Coulomb effects and hopping transport in granular metals

I. S. Beloborodov, A. V. Lopatin, and V. M. Vinokur

Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

(Received 6 April 2005; revised manuscript received 28 July 2005; published 22 September 2005)

We investigate effects of Coulomb interaction and hopping transport in the insulator phase of granular metals and quantum dot arrays considering both spatially periodic as well as irregular grain or dot arraignments. We study the Mott transition between the insulating and metallic phases in a strictly periodic system and find the dependence of the Mott gap on the intergranular coupling. In this case the conductivity of the insulating state has the activation form with the Mott gap entering the exponent. In the irregular arrays the electrostatic disorder induces the finite density of states near the Fermi level giving rise to the variable range hopping conductivity. We derive the transport properties of the irregular array in the dielectric, low coupling limit and show that the conductivity follows the Efros-Shklovskii law. We develop a theory of tunneling through a chain of grains and discuss in detail both elastic and inelastic cotunneling mechanisms; the former dominates at very low temperatures and very low applied electric fields, while the inelastic mechanism controls tunneling at high temperature or fields. Our results are obtained within the framework of the technique based on the mapping of the quantum electronic problem onto the classical gas of Coulomb charges. The processes of quantum tunneling of real electrons are represented in this technique as trajectories (world lines) of charged classical particles in $d+1$ dimensions. The Mott gap is related to the dielectric susceptibility of the Coulomb gas in the direction of the imaginary time axis.

I. INTRODUCTION

Electron transport in granular metals is a subject of intense theoretical and experimental research.[1–9] The interest is motivated by the fact that granular metals represent a unique tunable system that captures all the essential features of general disordered systems and allows for the study of the effects of interference of strong Coulomb interaction and disorder. In particular, depending on the intergranular conductance, granular metals exhibit the wide spectrum of transport behaviors ranging from the hopping conductivity to those typical to disordered Fermi liquids.

At present the metallic strong coupling regime where the Coulomb repulsion is substantially screened is understood fairly well.[3,5] The key energy scale in this regime is the inverse escape rate from a grain $\Gamma = g\delta$, where δ is the mean energy level spacing in the individual granule, and g is the intergranular conductance. Generally, on the energy scales larger than Γ the granularity is essential, while at lower energy scales the system resembles an ordinary disordered metal with an effective diffusion coefficient determined by the coupling between the grains. As a result at temperatures larger than Γ the conductivity exhibits the logarithmic temperature dependence in all dimensions, while the low temperature phase at $T < \Gamma$, (Ref. 3) can be viewed as the disordered Fermi liquid[5,6] with conductivity described by the Altshuler-Aronov[10] and localization type corrections.[11]

The dielectric regime where the conductivity is known to depend on the temperature as[12]

$$\sigma = \sigma_0 \exp[-(T_0/T)^{1/2}], \qquad (1)$$

with σ_0 being the high temperature conductivity and T_0 being the characteristic energy scale defined below, has been long posing a puzzle. This form of the observed behavior

suggested the primary role of Coulomb correlations and several models were proposed, yet the problem of conduction in granular metals remained unsolved; see Ref. 13 for a review. Building on the constructive critique of Ref. 13, the important development towards understanding the observed behavior (1) in granular metals was made in Ref. 4, where the role of electrostatic disorder in formation of random potentials on individual granules was fully recognized and explored. This suggested that the dependence (1) can indeed be viewed as Mott-Efros-Shklovskii variable range hopping in the presence of the Coulomb gap.[14,15] The remaining problem however was that in order to realize variable range hopping conductivity an electron would have to tunnel through the sequence of granules in one hop and no mechanism that could provide such a tunneling process was described.

In a Letter[16] we have reported the mechanism, the multiple cotunneling, via which the correlated hopping in a Coulomb gap occurs in granular conductors. We developed the technique allowing rigorous consideration of Coulomb effects thus putting the theory of hopping transport in granular conductors on a firm quantitative theoretical basis. In the present work we investigate in detail the effects of the Coulomb interaction on transport properties of granular metallic arrays concentrating on the regime of low and intermediate intergranular coupling. We discuss both the model of a periodic granular array as well as the irregular model that includes the random on-site electrostatic potential. We begin with the periodic model and show that the regular array of metallic granules is a Mott insulator at zero temperature, provided the tunneling conductance is less than the critical value $g_c = (1/\pi z)\ln(E_0^c/\delta)$ where z is the coordination number (number of neighbors for a site on the array), δ is the mean energy level spacing, and E_0^c is the Coulomb charging energy in a single grain. The Mott gap, the characteristic

property of the insulating state, decreases with the increase of the intergranular conductance and disappears when g reaches the critical value g_c.

We develop an approach to the insulating state of granular conductors based on mapping the original quantum model describing electrons in the granular system on the classical gas of the interacting Coulomb charges. More specifically, we first absorb the Coulomb interaction into the phase field, similarly to Ref. 17 and next map the quantum problem that involves functional integration over the phase fields onto the classical model that has an extra time dimension—in an analogy with the method of Ref. 18. The classical charges of the emerging effective model are connected via the electron world lines representing graphically tunneling processes and form necklacelike loops. In the insulating state only short loops are present; this means that electrons and holes form bound states and only virtual tunnelings over the short distances are allowed. In the metallic phase the loops of the infinite size are present meaning that electrons and holes are not bound together. We show that the Mott gap in the insulating state is given by the time-axis component of the inverse dielectric susceptibility of the classical charged gas.

Mott transition in a periodic granular array resembles the Mott transition in the Hubbard model.[19] However, there is an important difference between the two. Namely, in granular metals, even in the case of periodic samples, the electron motion within a grain as well as on the scales exceeding the intergranular distance is diffusive, while in the Hubbard model electrons move through a periodic lattice that allows one to label their states by quasimomenta. The model of a granular metal also has an extra physical parameter δ, the mean energy levels spacing in a single grain, that has no analog in the Hubbard model. Yet, the physics of the transition is similar in both cases, and one may expect that the methods developed for the study of the Mott transition in the system of strongly interacting electrons on the lattice (such as dynamical mean field theory,[19] for example) can be applied to the granular metals as well. At present, however, the description of the Mott gap in granular arrays in the very vicinity of the Mott transition is not available, in particular the order of the transition remains an unresolved problem.

The spatially regular granular arrays, including technologically important cases of arrays of nanocrystals or quantum dots, are now the subject of the extensive experimental study. One might expect that periodic arrays would show activation transport, characteristic for the systems with the hard gap in the electron spectrum. However, even seemingly perfectly periodic arrays[2,20,21] do not follow the expected activation conductivity behavior. Instead, the experimentally observed conductivity follows the Efros-Shklovskii (ES) law[22] expressed by Eq. (1). Dependence (1) cannot be observed in the strictly periodic structures without disorder,[23] where the hard gap in the excitation spectrum results in the activation conductivity behavior. This indicates that electrostatic disorder plays a primary role for transport properties of granular systems.

The two necessary ingredients for the variable-range-hopping type(VRH) conductivity are (i) the finite density of states at energies close to Fermi level and (ii) the exponentially falling probability of tunneling between the states close to the Fermi level. Furthermore, the transformation of the usual Mott's VRH dependence into the Efros-Shklovskii law reflects the presence of the so-called soft gap in the electron spectrum resulting from the strong long range Coulomb correlations. The problem of hopping transport in granular conductors is thus twofold: (i) to explain the origin of the finite density of states near the Fermi level and the role of Coulomb correlations and (ii) to elucidate the mechanism of tunneling through the dense array of metallic grains.

Earlier attempts to explain the behavior (1) were based on the assumption of the large dispersion in the granule's size.[12] Moreover, one more strong assumption was necessary in order to arrive at formula (1): that the distance between the grains and the grain sizes are not statistically independent but rather are linearly coupled with each other, and this hypothesis contradicted experimental findings.[13] Furthermore, recent observations of ES law were made in the arrays of quantum dots[20] and artificially manufactured granular systems[21] where the size of granules or periodicity in the dots arrangement was controlled within the few percent accuracy. Thus the basic assumption about the large size or distance dispersion does not apply to all these systems, yet the dependence (1) is observed. Moreover, as was pointed out in Refs. 4 and 13, the capacitance disorder alone can never lift the Coulomb blockade in a single grain completely, and thus this mechanism for sure cannot explain the finite density of states at the Fermi level, and as a consequence it cannot explain dependence (1) at low temperatures.

To address this problem properly, one notices that the insulating matrix in granular conductors that is typically formed by the amorphous oxide may contain a deep tail of localized states due to carrier traps. The traps with energies lower than the Fermi level are charged and induce the potential of the order of $e^2/\kappa r$ on the closest granule, where κ is the dielectric constant of the insulator and r is the distance from the granule to the trap. This compares with the Coulomb blockade energies due to charging metallic granules during the transport process. This mechanism was considered in Ref. 4. If granules are very small, surface effects also contribute to random changes in associated Coulomb energies. And, finally, talking about the 2D granular arrays and/or arrays of quantum dots, one expects that the source of the induced random potential are imperfections of and the charged defects in the substrate. Thus in granular conductors the role of the finite density of impurity states is taken up by random shifts in the granule Fermi levels due to all the above reasons. We describe these shifts as the external random potential μ_i, where i is the grain index, applied to each site of a granular system.

Tunneling over the distances well exceeding the average granule size in a dense granular array is a fundamental physical problem. The virtual electron tunneling through a *single* granule (quantum dot) was considered in Ref. 24 where two different mechanisms for a charge transport through a single quantum dot in the Coulomb blockade regime were identified. The first mechanism—the so-called elastic cotunneling—transfers the charge via tunneling of an electron through an intermediate virtual state on the dot such that the electron comes out of the dot with the same energy it came in. In the second mechanism—inelastic

cotunneling—an electron that comes out of the dot has a different energy from the one that came in. After such a tunneling process an electron-hole excitation is left in the dot which absorbs the energy difference. Note that both these processes go via the classically inaccessible intermediate states, i.e., both mechanisms occur as the charge transfer via a virtual state. Inelastic cotunneling dominates the elastic one in the region where the temperature or the applied voltage is larger than $\sqrt{E_0^c \delta}$.

Coulomb charge mapping technique, developed in our work, allows for a simple and transparent analytic description of both elastic and inelastic cotunneling processes offering thus a ground for a comprehensive theory of the hopping transport in granular metals. The advantage of our approach is that the higher order tunneling processes can be included in a straightforward way along with the nearest neighbor hoppings. The obtained results for the hopping conductivity dependencies as well as results concerning regular arrays are summarized in the section below.

Our paper is organized as follows: We summarize our results in Sec. II and introduce our model in detail in Sec. III. In Sec. IV we find the corrections to the Mott gap due to the weak intergranular coupling. In Sec. V we introduce the charge representation for the model of a granular metal and discuss the physical quantities in this representation in Sec. VI. The Mott gap for the periodic sample in the case of moderately large tunneling conductance is found in Sec. VII. In Sec. VIII we generalize our technique to include the high order tunneling processes and apply this technique to find the hopping conductivity in Sec. IX and to study the Mott transition in Sec. X. Our conclusions are presented in Sec. XI while some technical details are relegated into the Appendixes.

