

$$\int_0^{\varepsilon} g(\varepsilon')d\varepsilon' \leqslant \frac{(2\varepsilon)^D}{S(D)}, \tag{22}$$

which is valid for all ε if

$$g(\varepsilon) \leqslant \frac{D\times 2^D}{S(D)}|\varepsilon|^{D-1}. \tag{23}$$

The UH then is a limit case of Eq. (23). The density of sites with energies $\varepsilon_i^1 \in [-\varepsilon,0]$ indeed decreases when $\varepsilon\to 0$, so

TABLE II. Some donor-acceptor pairs for which the difference between the donor and acceptor energy levels does not exceed 10 meV. E_g, E_v, and E_c are, respectively, the energy gap, the top of the valence band, and the bottom of the conductivity band. If the minimal/maximal concentrations of both donor and acceptor are known, the parameter E_0 is calculated using the data for the less soluble of the pair.

Donor	Acceptor	Concentration, cm^{-3} min	Concentration, cm^{-3} max	E_0, meV min	E_0, meV max	E_j, meV	Δ, meV
			Si($E_g=1124$ meV, $\chi=12$)				
Fe		1.2×10^{16}	4×10^{16}	0.6	4	E_c-796	8
	Zn	2.3×10^{16}	8×10^{16}			E_v+320	
Ni		10^{18}	10^{13}	12	<1	$E_v+(160\div190)$	3.1 – 33.1
	In	3×10^{17}	4×10^{18}			$E_v+156.9$	
			Ge($E_g=740$ meV, $\chi=15.9$)				
S		no reliable data				E_c-296	4
	Ni	4.8×10^{15}	8×10^{15}	1.5	1.8	E_c-300	
			GaAs($E_g=1520$ meV, $\chi=12.5$)				
Ti		2×10^{16}		3.1		E_c-1000	0
	Fe					E_v+520	

the assumption (21) for the spheres with *finite* radii seems to be plausible. However, simultaneously $R_{sp}\to\infty$ and consequently the plausibility of the assumption (21) and thereby of Eq. (1) becomes questionable.

Finally, Eq. (1) can be also obtained as the asymptotic behavior of a nonlinear integral equation for $g(\varepsilon)$ as $\varepsilon\to0$, the equation which, in turn, is heuristically obtained from Eq. (2). The derivation of this integral equation (given, for example, in Ref. 19) is based on the implicit assumption that the sites with charged donors are randomly distributed in space according to the Poisson statistics. However, it was unequivocally demonstrated in computer studies of the Coulomb gap[12] that charged donor sites with energies close to μ tend to form clusters (Ref. 12, Fig. 6).

We conclude that $g(\varepsilon-\mu)$ in the region of the Coulomb gap in model (3) has a power law behavior for all energies down to μ, and that the UH of Efros (1) is questionable. Note that our results are in contradiction not only to the UH (1), but to the inequality (23) as well ($\gamma<D-1$). Up to now, all exponents found are in good agreement with this inequality. For example, in Ref. 13 specimens of 40 000 and 125 000 sites for two- and three-dimensional samples were investigated in the Efros' lattice model,[16] and it was found there that if $g(\varepsilon)$ follows a power law, $\gamma=1.2\pm0.1$ and $\gamma=2.6\pm0.2$ for two and three dimensions, respectively.

Energy levels of donor (acceptor) impurities are usually close to the bottom (top) of the conduction (valence) band. Since in the most common semiconductors the energy gap $E_g\sim10^4$ K and $E_0\sim20$ K, $\Delta\gg1$, one may ask what physical relevance does the model (3) with a finite $\Delta\lesssim10$ have, except for being a pure academic exercise? However, in the case of deep impurities the energy levels for some donor-acceptor pairs are extremely close to each other, not excluding even the case $\Delta=0$.[27] Table II shows some donor-acceptor pairs with $\Delta\lesssim10$ in the most common semiconductors. The solubilities of these impurities are rather low, thereby reducing the temperature at which the Coulomb gap with features described by the model (3) can be observed. Fortunately, these temperatures are high enough ($\sim10\div20$ K) for modern experimental techniques and hence experimental observation of the Coulomb gap in the semiconductors with deep impurities is possible to accomplish.

VI. SUMMARY

We have studied a model of impurities in semiconductors with infinite-range Coulomb interactions between donors, between acceptors, and between donors and acceptors. A new parameter introduced in the model is the finite energy Δ of charge transfer between donors and acceptors, a parameter which enables processes of ionization of neutral impurities and of recombination of charged impurities. In the particular case of equal amounts of donor and acceptor impurities, we derived rigorous relations for the symmetry of the model with respect to exchange of donor and acceptor sites. We also extended the previously known algorithm to find the ground state including the stability relations with respect to ionization and recombination processes and performed computer studies of the model proposed at zero temperature on a number of two- and three-dimensional samples with randomly distributed N donors and N acceptors. We explored the energy region around the Fermi energy μ where the Coulomb gap in the single-particle density of states $g(\varepsilon)$ is observed. The analysis of the calculated histograms $g(\varepsilon)$ revealed that the behavior of $g(\varepsilon)$ obtained from the simulations on finite samples in the immediate neighborhood of μ is determined solely by the finite size effects. In the

region where finite size effects become negligible, $g(\varepsilon)$ is described by a power law with an exponent considerably depending on the parameter Δ and on the sign of $\varepsilon-\mu$. Our findings challenge the Efros universality hypothesis. Moreover, our results are in contradiction to the main inequality (23) of which Efros' universality hypothesis is a particular case. We have reexamined the heuristic derivation of the Efros hypothesis and shown that some implicit assumptions which lead to universality are questionable. From the analysis of experimental data on admixtures in semiconductors we put forward possible experimental situations where one could observe the Coulomb gap with the features being the same as those of the model with a finite Δ.