II. SUMMARY OF THE RESULTS AND DISCUSSION

To treat the multigranule tunneling and to describe the insulating state we have developed an approach based on the mapping of the quantum model of granular metal onto the classical gas of Coulomb charges in $d+1$ dimensions. The Coulomb charges of the effective classical model are connected by electron world lines and form necklacelike loops. First we apply our technique to calculate the Mott gap Δ_M in the electron spectrum of a periodic granular array and show that it is simply related to the dielectric constant κ_τ of the classical charge gas in the extra time direction. In the insulating regime where virtual tunnelings on large distances are suppressed, the charge model approach is reduced to consideration of the short loops, representing the nearest neighbor hopping. Close to the metal-insulator transition the larger loops that describe tunneling on larger distances become important, and in the metallic regime the loops of an infinite size appear. In this picture, the Mott transition corresponds to the transition between the state where large loops are energetically suppressed and the state where loops of an infinite size do proliferate. It is important that despite the fact that the large loops representing tunneling over large distances are energetically suppressed and do not contribute to thermodynamics of the insulating state, they are essential for trans-

port and govern hopping conductivity in the arrays subject to electrostatic disorder.

We will discuss the physics of the Mott insulator phase and describe the metal to insulator transition in the context of the periodic granular model. The subsequent derivation of hopping conductivity in the weakly coupled arrays subject to external electrostatic disorder does not require special assumptions about the spatial periodicity and holds both for regular and irregular arrangements.

A. Periodic granular arrays

An array of weakly coupled metallic grains, $g \ll 1$, is the charge insulator at zero temperature. The electron transport is blocked due to the presence of the Coulomb gap Δ_M in the electron excitation spectrum. In the limit of vanishingly small intergranular tunneling conductances this gap is given by the Coulomb charging energy of a single grain $\Delta_M = E_0^c$. At finite tunneling conductance the Mott gap Δ_M is reduced due to the virtual electron tunneling to neighboring grains. Suppression of the Mott gap Δ_M due to such virtual processes in the limit of weak coupling can be obtained from the perturbation theory and is given by

$$\Delta_M = E_0^c - \frac{2z}{\pi} g E_{eh} \ln 2, \quad gz \ll 1, \tag{2}$$

where z is the coordination number and E_{eh} is the energy to create an electron-hole excitation in the system by removing an electron from a given site and placing it on one of the neighboring sites. In the case of diagonal Coulomb interaction $E_{ij} = E_0^c \delta_{ij}$ the energy of an electron-hole excitation E_{eh} is simply twice the Coulomb charging energy E_0^c.

From Eq. (2) it follows that the Mott gap Δ_M is reduced significantly at the values of the tunneling conductance $gz \sim 1$, where at the same time the perturbation theory becomes unapplicable.

In order to find the Mott gap Δ_M at larger tunneling conductances, we make use of mapping of the original model onto the model of classical charges and construct a mean field theory considering the coordination number z as a large parameter, $z \gg 1$. Within this theory we obtain that the Mott gap Δ_M is exponentially suppressed with the tunneling conductance

$$\Delta_M = c g E_0^c e^{-\pi z g}, \quad \Delta_M \gg g \delta, \quad gz \gtrsim 1, \tag{3}$$

where c is the number of the order of unity. When deriving Eq. (3), only nearest neighbor virtual electron tunneling were taken into account. In the related effective classical model this corresponds to accounting for the short quadrupole loops only. For this reason the result (3) holds only as long as the Mott gap Δ_M is larger than the inverse electron escape time from the grain $\Gamma = g \delta$. Solving the inequality $\Delta_M > g \delta$ together with the expression for the Mott gap (3) within the logarithmic accuracy we obtain that the latter can be used as long as the tunneling conductance g is less than the critical conductance given by the expression
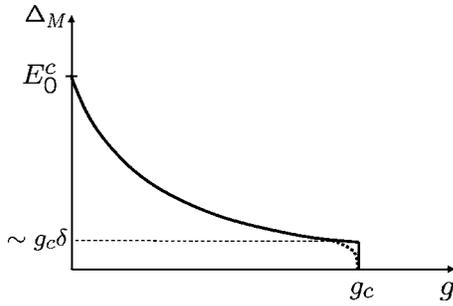
FIG. 1. Schematic dependence of the Mott gap on the tunneling conductance g for a periodic granular array of the dimensionality $d > 2$. At zero tunneling conductance the Mott gap is given by the charging energy in a single grain and it suppressed exponentially with g at $gz \gtrsim 1$. The Mott transition takes place at $\Delta_M \sim g\delta$, where δ is the mean energy level spacing in a single grain. The possible continuous transition that cannot be ruled out at the present time is shown by the dotted line.

$$g_c = \frac{1}{z\pi} \ln(E_0^c/\delta). \tag{4}$$

In the regime where the Mott gap Δ_M becomes of the order of the inverse escape time of an electron from a grain $g\delta$ the large loops become important showing tendency to the formation of the metallic state. This means that the value (4) determines the boundary of stability of the insulating state. At the same time the critical tunneling conductance (4) coincides with the boundary of the stability of the metallic phase obtained in Ref. 5 meaning that the critical value (4) indeed marks the boundary between the metallic and insulating phases at zero temperature.

Considerations leading to the result (4) for the critical tunneling conductance g_c are based on the investigation of the behavior of the system from the insulating and metallic sides implemented in the present paper and Ref. 5, respectively. However, both these considerations do not apply in the immediate vicinity of the phase transition. At present the quantitative description of the Mott transition is not available; in particular, the order of the transition remains an open question. Two possible options are illustrated in Fig. 1.

Presented description of the Mott transition may not apply in one and two dimensions where even the metallic state $g > g_c$ is poorly defined due to divergent Althshuler-Aronov and localization corrections to conductance at $T=0$. The metal to insulator transiton or crossover at higher temperatures, $T > g\delta$, in low dimensional metallic arrays was considered in Refs. 25–27.

The temperature dependence of the conductivity of granular arrays which are insulators at $T=0$ (i.e., with $g < g_c$) is given by the activation law[3]

$$\sigma \sim \sigma_0 e^{-\Delta_M/T}, \quad T \ll \Delta_M, \tag{5}$$

with σ_0 being the high temperature conductivity as long as the temperature is less than the Mott gap. Indeed, the finite temperature conductivity is due to the electrons and holes that are present in the system as real excitations. Their density is given by the Gibbs distribution that results in the

conductivity dependence (5). Note that the expression (5) contains the renormalized Mott gap Δ_M given by Eqs. (2) and (3).

It is important to note that at finite temperatures the continuous transition between the insulating and metallic phases, is, in fact, a crossover rather than the true phase transition. Indeed, the conductivity at nonzero temperature is finite in both phases; besides the Mott transition does not break any symmetry of the system. However, if the transition is discontinuous at $T=0$ it has to remain discontinuous in some finite temperature interval. In this case the Mott transition would be observable even at nonzero temperature via the jumps of the physical quantities such as conductivity.

The results presented above hold only for periodic arrays. Most realistic granular or quantum dots arrays contain disorder, in particular, electrostatic disorder.

B. Granular arrays with electrostatic disorder

Electrostatic disorder causes fluctuations in electrostatic energy of granules and can thus lift the Coulomb blockade at some sites of a granular sample. This results, in its turn, in the finite density of states at the Fermi level and in the variable range hopping appearing as the dominating mechanism of conductivity. The bare density of states induced by the random potential can be substantially suppressed due to the presence of the long-range Coulomb interaction as it takes place in semiconductors where the density of states (DOS) is given by the Efros-Shklovskii expression[15,22]

$$\nu_g(\varepsilon) \sim (\tilde{\kappa}/e^2)^d |\varepsilon|^{d-1}, \tag{6}$$

with e being the electron charge and $\tilde{\kappa}$ being the dielectric constant. This result was recently confirmed in Ref. 28 (see also Ref. 29) within the analytic approach of the locator approximation.

In the limit of weak coupling between the grains, $g \ll 1$, the model of granular array can be described in terms of the classical model that deals with the total charges on the grains only

$$H = \sum_i \mu_i n_i + \sum_{ij} n_i E_{ij}^c n_j, \tag{7}$$

where μ_i is the random external potential, n_i is the electron density on the site i, and E_{ij}^c is the Coulomb interaction between the sites i and j. Taking into account the asymptotic behavior of the Coulomb interaction matrix $E_{ij}^c \sim e^2/2r_{ij}\tilde{\kappa}$, where r_{ij} is the distance between the sites i and j and $\tilde{\kappa}$ is the effective dielectric constant in a granular sample we see that the classical model (7) is essentially equivalent to the one studied by Efros and Shklovskii. Therefore, their results must be applicable to the model of a granular array (7) as well.

At the same time one expects that the DOS in an array of metallic granules is larger than that in a semiconductor since each metallic grain has a dense electron spectrum. Indeed, one has to remember that there are many electron states in a grain that correspond to the same grain charge while in the model of impurity levels in semiconductors the charge is uniquely (up to the spin) identified with the electron state. The energy of the unoccupied state ε_i in the model (7) is by

definition the energy of an electron placed on this state. In the granular metal any state with energy larger than ε_i will also be available for electron placement. Thus, in order to translate the ES result (6) to the density of electron spectrum in granular metals one has to integrate the dependence (6) by the energy ε and multiply it by the bare DOS in a single grain. As a result we obtain

$$\nu(\varepsilon) \sim \nu_0(|\varepsilon|\tilde{\kappa}/e^2)^d, \tag{8}$$

where ν_0 is the average DOS in a single grain (defined as the number of states per energy). However, the above expression cannot be used in the Mott argument for the hopping conductivity where one needs to estimate the distance to the first available state r within the energy shell ε via the relation

$$r^d \int_0^\varepsilon d\varepsilon' \, \nu_g(\varepsilon') \sim 1. \tag{9}$$

The problem with using the expression for DOS (8) in Eq. (9) is that expression (8) assumes strong correlations of the electron states in the space of the array coordinates. Indeed, expression (8) takes into account the fact that if there is a state available for placement of an electron with a given energy, then, typically *on the same grain* there will be plenty of other states available for electron placement. But, for application to the hopping conductivity we should not count different electron states that belong to the same grain since it is enough to find at least one to insure the transport. Thus, when finding DOS relevant for the hopping transport the lowest energy states within the each grain are to be counted only. We arrive at the conclusion that while in granular metals the electron DOS is modified according to (8); in order to find the distance to the first available state within a given energy shell via Eq. (9), even in granular metals, one has to use the ES expression for DOS in its form (6). Similar considerations were presented in Ref. 4. Following this reference we will call the DOS that counts only lowest excited states in each grain and is relevant for the hopping conductivity by the density of *ground* states and will denote by a subscript g as in Eq. (6).

Apart from the knowledge of the spectrum of electron excitations, description of the hopping conductivity requires understanding the tunneling mechanism of an electron between the two distant grains with energies close to the Fermi level through an optimal chain of other grains where the Coulomb blockade is present. At low temperatures such a tunneling process can be realized by means of the virtual electron hopping (cotunneling). As we mentioned in the Introduction one has to distinguish the elastic and inelastic cotunneling processes. The elastic cotunneling is the dominant mechanism for hopping conductivity at temperatures $T < \sqrt{E_0^c \delta}$, while at larger temperatures the electron transport goes via inelastic cotunneling processes.

1. Hopping via elastic cotunneling

Considering the probability for an electron to tunnel from the site i_0 to the site i_{N+1} it is convenient to put the Coulomb interaction energy at these sites to zero and count the initial, ξ_0, and final, ξ_{N+1}, electron energies from the Fermi level.

The presence of the electrostatic disorder on the grains is modeled by the random potential μ_i, such that the energy of the electron (hole) excitation on the site i is

$$E_i^\pm = E_i^c \pm \mu_i. \tag{10}$$

The probability P_{el} of a tunneling process via the elastic cotunneling can be easily found for the case of the diagonal Coulomb interaction $E_{ij}^c = E_i^c \delta_{ij}$:

$$P_{el} = w\bar{g}^{N+1}\left(\frac{\bar{\delta}}{\pi\bar{E}}\right)^N \delta(\xi_{N+1} - \xi_0), \tag{11}$$

where the factor $w = n(\xi_0)(1 - n(\xi_{N+1}))$ takes into account the occupations $n(\xi_0)$ and $n(\xi_{N+1})$ of the initial and final states, respectively, and the bar denotes the geometrical average of the physical quantity along the tunneling path. For example the average tunneling conductance $\bar{g}$ is defined as

$$\ln \bar{g} = \frac{1}{N+1}\sum_{k=0}^{N} \ln g_k, \tag{12}$$

where the summation goes over the tunneling path and g_k is the tunneling conductance between the kth and $(k+1)$st grains. The energy $\bar{E}$ is the geometrical average $\ln \bar{E} = (1/N)\Sigma_{k=1}^{N} \ln \tilde{E}_k$, of the following combination of electron and hole excitation energies

$$\tilde{E}_k = 2(1/E_k^+ + 1/E_k^-)^{-1}. \tag{13}$$

The presence of the delta function in Eq. (11) reflects the fact that the tunneling process is elastic.

The result (11) for the tunneling probability P_{el} implies that on average the probability falls off exponentially with the distance s along the path, $P_{el} \sim e^{-2s/\xi_{el}}$, where the dimensionless localization length ξ_{el} is given by

$$\xi_{el} = \frac{2}{\ln(\bar{E}\pi/c\bar{g}\bar{\delta})}. \tag{14}$$

The numerical constant c in the above expression is equal to unity for the model of diagonal Coulomb interaction, $E_{ij}^c = E_i^c \delta_{ij}$. We show below that the inclusion of the off-diagonal part of the Coulomb interaction results in the renormalization of the constant to some value $0.5 \lesssim c < 1$.

In a full analogy with the hopping conductivity in semiconductors[22] in the case of granular metals inelastic processes are required to allow an electron to tunnel to a state with higher energy. In granular metals such processes occur via phonons as well as via the inelastic collisions with other electrons; these processes are not reflected in the formula (11). Applying the conventional Mott-Efros-Shklovskii arguments, i.e., optimizing the full hopping probability $\propto \exp[-(2s/\xi_{el}) - (e^2/\tilde{\kappa}Tas)]$, one obtains Eq. (1) for the hopping conductivity with the characteristic temperature

$$T_0 \sim e^2/a\tilde{\kappa}\xi_{el}, \tag{15}$$

where $\tilde{\kappa}$ is the effective dielectric constant of a granular sample, a is the average grain size, and ξ_{el} is given by Eq.

(14) [when deriving (15) we considered the tunneling path as nearly straight].

In the presence of the strong electric field $\mathcal{E}$, the direct application of the consideration of Ref. 30 for the case of low temperatures $T/e\xi_{el}a \ll \mathcal{E} \ll \sqrt{\delta E_0^c}/ea$ gives the nonlinear current dependence

$$j \sim j_0 e^{-(\mathcal{E}_0/\mathcal{E})^{1/2}}, \tag{16}$$

where the characteristic electric field $\mathcal{E}_0$ is given by the expression

$$\mathcal{E}_0 \sim T_0/ea\xi_{el}. \tag{17}$$

The results presented in this subsection are valid as long as the contribution of the inelastic cotunneling to the hopping conductivity can be neglected. This is the case at low temperatures and electric fields $T, \mathcal{E}ea \ll \sqrt{\delta E_0^c}$. Below we present our results for the opposite case where inelastic processes give the main contribution to hopping conductivity.

2. Hopping via inelastic cotunneling

Inelastic cotunneling is the process where the charge is transferred by means of different electrons on each elementary hop. During such process the energy brought to a grain by incoming electron differs from the energy taken by the outcoming one.

As in the case of elastic cotunneling we assume that the electron tunnels from the grain i_0 with energy ξ_0 to the grain i_{N+1} with energy ξ_{N+1}. For the probability of such a tunneling process through a chain of grains in the tunneling path we find the general expression[31]

$$P_{in} = \frac{w}{4\pi T} \frac{\bar{g}^{N+1}}{\pi^{N+1}} \left[\frac{4\pi T}{\bar{E}}\right]^{2N} \frac{\left|\Gamma\left(N + \frac{i\Delta}{2\pi T}\right)\right|^2}{\Gamma(2N)} e^{-\Delta/2T}, \tag{18}$$

where $\Gamma(x)$ is the Gamma function and $\Delta = \xi_{N+1} - \xi_0$ is the energy difference between final and initial states. The appearance of the factor $e^{-\Delta/2T}$ is consistent with the detailed balance principle. Indeed, at finite temperatures an electron can tunnel in both ways with either increase or decrease of its energy. According to the detailed balance principle the ratio of such probabilities must be $e^{\Delta/T}$; that is indeed the case for the function P_{in} (18), since apart from the factor $e^{-\Delta/2T}$ the rest of the equation is even in Δ.

To obtain the expression for the hopping conductivity one has to optimize Eq. (18) with respect to the hopping distance N under the constraint

$$N a\tilde{\kappa}\Delta/e^2 \sim 1, \tag{19}$$

following from the ES expression for the density of ground states (6). Optimization of Eq. (18) is a bit more involved procedure than the standard derivation of Mott-Efros-Shklovskii law based on the Gibbs energy distribution function (see Appendix A), however it leads to essentially the same result, i.e., the ES law with

$$T_0(T) \sim e^2/a\tilde{\kappa}\xi_{in}(T), \tag{20}$$

with the dimensionless localization length $\xi_{in}(T)$ being weakly temperature dependent

$$\xi_{in}(T) = \frac{2}{\ln[\bar{E}^2/16c\pi T^2\bar{g}]}, \tag{21}$$

where the coefficient c equals unity for a model with diagonal Coulomb interaction matrix. As in the case of the inelastic cotunneling, the long range interaction results in the reduction of the coefficient to some $1/4 \lesssim c < 1$.

At zero temperature Eq. (18) for inelastic probability P_{in} for $\Delta < 0$ is simplified to

$$P_{in}(T=0) = \frac{w2^{2N}\pi}{(2N-1)!} \frac{|\Delta|^{2N-1}}{\bar{E}^{2N}} \left(\frac{\bar{g}}{\pi}\right)^{N+1}, \tag{22}$$

while for $\Delta > 0$ Eq. (18) gives zero since an electron tunneling process with the increase of electron energy is prohibited at zero temperature.

To obtain the hopping conductivity in the regime of strong electric field $\mathcal{E}$ and low temperatures, following Ref. 30, we use the condition (19) that defines the distance to the first available electron site within the energy shell Δ together with equation

$$e\mathcal{E}r \sim \Delta, \tag{23}$$

that relates Δ with the electric field $\mathcal{E}$ and the distance between the initial and final tunneling sites $r \sim Na$. Together Eqs. (19) and (23) define $\Delta \sim \sqrt{\mathcal{E}e^3/\tilde{\kappa}}$ and $N \sim \sqrt{e/\tilde{\kappa}\mathcal{E}a^2}$ that inserted into Eq. (22) result in the expression for the current

$$j \sim j_0 e^{-(\mathcal{E}_0/\mathcal{E})^{1/2}}, \tag{24a}$$

where the characteristic electric field $\mathcal{E}_0$ is a weak function of the applied field $\mathcal{E}$

$$\mathcal{E}_0(\mathcal{E}) \sim \frac{e}{\tilde{\kappa}a^2} \ln^2[\bar{E}^2/e^2\mathcal{E}^2a^2\bar{g}]. \tag{24b}$$

III. THE MODEL

Above presented results were derived within the following model: We consider a d-dimensional array of metallic grains. The electron motion within each grain is diffusive and the grains are assumed to be weakly coupled such that the intergranular tunneling conductance g is much less than the intragranule one. The system is described by the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_c + \sum_{ij} t_{k_i k_j}^{ij} \hat{\psi}_i^\dagger(k_i)\hat{\psi}_j(k_j), \tag{25a}$$

where $t_{k_i k_j}^{ij}$ is the matrix element representing the electron tunneling amplitude from the state k_i of the grain i to the state k_j of the neighboring grain j. The Hamiltonian $\hat{H}_0$ in the right-hand side (RHS) of Eq. (25a) describes noninteracting isolated disordered grains. The term $\hat{H}_c$ includes the electron Coulomb interaction and the local external electrostatic potential μ_i on each grain

$$\hat{H}_c = \sum_{ij} \hat{n}_i E_{ij}^c \hat{n}_j + \sum_i \mu_i \hat{n}_i, \tag{25b}$$

where E_{ij}^c is the matrix of Coulomb interaction and $\hat{n}_i$ is the operator of electron number in the ith grain. The Coulomb interaction matrix is related to the capacitance matrix in the usual way

$$E_{ij}^c = (e^2/2)C_{ij}^{-1}. \tag{26}$$

As we discussed in the Introduction the random potential μ_i that models the electrostatic disorder lifts the Coulomb blockade on some grains.

We let the tunneling elements be random Gaussian variables defined by their correlators:

$$\langle t_{k_1,k_2}^{*ij} t_{k_1',k_2'}^{ij} \rangle = t_{ij}^2 \delta_{k_1 k_1'} \delta_{k_2 k_2'}, \tag{27}$$

where i,j are the coordinates of the nearest neighbor grains. The intergranular conductance between the grains i and j is related to average matrix elements as

$$g_{ij} = 2\pi \nu_i \nu_j t_{ij}^2, \tag{28}$$

where ν_i is the density of states in the grain i. The conductance g_{ij} is defined per one spin component, such that, for example, the high temperature (Drude) conductivity of a periodic granular sample is $\sigma = 2e^2 g a^{2-d}$, with a being the size of the grain. In the following we will consider both cases of periodic (in particular, it means that $\mu_i = 0$, $\nu_i = \nu_0$, $g_{ij} = g$) and irregular arrays beginning with the former case.