ACKNOWLEDGMENTS

This research was supported by The Swedish Natural Science Council and by The Swedish Royal Academy of Sciences. S.A.B. and V.A.O. were partially supported by Grant No. RFFR 00-15-97334.

[1] M. Pollak, Discuss. Faraday Soc. **50**, 13 (1970); G. Srinivasan, Phys. Rev. B **4**, 2581 (1971); V. Ambegaokar, B.J. Halperin, and J.S. Langer, *ibid.* **4**, 2612 (1971).

[2] B.L. Altshuler and A.G. Aronov, Solid State Commun. **30**, 115 (1979).

[3] 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).

[4] Y. Imry and Z. Ovadyahu, Phys. Rev. Lett. **49**, 841 (1982); A.E. White, R.C. Dynes, and J.P. Garno, Phys. Rev. B **31**, 1174 (1985).

[5] N.F. Mott, J. Non-Cryst. Solids **1**, 1 (1968).

[6] A.L. Efros and B.I. Shklovskii, J. Phys. C **8**, L49 (1975); B.I. Shklovskii and A.L. Efros, Fiz. Tekh. Poluprovodn. **14**, 825 (1980) [Sov. Phys. Semicond. **14**, 487 (1980)].

[7] M.A. Pollak, J. Non-Cryst. Solids **11**, 1 (1972); E.M. Hamilton, Philos. Mag. **26**, 1043 (1972).

[8] W. Schoepe, Z. Phys. B: Condens. Matter **71**, 455 (1988); A. Andronard, A. Kouzoun, N. Cherradi, G. Marshal, and M. Gerl, Philos. Mag. B **59**, 207 (1989).

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

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

[11] A.L. Efros, N. van Lien, and B.I. Shklovskii, J. Phys. C **12**, 1869 (1979); S.D. Baranovskii, A.A. Usakov, and A.L. Efros, Zh. Éksp. Teor. Fiz. **83**, 756 (1982) [Sov. Phys. JETP **56**, 422 (1982)].

[12] J.H. Davies, P.A. Lee, and T.M. Rice, Phys. Rev. B **29**, 4260 (1984).

[13] A. Möbius and M.R. Richter, J. Phys. C **20**, 539 (1988); A. Möbius, M.R. Richter, and B. Drittler, Phys. Rev. B **45**, 11 568 (1992).

[14] K. Tenelsen and M. Schreiber, Europhys. Lett. **21**, 697 (1993); Phys. Rev. B **49**, 12 662 (1994).

[15] M. Chickon, M. Ortuno, and M. Pollak, Phys. Rev. B **37**, 10 520 (1988).

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

[17] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Berlin, 1984), p. 409.

[18] A.L. Burin, J. Low Temp. Phys. **100**, 309 (1995).

[19] A.A. Mogilyansky and M.E. Raich, Zh. Éksp. Teor. Fiz. **95**, 1870 (1989) [Sov. Phys. JETP **68**, 1081 (1989)].

[20] M. Sarvestani, M. Schreiber, and T. Vojta, Phys. Rev. B **52**, R3820 (1995); F. Epperlein, M. Schreiber, and T. Vojta, *ibid.* **56**, 5890 (1997).

[21] E.R. Grannan and C.C. Yu, Phys. Rev. Lett. **71**, 3335 (1993).

[22] V.N. Likhachev and V.A. Onishchouk, Phys. Lett. A **244**, 437 (1998).

[23] T. Vojta and M. Schreiber, Phys. Rev. Lett. **73**, 2933 (1994).

[24] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glasses Theory and Beyond* (World Scientific, Singapore, 1987).

[25] A.E. White, R.C. Dynes, and J.P. Garno, Phys. Rev. Lett. **56**, 532 (1986); J.G. Massey and M. Lee, *ibid.* **75**, 4266 (1995).

[26] M. Lee, J.G. Massey, V.L. Nguyen, and B.I. Shklovskii, Phys. Rev. B **60**, 1582 (1999).

[27] *Semiconductors: Impurities and Defects in Group IV Elements and III-V Compounds*, Landolt-Börnstein New Series, v. III/22, edited by M. Schulz (Springer-Verlag, Berlin, 1989); A. Bargys and J. Kundrotas *Handbook on Physical Properties of Ge, Si, GaAs and InP* (Science and Encyclopedia Publ., Vilnius, 1994).

[28] For the sake of simplicity, we took $\mu=0$ in Eqs. (21)–(23). In the general case $\mu\neq 0$, one should replace ε in Eq. (23) by $\mu-\varepsilon$.