IV. MOTT GAP AT SMALL TUNNELING CONDUCTANCES

In the limit of small tunneling conductances, $g \ll 1$, and zero temperature a *periodic* granular array is a charge insulator. The key characteristic of the insulating state is the Mott gap Δ_M in the electron excitation spectrum that can be defined as

$$\Delta_M = E_{N=1} - \mu - E_{N=0}, \tag{29}$$

where μ is the chemical potential, $E_{N=0}$ is the ground state energy of the charge neutral array and $E_{N=1}$ is the minimal energy of the system with an extra electron added to the neutral state. In the limit of vanishing tunneling conductance, $g \rightarrow 0$, the Mott gap Δ_M coincides with the Coulomb energy $E_0^c \equiv E_{i,i}^c$ in a single grain

$$\Delta_{Mg \rightarrow 0} = E_0^c. \tag{30}$$

At finite intergranular coupling the Mott gap is reduced as a result of the virtual presence of an electron on the neighboring grains. In the case of weak coupling, $g \ll 1$, the resulting suppression of the Mott gap can be easily calculated within the perturbation theory. The first nonvanishing correction appears only in the second order in tunneling matrix elements and can be found with the help of the textbook formula[32]
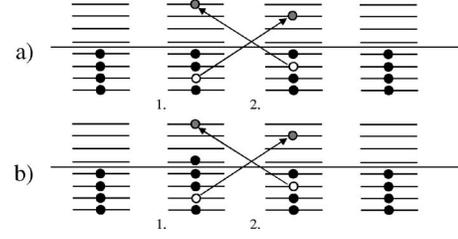

FIG. 2. This figure illustrates the virtual electron tunneling processes between neighboring grains that contribute to the reduction of the Mott gap. The tunneling processes shown on (a) determine the correction to the ground state energy of the charge neutral state $E_{N=0}$. The correction to the energy of the ground state with one extra electron (placed on grain 1) $E_{N=1}$ is illustrated in (b).

$$\Delta E = -\sum_k \frac{|V_{k,0}|^2}{E_k - E_0}, \tag{31}$$

where ΔE is the energy correction to the ground state and the matrix elements of the perturbation V are taken between the ground state 0 and excited states k.

Correction due to finite intergranular coupling should be included in both terms $E_{N=1}$ and $E_{N=0}$ in Eq. (29). The ground state with one extra electron $N=1$ in the limit of zero coupling is degenerate since the extra electron can be present on any grain. Further, we will assume that it occupies grain 1 as shown in Fig. 2(b). Since we want to find the difference between the corrections to $N=1$ and $N=0$ ground states, it is clear that within the second order perturbation theory in tunneling elements t_{k_1,k_2} only electron hops between grain 1 and its neighbors result in nonzero contribution. All other possible hops give rise to equal corrections to the energies $E_{N=1}$ and $E_{N=0}$ that are mutually cancelled in Eq. (29). Further, since all the neighbors of grain 1 are equivalent, we will consider only electron hops between grain 1 and one of its neighbors that we denote as grain 2. Contribution of the hops between grain 1 and all its other neighbors will simply result in the factor z (coordination number) in the final answer.

Now let us consider the correction to the ground state energy $E_{N=0}$. In this case the matrix elements $V_{k,0}$ correspond to the processes of electron tunneling between the neutral neighboring grains. Consider the tunneling process from grain 1 to grain 2 shown in Fig. 2(a): The excitation energy of this process is $\varepsilon_{k_1} + \varepsilon_{k_2} + E_{eh}$, where ε_{k_2} is the bare (with no Coulomb energy included) energy of an electron excitation in grain 2, ε_{k_1} is the bare energy of a hole excitation in grain 1, and E_{eh} is the Coulomb part of the energy of the electron-hole excitation $E_{eh} = E_{11}^c + E_{22}^c - 2E_{12}^c$ that for the periodic case under consideration is reduced to

$$E_{eh} = 2E_0^c - 2E_{12}^c. \tag{32}$$

The energy correction corresponding to such process is

$$-\Delta E_{N=0} = 2\sum_{k_1,k_2} \frac{|t_{k_1,k_2}|^2}{\varepsilon_{k_1} + \varepsilon_{k_2} + E_{eh}}, \tag{33}$$

where the factor 2 takes into account the similar process of electron hopping from grain 2 to grain 1.

Analogously we find the correction to the energy $E_{N=1}$ as illustrated on Fig. 2(b): The excitation energy of the process of electron tunneling from grain 2 to grain 1 is $\varepsilon_{k1}+\varepsilon_{k_2}+3E_{11}^c+E_{22}^c-4E_{12}^c=\varepsilon_{k_1}+\varepsilon_{k_2}+2E_{eh}$, while the excitation energy of the process of electron tunneling from grain 1 to grain 2 is $\varepsilon_{k_1}+\varepsilon_{k_2}+E_{11}^c-E_{22}^c=\varepsilon_{k_1}+\varepsilon_{k_2}$. The corresponding correction to the energy $E_{N=1}$ reads

$$-\Delta E_{N=1}=\sum_{k_1,k_2}\frac{|t_{k_1,k_2}|^2}{\varepsilon_{k_1}+\varepsilon_{k_2}+2E_{eh}}+\frac{|t_{k_1,k_2}|^2}{\varepsilon_{k_1}+\varepsilon_{k_2}}. \quad (34)$$

One can see that corrections (33) and (34) are ultraviolet divergent. However, their difference must be finite since it represents the measurable quantity—correction to the Mott gap. This is, indeed, the case: Subtracting Eq. (33) from Eq. (34), presenting summation over states as integrals $\varepsilon_{k_1(k_2)}\rightarrow\varepsilon_{1(2)}$ and introducing further the variable $\varepsilon=\varepsilon_1+\varepsilon_2$ we obtain

$$\Delta_M-E_0^c=z\nu_0^2t^2\int_0^\infty\varepsilon\,d\varepsilon\left[\frac{1}{\varepsilon+2E_{eh}}-\frac{2}{\varepsilon+E_{eh}}+\frac{1}{\varepsilon}\right],$$

where ν_0 is the density of states in a single grain and t^2 defined by Eq. (27) appears from the averaging over matrix elements. Taking the integral in the above expression and multiplying by the factor 2 coming from the spin degeneracy we obtain the result (2).

In the derivation of the reduction of the Mott gap presented in this section we neglected the fact that an extra electron added to the neutral system in the presence of finite intergranular coupling can move diffusively over the sample. Contributions corresponding to such processes, as we will see below, are suppressed by an extra small factor $\delta/E_0^c\ll1$ and, thus, can be neglected.

V. PHASE AND CHARGE REPRESENTATIONS

The perturbation theory presented in Sec. IV is limited to the regime of low tunneling conductance $g\ll1$. Moreover, the high order tunneling processes that are very important for understanding the physics of the Mott transition and hopping conductivity cannot be included in this way. For this reason one has to seek a convenient approach applicable at intermediate tunneling conductances and capable of description of the effects of high order tunneling processes. In view of this, a useful tool in approaching this problem is the phase functional approach[17] that allows further mapping of the system onto the Coulomb gas of charged particles in $d+1$ dimensions.[18,33]

A. Phase representation

The phase functional approach is based on the decoupling of the Coulomb interaction term (25b) in the Hamiltonian (25a) by means of the potential field $V(\tau)$ that is further absorbed into the gauge transformation of the fermionic fields

$$\psi_i(\tau)\rightarrow\psi_i(\tau)e^{-i\phi_i(\tau)},\quad\phi_i(\tau)=\int_{-\infty}^\tau V_i(\tau')d\tau'. \quad (35)$$

Since the action of an isolated grain is gauge invariant the phases $\phi_i(\tau)$ enter the whole Lagrangian of the system only through the tunneling matrix elements

$$t_{ij}\rightarrow t_{ij}e^{i\phi_{ij}(\tau)}, \quad (36)$$

where $\phi_{ij}(\tau)=\phi_i(\tau)-\phi_j(\tau)$ is the phase difference of the ith and jth grains.

A simple description in terms of phase variables can be obtained via integrating out the fermionic fields and expanding the resulting rather complicated action to the lowest non-vanishing order in tunneling elements. As a result one obtains the Ambegaokar-Eckern-Schön (AES) action[17]

$$S_{AES}=-\frac{1}{2e^2}\sum_{ij}\int d\tau(\dot{\phi}_i+i\mu_i)C_{ij}(\dot{\phi}_j+i\mu_j)+S_t[\phi], \quad (37)$$

where the tunneling term S_t at zero temperature is

$$S_t[\phi]=\frac{1}{2\pi}\sum_{\langle ij\rangle}g_{ij}\int_{-\infty}^{+\infty}d\tau_1d\tau_2\frac{e^{i\phi_{ij}(\tau_1)-i\phi_{ij}(\tau_2)}}{(\tau_1-\tau_2)^2}, \quad (37b)$$

where the angular brackets mean summation over the nearest neighbors. The higher order terms in the expansion of the total action in the tunneling matrix elements cannot always be neglected and, in general, must be taken into account along with the nearest neighbor action (37). However, prior to considering these more complicated higher order terms, it is instructive to analyze, first, the action (37) in detail. Moreover, as we will show, even at $T=0$ there is a regime where the AES action is applicable. We, thus, proceed to the description of the charge mapping beginning with the AES action, and later, in Sec. VIII, we will show how the higher order tunneling terms can be included in the mapping. Also, for simplicity, we will first consider the case of the regular periodic array.

B. Charge representation

Proceeding to the mapping of the action (37) onto the model of the classical Coulomb gas and using the method of Ref. 18, we first expand the partition function Z to all orders in the tunneling part of the action S_t

$$Z=Z_0\sum_{N=1}^\infty\langle S_t^N[\phi]\rangle/N!. \quad (38)$$

Here Z_0 is the partition function of a system of isolated grains and the angle brackets assume averaging with respect to the Coulomb phase action that for the case $\mu_i=0$ under consideration reads

$$S_c=-\frac{1}{2e^2}\sum_{ij}\int d\tau\dot{\phi}_iC_{ij}\dot{\phi}_j. \quad (39)$$

On the next step, for each term in the expansion (38) we write the Nth power of the tunneling action S_t explicitly as the product
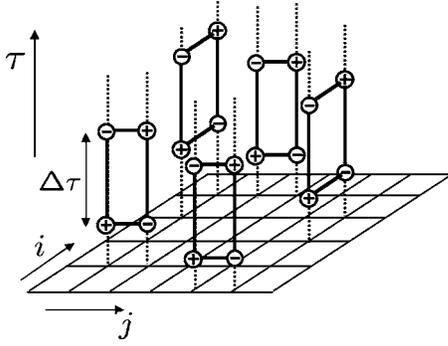
FIG. 3. Illustration of the mapping of a quantum granular array model in the nearest neighbor hopping approximation onto the classical model of gas of quadrupoles in $d+1$ dimensions.

$$S_t^N[\phi] = \prod_{n=1}^{N} S_t^{(n)}, \qquad (40a)$$

where the index n of the tunneling action S_t reminds one that the time integrations in each term in the above product are independent

$$S_t^{(n)} = \frac{g}{2\pi} \sum_{\langle ij \rangle} \int_{-\infty}^{+\infty} \frac{d\tau_n d\tau_n'}{(\tau_n - \tau_n')^2} e^{i\phi_{ij}(\tau_n) - i\phi_{ij}(\tau_n')}. \qquad (40b)$$

Writing the powers of the tunneling action as products (40a) allows one to implement the integrations over the phase fields $\phi_i(\tau)$ exactly for each term in the expansion (38) with the help of the formula

$$\langle e^{i\Sigma_n \phi_i(\tau_n) e_n} \rangle = e^{-U^c}, \qquad (41)$$

where U^c given by the expression

$$U^c = \frac{1}{2} \sum_{n_1, n_2} E_{i_{n_1} i_{n_2}}^c |\tau_{n_1} - \tau_{n_2}| e_{n_1} e_{n_2}, \qquad (42)$$

can be viewed as a Coulomb energy of a system of classical charges in a space with an extra time dimension.[34]

The classical charges $e_n = \pm 1$ interact via the 1D Coulomb potential along the time direction while the strength of the interaction is given by the "real" Coulomb energy of original quantum electrons E_{ij}^c.

From Eq. (40b) we see that each tunneling term $S_t^{(n)}[\phi]$ contains the sum of four phase fields in the exponent

$$i\phi_i(\tau_n) - i\phi_j(\tau_n) - i\phi_i(\tau_n') + i\phi_j(\tau_n'). \qquad (43)$$

Each of these fields, according to Eqs. (41) and (42), gives rise to a classical charge located at coordinates corresponding to the arguments of the phase fields. This means that the classical charges appear as quadrupoles that occupy only the neighboring sites as it is shown in Fig. 3.

The time dependent kernel in Eq. (40b) can be written in the form

$$\frac{g}{2\pi} \frac{1}{(\tau_n - \tau_n')^2} = e^{-\mathcal{U}_n^q(\tau_n - \tau_n')}, \qquad (44)$$

such that $\mathcal{U}_n^q$ defines the "internal" energy of nth quadrupole

$$\mathcal{U}_n^q = \ln[2\pi(\Delta\tau_n)^2/g], \qquad (45)$$

with $\Delta\tau_n$ being the size of the nth quadrupole in the τ direction as shown in Fig. 3.

Thus the partition function, Eq. (38), of the quantum problem under consideration in the nearest neighbor electron tunneling approximation can be presented as a partition function of classical charges that appear in quadrupoles, all charges being subject to the Coulomb interaction (42). The total energy of the system of N quadrupoles can be written as

$$U^q = \sum_{n=1}^{N} \mathcal{U}_n^q + U_{4N}^c, \qquad (46)$$

where $\mathcal{U}_n^q$ is the internal energy of the nth quadrupole and U_{4N}^c is the Coulomb energy of $4N$ charges that form N quadrupoles defined by Eq. (42). We notice that different quadrupoles interact with each other only via the Coulomb part of the energy U^c.

In the following sections we will consider the thermodynamic properties of the classical model and set correspondence between the parameters of the original quantum model and the effective classical one.

VI. THERMODYNAMICS OF THE CLASSICAL MODEL IN THE LIMIT OF LOW QUADRUPOLE DENSITY

The thermodynamic potential Ω of the classical model (46) can be found in the limit of the low quadrupole density that, as we will show below, corresponds to the case of the low tunneling conductance $g \ll 1$ in the quantum model. In this limit one can neglect the mutual interaction of different quadrupoles considering them as free independent objects and the thermodynamic potential Ω of the whole system is reduced to the sum of the free energy $\mathcal{F}_0$ of a single quadrupole multiplied by the number of the quadrupoles N and the entropy S of the quadrupole configuration space

$$\Omega = N\mathcal{F}_0 - S. \qquad (47)$$

The entropy of N independent quadrupoles is given by the expression

$$S = N \ln V - \ln N!, \qquad (48)$$

where $V = LM$ is the volume of the $(d+1)$-dimensional system with L being the size of the system in τ direction and M being the total number of the sites (grains) in the array. The term $\ln N!$ in Eq. (48) appears from the term $1/N!$ in the expansion (38). In the classical statistical mechanics such a term appears due to equivalence of the classical configurations that differ only by mutual permutations of particles. Thus, in the thermodynamic limit $N \rightarrow \infty$ for the thermodynamic potential Ω in the case of low quadrupole density, we obtain the expression

$$\Omega = N\mathcal{F}_0 + N \ln \rho - N, \qquad (49)$$

where $\rho = N/V$ is the density of quadrupoles in $d+1$ dimensions.

The partition function of a single quadrupole is given by the expression

$$\mathcal{Z}_0 = \frac{g}{\pi} \int_0^\infty \frac{d\tau}{\tau^2} e^{-E_{eh}\tau}, \tag{50}$$

where the energy of the electron-hole excitation is defined by Eq. (32). Equation (50) is divergent on the low integration limit and in order to regularize it we introduce the time cutoff λ obtaining $\mathcal{Z}_0 = g/\pi\lambda$ so that the free energy of a single quadrupole $\mathcal{F}_0 = -\ln \mathcal{Z}_0$ becomes

$$\mathcal{F}_0 = -\ln(g/\pi\lambda). \tag{51}$$

A. Density of classical charges

The thermodynamic potential given by Eqs. (49) and (51) determines all the thermodynamic quantities of the classical model in the limit of low quadrupole density. In particular, the density itself can be obtained via minimization of Ω with respect to the number of quadrupoles N resulting in

$$\rho = g/\pi\lambda. \tag{52}$$

We would like to note that the obtained density coincides with the single quadrupole partition function $\rho = \mathcal{Z}_0$. The result for the density ρ (52) can be better understood considering a discrete time axis. In this case the cutoff distance λ would be the distance between the neighboring sites in τ direction and the dimensionless density according to Eq. (52) would be g/π. Thus, it becomes clear that the limit of low conductance $g \ll 1$ corresponds to the limit of the low occupation number of classical charges.

B. Relation between the Mott gap and dielectric constant of classical gas

In order to make the developed mapping useful one has to express the physical observables of the original quantum model in terms of the corresponding thermodynamic quantities of the classical model. Below we show that the key characteristic of the insulating state, the Mott gap Δ_M, is related to the dielectric constant κ_τ of the effective classical model measured in the direction of τ axis as

$$\Delta_M = E_0^c/\kappa_\tau. \tag{53}$$

To prove relation (53) we first note that the gap Δ_M in the electron spectrum determines the asymptotic behavior $(\tau \to \infty)$ of the single electron Green function

$$G_{ii}(\tau) \sim e^{-\Delta_M |\tau|}. \tag{54}$$

At the same time the asymptotic behavior of the Green function $G_{ii}(\tau)$ with exponential accuracy coincides with the asymptotic behavior of the phase correlation function

$$G_{ii}(\tau_1 - \tau_2) \sim \Phi(\tau_1 - \tau_2) \equiv \langle e^{i\phi_i(\tau_1) - i\phi_i(\tau_2)} \rangle, \tag{55}$$

where the angular brackets stand for averaging with respect to the phase functional (37), written explicitly

$$\Phi(\tau_1 - \tau_2) = Z^{-1} \int D[\phi] e^{i\phi_i(\tau_1) - i\phi_i(\tau_2)} e^{S[\phi]}. \tag{56}$$

All the considerations that were previously used in order to derive the effective model (46) can be applied to the corre-
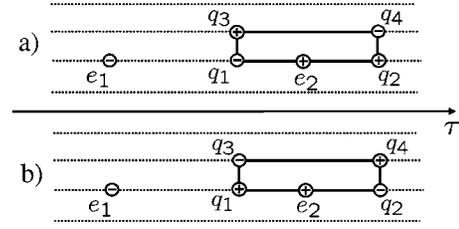

FIG. 4. An electric field that charge e_2 "feels" consists of the field created by the charge e_1 and the fields created by the charges of the quadrupole $q_1, \ldots, q_4$.

lation function (56) as well. Repeating all the necessary steps we notice that phases $\phi_i(\tau_1)$ and $\phi_i(\tau_2)$ that enter the exponent in Eq. (56) result in appearance of two external charges of opposite signs located on the site i at time coordinates τ_1 and τ_2. Further, we see that the phase correlation function Φ is related to the thermodynamic potential of the system with two external charges $\Omega_2(\tau_1 - \tau_2)$ as

$$\Phi(\tau) = e^{-\Omega_2(\tau)}. \tag{57}$$

The function $\Omega_2(\tau)$ being normalized such that $\Omega_2(0) = 0$ is nothing but the effective interaction of two external charges in the presence of quadrupoles. The initial electrostatic interaction is reduced in the presence of quadrupoles and the asymptotic behavior of the function Ω_2 takes the form

$$\Omega_2(\tau) = |\tau| E_0^c/\kappa_\tau, \quad \tau \to \infty, \tag{58}$$

where κ_τ is the dielectric constant. This proves the relation between the Mott gap and dielectric constant of the classical Coulomb gas in τ direction (53).

C. Mott gap in the limit of $gz \ll 1$

Finally, in order to better understand the relationship between the original quantum model and the effective classical one we will rederive the result (2) for the Mott gap in the case of small tunneling conductance $gz \ll 1$ within the framework of the effective classical model.

Let us consider two charges e_1 and e_2 (see Fig. 4) of opposite signs located on the site i at time coordinates τ_1 and τ_2, respectively. In the absence of surrounding quadrupoles their interaction energy would be $E_0^c|\tau_1 - \tau_2|$ that gives the zero-order value of the Mott gap $\Delta_M = E_0^c$. The bare interaction energy E_0^c of two charges is reduced due to the presence of quadrupoles. In order to find this reduction it is convenient to consider the electric field created by the charge 1 and partially screened by the surrounding quadrupoles at the point τ_2 of location of charge 2. In the limit of low quadrupole density the screening of the interaction due to different quadrupoles can be considered independently.

It is easy to see that the electric field at the point τ_2 is affected by the presence of the quadrupole only if the latter is placed such that charge e_2 is inside the quadrupole as it is shown in Fig. 4. Depending on the arrangements of quadrupole charges one can distinguish two situations shown in Figs. 4(a) and 4(b). In case (a) the electric field that charge e_2 "feels" is

$$E^{(a)} = -3E_0^c + 2E_{12}^c = -E_0^c - E_{eh}, \qquad (59a)$$

while in case (b) it is

$$E^{(b)} = E_0^c - 2E_{12}^c = -E_0^c + E_{eh}. \qquad (59b)$$

We see from Eqs. (59a) and (59b) that the electric field at the point τ_2 of location of charge 2 gets an additional contribution $\pm E_{eh}$ depending on the arrangements of quadrupole charges. The probability of the configuration (a) is given by the Gibbs weight as

$$P_a(\tau_{q_1}, \tau_{q_2}) = \mathcal{Z}_0^{-1} e^{-\mathcal{U}^q(\tau_{q_2} - \tau_{q_1}) - E_{eh}(\tau_{q_2} - \tau_{q_1}) + \tau_{q_1} E_{eh}},$$

where the time coordinates of quadrupole charges τ_{q_1} and τ_{q_2} are taken with respect to coordinate τ_2 where charge e_2 is located such that $\tau_{q_2} > 0$ and $\tau_{q_1} < 0$. Analogously we find the probability of the configuration (b)

$$P_b(\tau_{q_1}, \tau_{q_2}) = \mathcal{Z}_0^{-1} e^{-\mathcal{U}^q(\tau_{q_2} - \tau_{q_1}) - E_{eh}(\tau_{q_2} - \tau_{q_1}) - \tau_{q_1} E_{eh}}.$$

The probabilities P_a and P_b must be further multiplied by the quadrupole density ρ, coordination number z, and combinatoric factor 2. Multiplication by ρ simply cancels the factor $\mathcal{Z}_0^{-1}$ due to the relation $\rho = \mathcal{Z}_0$ that we pointed out below Eq. (52). The additional contribution to the electric field at coordinate τ_2 of charge 2 due to the presence of the quadrupoles becomes

$$\Delta E = 2z E_{eh} \int_{-\infty}^{0} d\tau_{q_1} \int_{0}^{\infty} d\tau_{q_2} (P_b - P_a). \qquad (60)$$

The dielectric constant κ_τ is given by the expression $\kappa_\tau^{-1} = (E_0^c - \Delta E)/E_0^c$. Taking integrals in Eq. (60) we finally get the dielectric constant

$$\kappa_\tau^{-1} = 1 - \frac{2gz}{\pi} \frac{E_{eh}}{E_0^c} \ln 2, \qquad (61)$$

that taking into account relation (53) between the Mott gap and dielectric constant is indeed consistent with result (2) previously obtained within the direct perturbation theory in tunneling matrix elements.

VII. MOTT GAP AT LARGE TUNNELING CONDUCTANCES

In this section we continue our investigation of the quadrupole model (46) now coming to the regime of moderately large tunneling conductance $gz \gtrsim 1$. We show that within this model the Mott gap Δ_M is exponentially reduced with growth of g but nevertheless it always remains finite and, thus, the quantum system represented by the model (46) is always an insulator at $T=0$ even at arbitrary large tunneling conductance g. This statement, however, turns out to be an artifact of neglecting the high order tunneling terms in derivation of the phase action (37) and the model (46) following it. Below we show that models (37) and (46) are applicable as long as the Mott gap Δ_M is larger than the inverse escape time of an electron from a grain $\Gamma = g\delta$. Thus the results of this section are applicable as long as the obtained Mott gap Δ_M is larger than Γ.
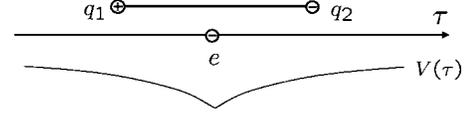

FIG. 5. Screening of the external charge e in the dipole model. The Coulomb potential of the charge e is screened by the dipole charges q_1 and q_2 resulting in the effective screened potential $V(\tau)$.

A. Mean field approximation

The model (46) can hardly be solved in a general case and further in this section we make two assumptions that will allow us to construct a mean field approach.[36] Namely, (i) we take the diagonal matrix of Coulomb interaction $E_{ij}^c = E_0^c \delta_{ij}$ and (ii) we assume that the coordination number z is large. These two assumptions allow us to reduce the quadrupole problem in $d+1$ dimensions to the problem of one dimensional dipoles with self consistent dipole interaction.

To construct a mean field theory let us pick up a site i and consider the influence of the neighboring sites on it in a mean field manner. The quadrupole that occupies the site i and a neighboring site j is reduced to the dipole located on the site i. The mutual interaction of two charges of the resulting dipole consists of the quadrupole interaction $\mathcal{U}^q$ and of the Coulomb interaction with all other charges located at the site j. The latter must be taken into account in a self-consistent way. The mutual influence of the neighboring sites due to direct Coulomb interaction is absent due to our choice of the diagonal Coulomb matrix E_{ij}^c. Thus the internal dipole interaction energy becomes

$$\mathcal{U}_d = \ln[\pi(\Delta\tau_n)^2/gz] + \Omega_2(\Delta\tau_n), \qquad (62)$$

where the interaction energy $\Omega_2(\tau)$ of two external charges must be calculated self consistently in terms of the dipole model. The asymptotic behavior of this function at the same time according to Eq. (58) determines the dielectric constant κ_τ that is directly related to the Mott gap Δ_M of the original model.

Thus we arrive at the simplified one dimensional model of dipoles bounded by the internal interaction (62). Apart from the bounding interaction all charges are subject to the 1D Coulomb interaction

$$U_{1d}^c = \frac{E_0^c}{2} \sum_{n_1,n_2} |\tau_{n_1} - \tau_{n_2}| e_{n_1} e_{n_2}. \qquad (63)$$

In the limit of large density $\rho \gg 1$ that corresponds to the case of large tunneling conductance $g \gg 1$ the Coulomb interaction gets sufficiently screened. For this reason the expression for the density (52) as well as the identity $\mathcal{Z}_0 = \rho_0$ are applicable in the dense limit too. In the dense limit the screening effects can usually be well described within the Debye theory.[35]

Consider the charge e placed at $\tau=0$ shown in Fig. 5. The Coulomb potential of this charge screened by the dipole charges obeys the Poisson equation

$$\partial_\tau^2 V(\tau) = -2E_0^c(q(\tau) - \delta(\tau)), \qquad (64)$$

where $q(\tau)$ is the average density of all dipole charges and $\delta(\tau)$ is the delta function. At the same time the density of dipole charges is given by the Gibbs distribution: The contribution of two charges q_1 and q_2 of a single dipole (located at coordinates τ_1 and τ_2, respectively) to the charge density $q(\tau)$ is

$$\mathcal{Z}_0^{-1}[\delta(\tau - \tau_1) - \delta(\tau - \tau_2)]e^{-\tilde{\mathcal{U}}_d(\tau_1 - \tau_2) - V(\tau_1) + V(\tau_2)}, \qquad (65)$$

where the energy $\tilde{\mathcal{U}}_d(\tau)$ that enters the Gibbs exponent includes the dipole interaction energy $\mathcal{U}_d(\tau)$ and the screened Coulomb interaction of charges q_1 and q_2

$$\tilde{\mathcal{U}}_d(\tau) = \mathcal{U}_d(\tau) + \Omega_2(\tau). \qquad (66)$$

We note that within the approximation that we use the screened potential $V(\tau)$ is nothing but the effective screened interaction of two charges $\Omega_2(\tau)$ so that

$$V(\tau) = \Omega_2(\tau). \qquad (67)$$

Multiplying Eq. (65) by the dipole density ρ, making use of the identity $\mathcal{Z}_0 = \rho$ and integrating over the coordinates τ_1, τ_2, we obtain the averaged charge density

$$q(\tau) = \frac{2gz}{\pi} \int_{-\infty}^{\infty} d\tau' \frac{e^{-2V(\tau - \tau')}}{(\tau - \tau')^2} \sinh[V(\tau') - V(\tau)]. \qquad (68)$$

Equations (64) and (68) form a closed system of equations that determine the potential $V(\tau)$. In the next subsection we present the solution of these equations.

B. Solution of the mean field equations

Numerical solution of Eqs. (64) and (68) (see Appendix B) for intermediate values of gz shows that the external charge is never screened completely and the long time asymptotic behavior of the function $V(\tau)$ in Eq. (67) has a linear form

$$V(\tau) = |\tau| E_0^c \kappa_\tau^{-1}, \quad \tau \to \infty, \qquad (69)$$

where the inverse dielectric constant κ_τ^{-1} becomes very small already at $gz \sim 2$. At larger values of gz the numerical investigation becomes difficult since the crossover time scale at which the potential turns into the insulating behavior (69) grows exponentially with gz (see below). Fortunately, at large gz Eqs. (64) and (68) can be considered within an analytic approach: To find the potential $V(\tau)$ on intermediate times where nonlinear effects are not important we consider Eq. (68) in the linear approximation with respect to $V(\tau)$. As a result in the Fourier representation we obtain

$$q_\omega = -2zg|\omega|V_\omega. \qquad (70)$$

Together with the Poisson equation (64) written in the Fourier representation $\omega^2 V(\omega) = 2E_0^c(q_\omega - 1)$ Eq. (70) defines the potential

$$V_\omega = -\frac{1}{\omega^2/2E_0^c + 2zg|\omega|}, \qquad (71)$$

that being transformed to the time representation for $\tau > (E_0^c g)^{-1}$ results in

$$V(\tau) = \frac{\ln(\tau E_0^c g)}{2\pi z g}. \qquad (72)$$

The linear approximation that we used in Eq. (68) is valid as long as the potential $V(\tau)$ is small. This assumes that dependence (72) is applicable for not very long times $\tau \ll \exp(\pi z g)/g E_0^c$. The logarithmic form of the answer (72) assumes that the change in the charging energy

$$E_0^c \to \xi E_0^c, \qquad (73)$$

in Eq. (72) results in a constant shift of the potential $V(\tau) \to V(\tau) + (1/2\pi z g)\ln \xi$. At the same time the only term in Eq. (68) that is sensitive to this shift is the potential $V(\tau)$ in the exponent that provides the renormalization of conductance

$$g \to g - (1/\pi z)\ln \xi. \qquad (74)$$

This consideration allows us to solve the system of Eqs. (64) and (68) for the case of large tunneling conductance in the renormalization group(RG) scheme. Since the charging energy E_0^c is the only energy scale in the problem, the scaling of this energy with conductance determines the scaling of the Mott gap. Thus, from the scaling Eqs. (73) and (74), we finally obtain the expression for the Mott gap (3).

Equation (3) agrees with the result of RG approach of Ref. 6 where the opening of the Mott gap Δ_M was assumed to happen at the point where the tunneling conductance running under the RG scheme flowed to values of the order one. We would like to note that the present approach does not rely on any assumptions of this kind: we first apply the RG scheme to reduce the conductance to values $gz \gtrsim 1$, and further we solve the equation numerically and show that opening of the gap indeed takes place, i.e., the asymptotics of the potential $V(\tau)$ has a linear form indeed.

VIII. INCLUSION OF HIGHER ORDER TUNNELING TERMS

So far we considered the quadrupole model (46) that includes only the nearest neighbor electron tunnelings. In the present section we describe how the high order tunneling terms can be included in our mapping scheme and in the following sections we apply this technique to the problem of hopping conductivity and Mott transition in granular arrays.

Keeping in mind the applications to the irregular arrays, as in the case of hopping conductivity, in this section we include consideration of the random on-site potential μ_i. Also we take into account variations of tunneling conductance g from grain to grain as well as variations of charging energies E_i^c and the mean energy level spacings δ_i.

In order to construct the effective classical model that takes into account higher order tunneling terms we proceed in the same way as in the derivation of AES functional Eqs.
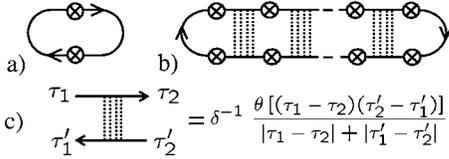
FIG. 6. Diagram (a) represents the lowest order correction in tunneling conductance g to the partition function, the first term in the expansion (38). The solid lines denote the propagator of electrons; the crossed circles stand for the phase dependent tunneling matrix elements $t_{ij}e^{i\phi_{ij}(\tau)}$. Diagram (b) describes a higher order correction in tunneling matrix elements t_{ij}. The dotted lines represent impurity scattering. Diagram (c) shows the zero dimensional diffusion in time representation.

(37). First we decouple the Coulomb interaction term in Eq. (25a) with the help of potential field $V(\tau)$

$$\mathcal{L}_c = -\frac{1}{2e^2}\sum_{ij}[V_i + i\mu_i]C_{ij}[V_j + i\mu_j] - i\sum_i n_i V_i, \quad (75)$$

where $n_i = n_i(\tau)$ is the electron density field. The term $-in_i V_i$ can be absorbed by the fermion gauge transformation given by Eqs. (35) such that the action that governs dynamics of the potential field $V(\tau)$ becomes

$$\mathcal{L}_c = -\frac{1}{2e^2}\sum_{ij}[V_i + i\mu_i]C_{ij}[V_j + i\mu_j]. \quad (76)$$

Further, following Ref. 37 we proceed with diagrammatic expansion of the partition function Z in tunneling matrix elements $t_{ij}e^{i\phi_{ij}(\tau)}$. In the lowest order we reproduce the results obtained within the phase functional. Indeed, the lowest order correction to the partition function Z shown in Fig. 6(a) simply results in the first term in the expansion (38). All disconnected diagrams of the type shown in Fig. 6(a) must also be included in the partition function; this reproduces all high order terms in the expansion (38). Integration over the phase fields $\phi_i(\tau)$ can be implemented exactly as it was described in Sec. V B resulting in the appearance of classical charges in $d+1$ dimensions with the Coulomb interaction energy

$$U = \frac{1}{2}\sum_{n_1,n_2} E^c_{i_{n_1}i_{n_2}}|\tau_{n_1} - \tau_{n_2}|e_{n_1}e_{n_2} + \sum_n \mu_{i_n}\tau_n e_n, \quad (77)$$

that differs from Eq. (42) only by the presence of the random potential μ_i not previously taken into account.

Thus, inclusion of all disconnected diagrams, Fig. 6(a), and their further averaging over the phases reproduces the classical quadrupole model (46). The expression for the internal quadrupole energy for the case where conductance g varies from contact to contact is the straightforward generalization of Eq. (45)

$$\mathcal{U}_n^q = \ln[2\pi(\Delta\tau_n)^2/g_{ij}]. \quad (78)$$

The advantage of the mapping of the original quantum model on the classical electrostatic system is that it allows us to include the higher order tunneling processes shown in Fig. 6(b) in essentially the same way as the nearest neighbor hop-
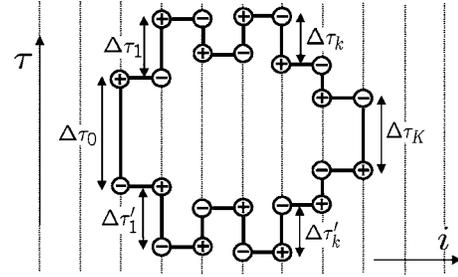

FIG. 7. Illustration of the mapping of the model of quantum electrons in a granular array onto the model of classical charges imbedded into the electron world lines forming loops. The illustration is shown for the case of $1+1$ dimensions (1D granular array); generalization to higher $d+1$ dimensions is straightforward.

pings. The higher order diagrams include the single grain diffusion propagator

$$D_0 = 2\pi\delta^{-1}/|\omega_1 - \omega_2|, \quad (79)$$

that being transformed into the time representation results in the expression shown in Fig. 6(c). Each tunneling matrix element in the diagram Fig. 6(b) includes the phase variables as $t_{ij}e^{i\phi_{ij}(\tau)}$. This allows us to present the diagram in Fig. 6(b) as a charge loop shown in Fig. 7. The charges interact electrostatically and are subject to the local potentials according to Eq. (77). The internal energy of a single Kth order loop is

$$E^{(K)} = \ln[2\pi/g_0] - \sum_{k=1}^{K-1}\ln[g_k\delta_k/2\pi]$$

$$+ \ln|\Delta\tau_0| + \ln|\Delta\tau_K| + \sum_{k=1}^{K-1}\ln(|\Delta\tau_k| + |\Delta\tau'_k|), \quad (80)$$

where the time intervals $\Delta\tau_k$ and $\Delta\tau'_k$ shown in Fig. 7 for each site k have opposite signs, i.e., $\Delta\tau_k\Delta\tau'_k < 0$.

We mapped the model of quantum electrons in a metallic array with Coulomb interaction onto the classical model of Coulomb charges bounded by the electron world lines and forming loops. First order loops that form quadrupoles represent the process of virtual electron tunnelings to the neighboring grains and back. Higher order loops represent electron motion on larger distances.

In the Mott insulator state electrons cannot move on large distances. This assumes that in the classical representation of the Mott insulator state the thermodynamics of the classical charges will be dominated by the contribution of the short loops, while high order loops will be energetically suppressed.

In the metallic state, on the contrary, diffusive motion of electrons on any distances is possible meaning that in the classical model the loops of infinite order are present. It is clear that the insulator to metal transition in terms of the classical model corresponds to the transition at which the infinite loops that were suppressed in the insulating state begin to proliferate. Infinite loops represent almost free charges that screen the Coulomb interaction. The inverse dielectric constant κ_τ^{-1} of this state that corresponds to the absences of the Mott gap Δ_M is zero.

The detailed description of the Mott transition is difficult since it requires the analysis of contribution of high order loops in the regime where the tunneling conductance g is not small and the Mott gap Δ_M is already strongly reduced due to the screening by quadrupoles according to Eq. (3). Such description, on the qualitative level, is performed in Sec. X, while in the next section we will study the high order electron tunnelings in a more simple regime of low tunneling conductance. As we discussed above, the high order tunnelings being not important for the thermodynamics in the regime of small g are crucial for the description of the hopping transport in the case of irregular array.

IX. HOPPING CONDUCTIVITY IN GRANULAR METALS

In Sec. VIII it was shown that the periodic granular array with low tunneling conductances g is the Mott insulator with the hard gap Δ_M in the electron excitation spectrum. This assumes the activation conductivity behavior, $\sigma \sim \exp(-\Delta_M / T)$, and excludes the possibility of hopping transport. However, as we pointed out in the Introduction, in realistic granular arrays there are many reasons for electrostatic disorder that compensate the Coulomb blockade on some sites. This gives rise to the finite density of states on the Fermi level and opens the possibility for the variable range hopping transport.

In this section we discuss the hopping conductivity in granular metals via elastic and inelastic electron cotunneling. We first concentrate on hopping via elastic cotunneling and later discuss the inelastic mechanism.

A. Hopping via elastic cotunneling

To analyze the hopping transport mechanism in granular array one has to consider the probability of electron tunneling between the two distant grains with energies close to the Fermi level. At very low temperatures such a tunneling can be realized only as a virtual tunneling of an electron through several grains being in the Coulomb blockade regime—such a process is called the elastic cotunneling.[24]

In the following we consider the probability of elastic cotunneling through several grains within the approach of an effective classical model. We assume that the electron tunneling takes place between two sites i_0 and i_{N+1} where the local Coulomb gap is absent. In our model it means that the Coulomb energy on these two sites is compensated by the external local potentials μ_i. Thus, removing the charge from the grain i_0 as well as placing it on the grain i_{N+1} does not cost any Coulomb energy. The probability of tunneling between these two states can be found as follows: In the basis of the eigenstates of a system of isolated grains the amplitude of the tunneling process from the site i_0 to the site i_{N+1} is

$$A_{i_0,i_{N+1}} = \langle 0|\hat{c}_{i_{N+1}}\hat{S}\hat{c}_{i_0}^\dagger|0\rangle, \tag{81}$$

where $\hat{S}$ is the evolution operator that written in the interaction representation takes into account only the tunneling part of the Hamiltonian (25a). Consideration of the amplitude
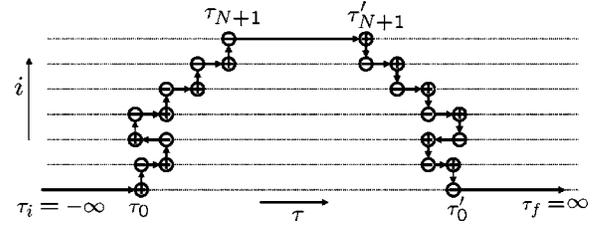

FIG. 8. The electron world line representing the probability of elastic cotunneling from the 0th to the $(N+1)$st grain.

$A_{i_0,i_{N+1}}$, however, is not convenient since the amplitude becomes zero after averaging over tunneling matrix elements t_{k_1,k_2}. This happens because the phase of the amplitude of the tunneling process is very sensitive to the particular realization of disorder. For this reason one has to consider the probability of the tunneling process $P = A^*A$. The conjugate amplitude A^* can be written as the probability of the inverse process of tunneling between i_{N+1}st and i_0th grains

$$A_{i_0,i_{N+1}}^* = A_{i_{N+1},i_0} = \langle 0|\hat{c}_{i_0}\hat{S}\hat{c}_{i_{N+1}}^\dagger|0\rangle. \tag{82}$$

The probability of the tunneling process $P = A^*A$ can be presented as an electron world line shown in Fig. 8. The first interval of the world line (τ_i, τ_0) represents the initial state of the system where the electron is located at the site i_0. The following part (τ_0, τ_{N+1}) represents the amplitude of the tunneling process $A_{i_0,i_{N+1}}$ from the site i_0 to the site i_{N+1}. The interval $(\tau_{N+1}, \tau'_{N+1})$ describes the system with the electron located on the site i_{N+1}. The next interval (τ'_{N+1}, τ'_0) represents the amplitude A_{i_{N+1},i_0} of the process of tunneling back to the site i_0, and the last interval (τ'_0, τ_f) represents the system back in the initial state where the electron is located at site i_0.

The world line that represents the tunneling probability $P = A^*A$ is very similar to the closed loops in Fig. 7. The only difference is that in the world line of Fig. 8 the Green functions on the sites i_0 and i_{N+1} must be taken in a different way: In order to capture the energy resolution of electron energies of the initial and final states we will assume that both sites i_0 and i_{N+1} have only one quantum state each with energies ξ_0 and ξ_{N+1} (counted with respect to the chemical potential), respectively.

When an electron tunnels from site i_0 to site i_1, the hole is created at site i_0. Such a process is described by the Green function of a negative time argument[38]

$$G_0(\tau'_0 - \tau_0) = n(\xi_0)e^{\xi_0(\tau'_0 - \tau_0)}, \tag{83}$$

where $n(\xi)$ is the electron occupation number that reminds one that the electron is assumed to be initially present at site i_0.

Similarly, the process when the electron is located on site i_{N+1} is described by the Green function of positive arguments

$$G_{N+1}(\tau'_{N+1} - \tau_{N+1}) = -[1 - n(\xi_{N+1})]e^{-\xi_{N+1}(\tau'_{N+1} - \tau_{N+1})}, \tag{84}$$

where the occupation number $1 - n(\xi_{N+1})$ insures that the final state must be empty for an electron to be able to tunnel there.

Before proceeding to the calculation of tunneling probability let us first fix notations for all time coordinates in Fig. 8: To distinguish the time coordinates that belong to the left part of the world line representing the direct tunneling process from the ones that belong to the right part representing the inverse process we denote the later ones by primes. Further, to distinguish the time coordinates of positive and negative charges we use the index $+$ and $-$, respectively.

The tunneling probability P represented by the world line shown in Fig. 8 can be calculated easily for the case of a diagonal (short range) Coulomb potential $E_{ij}^c = \delta_{ij} E_i^c$.

1. Short range Coulomb interaction

In this case time integrations on each site k can be implemented independently and the tunneling probability P is given by the product of probabilities P_k representing the single site contributions. To find a single site contribution P_k we note that there are two possible arrangements of the tunneling times: (i) $\tau_{k-} < \tau_{k+} < \tau_{k-}' < \tau_{k+}'$ and (ii) $\tau_{k+} < \tau_{k-} < \tau_{k+}' < \tau_{k-}'$. Let us first consider the case (i): The Coulomb energy of such configuration is given by

$$U_k^{c(i)} = E_k^+ (\Delta \tau_k + \Delta \tau_k'), \tag{85}$$

where the electron excitation energy E_k^+ is given by Eq. (10) and the time intervals $\Delta \tau_k, \Delta \tau_k' > 0$ are defined as $\Delta \tau_k = \tau_{k+} - \tau_{k-}$ and $\Delta \tau_k' = \tau_{k+}' - \tau_{k-}'$. The internal energy of the configuration according to Eq. (80) is

$$\mathcal{U}_k = -\ln[g_k \delta_k / 2\pi] + \ln(\Delta \tau_k + \Delta \tau_k'), \tag{86}$$

where g_k is the tunneling conductance between the kth and $(k+1)$st grains. One can see that the second arrangement of charges (ii) differs only by the expression for the Coulomb energy

$$U_k^{c(ii)} = E_k^- (|\Delta \tau_k| + |\Delta \tau_k'|). \tag{87}$$

Now with the help of Eqs. (85)–(87) we find the elementary probability P_k determined by the sum of the contributions (i) and (ii) as

$$P_k = \frac{g_k \delta_k}{2\pi} \int_0^\infty d\Delta \tau d\Delta \tau' \frac{e^{-E_k^+(\Delta\tau+\Delta\tau')} + e^{-E_k^-(\Delta\tau+\Delta\tau')}}{\Delta \tau + \Delta \tau'},$$

and taking the integrals we obtain

$$P_k = \frac{g_k \delta_k}{\pi \widetilde{E}_k}, \tag{88}$$

where the energy $\widetilde{E}_k$ is the combination of the energies E_k^+ and E_k^- defined by Eq. (13).

To find the time dependence of the tunneling process we notice that the time intervals $\tau_0' - \tau_0$ and $\tau_{N+1}' - \tau_{N+1}$ coincide within the accuracy of the inverse charging energy. In the limit that we consider, the charging energy is the largest scale in the problem that allows us to take both intervals to be equal to the "width" of the loop τ_w that at the same time represents the time that electrons spend on site i_{N+1}. Thus the time dependence of the tunneling process is simply given by

$$e^{-(\xi_{N+1} - \xi_0)\tau_w}. \tag{89}$$

Making an analytical continuation to real times $\tau_w = it_w$ and taking the last integral over t_w we obtain the delta function

$$2\pi \delta(\xi_{N+1} - \xi_0), \tag{90}$$

that indeed shows that the process is elastic, i.e., an electron can tunnel only to the state with exactly the same energy.

Finally, for the tunneling probability of the elastic process we obtain

$$P_{el} = w\delta(\xi_{N+1} - \xi_0)g_0 \prod_{k=1}^N P_k, \tag{91}$$

where the factor $w = n(\xi_0)(1 - n(\xi_{N+1}))$ takes into account the occupation numbers of the initial and final states. We see that the total probability P_{el} contains the product of all conductances, mean level spacings, and inverse energies along the tunneling path. For this reason it is convenient to define the geometrical averages of these quantities along the tunneling path that were introduced in Sec. II B 1. Writing the total probability P_{el} in terms of the quantities $\bar{g}, \bar{\delta}, \bar{E}$ we obtain the result (11).

2. Long range Coulomb interaction

In the presence of the long range Coulomb interaction the situation is more complicated since the integrals over variables $\tau_{\pm k}$ cannot be taken on each site k independently. However, one can estimate the tunneling probability P_{el} by finding its upper and lower limits: Indeed, neglecting the long range part of the potential (as we did above) one gets the upper boundary. The lower boundary can be obtained by considering such trajectories where all $\Delta \tau_k$ (as well as $\Delta \tau_k'$) are of the same sign. In this case the electric field created by the charges on step k does not interfere with other charges and thus one can implement integrations independently. This gives a lower boundary for the tunneling probability P_k which is about a factor 2 smaller than the tunneling probability P_k obtained neglecting the long range part of the Coulomb potential. This results in the boundaries of the factor $0.5 \lesssim c < 1.0$ in the expression for localization length ξ_{el} in Eq. (14).

The above derivation of tunneling probability P_{el} has been done for the zero temperatures case. Of course the finite temperatures and the presence of phonons are necessary for the realization of the Mott-Efros-Shklovskii mechanism. Calculation of the tunneling probability at zero temperature is justified as long as inelastic tunneling processes that are considered in the next section can be neglected.

B. Hopping via inelastic cotunneling

According to the theory of transport through a single quantum dot in the Coulomb blockade regime,[24] at temperatures $T > \sqrt{\delta E_C}$ the electron transport is dominated by inelastic cotunneling processes. In this case the process of charge transfer from one contact to the dot and the following process of charge transfer from the dot to the other contact is realized by means of different electrons. As a result, the en-
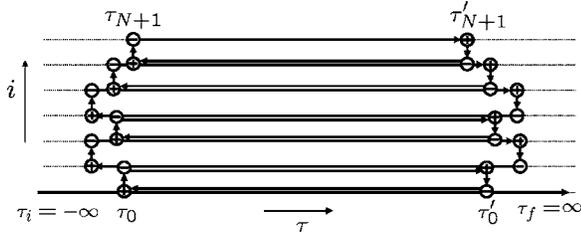
FIG. 9. Tunneling via inelastic cotunneling between 0th to the $(N+1)$st grains.

ergy of the outcoming electron differs from that of the incoming one and the dot energy is changed after each elementary charge transfer process. For this reason inelastic cotunneling processes are allowed only either at finite temperature or if a finite voltage is applied to contacts.

In order to include the inelastic cotunneling processes into the scheme that we described above, we first have to generalize the charge representation to finite temperatures. The main difference is that the imaginary time τ at finite temperatures is constraint to the interval $(0,\beta)$, where $\beta = 1/T$ is the inverse temperature. Also the expressions for internal energies must be properly modified. Using the Matsubara technique one can show that the proper expressions for the internal energies can be obtained from the corresponding expressions at zero temperature via the substitution $1/\tau \rightarrow \pi T / \sin(\pi T \tau)$, such that, for example, the quadrupole internal energy given by Eq. (45) becomes

$$\mathcal{U}_n^q = \ln[2 \sin^2(\pi T|\Delta \tau|)/g \pi T^2]. \qquad (92)$$

In general, one should also take into account the static charges that formally appear from the integer winding numbers in the phase representation at finite temperature. However, inclusion of static charges is not needed for the description of virtual cotunneling processes that we consider.

The representation of an inelastic cotunneling process in terms of the electron world lines is shown in Fig. 9. This process resembles the world line of the elastic cotunneling process shown in Fig. 8 since the electron charge is transferred from the site i_0 to the site i_{N+1} and then back to the site i_0 in essentially the same way. However, the important difference is that the diagram for the inelastic process is constructed out of quadrupole loops only. This implies that the charge is transferred by means of different electrons on each hop.

As in the case of the elastic cotunneling we assume that the sites i_0 and i_{N+1} have only one state each with energies ξ_0 and ξ_{N+1} respectively. Similarly to the previous subsection consideration, we write down the Green functions representing the initial and final states: when an electron tunnels from site i_0 to site i_1, a hole is created at site i_0. Such a process is described by the Green function of negative time arguments

$$G_0(\tau_0' - \tau_0) = n(\xi_0)e^{\xi_0(\tau_0' - \tau_0)}. \qquad (93)$$

The process where an electron is located on the site i_{N+1} is described by the Green function of positive arguments

$$G_{N+1}(\tau_{N+1}' - \tau_{N+1}) = -[1 - n(\xi_{N+1})]e^{-\xi_{N+1}(\tau_{N+1}' - \tau_{N+1})}. \qquad (94)$$

The time coordinates τ_{+k} and τ_{-k} are again confined within the temporal interval $\Delta \tau_k \sim 1/E_0^c$. This means that the time arguments of the quadrupole energies $\mathcal{U}_n^q$ in Eq. (92) are approximately the same and equal to the "width" of the loop formed by the world line in Fig. 9: $\tau_w \approx \tau_0' - \tau_0 \approx \tau_{N+1}' - \tau_{N+1}$. Thus the contributions of quadrupole internal energies $\mathcal{U}_n^q$ to the tunneling probability P_{in} result in the factor

$$\left(\frac{\bar{g} \pi T^2}{\sin^2(\pi T \tau_w)} \right)^N, \qquad (95)$$

where $\bar{g}$ is the geometrical average of the conductances over the electron path. Integrations over the intermediate times τ_k of the inelastic process can be easily done for the case of the diagonal Coulomb interaction $E_{ij}^c = E_0^c \delta_{ij}$.

1. Short range Coulomb interaction

In the case of the short range interaction the integrations over $\Delta \tau_k$ and $\Delta \tau_k'$ can be performed independently and each of the integrations results in the contribution

$$\int_{-\infty}^{\infty} d\Delta \tau_k e^{-E_k^c |\Delta \tau_k| + \mu_k \Delta \tau_k} = 2/\tilde{E}_k, \qquad (96)$$

where the energy $\tilde{E}_k$ is defined in Eq. (13). Collecting all the terms for the probability P_{in} one obtains

$$P_{in}(\tau_w) = \frac{w\bar{g}^{N+1}}{2\pi^{N+1}} \left(\frac{2\pi T}{\bar{E} \sin(\pi T \tau_w)} \right)^{2N} e^{-\Delta \tau_w}. \qquad (97)$$

Finally, in order to find the tunneling probability P_{in} via the inelastic cotunneling one has to make the analytical continuation in Eq. (97) to the real times $\tau_w = it_w$ and integrate over t_w arriving at

$$P_{in} = \frac{wg^{N+1}}{2\pi^{N+1}} \int_{-\infty}^{+\infty} dt \left[\frac{2\pi iT}{\bar{E} \sinh(\pi Tt)} \right]^{2N} e^{-i\Delta t}. \qquad (98)$$

Here the singularity of the function $\sinh^{-2N}(\pi Tt)$ is assumed to be in the upper half of the complex plain. At zero temperature one can easily calculate the integral over the variable t in Eq. (98) by shifting the contour of integration in the complex plain either to negative complex infinity in the case of positive Δ or to the positive complex infinity in the case of negative Δ. In the first case we obtain $P_{in} = 0$. This reflects the fact that the real tunneling process with an increase of the electron's energy is forbidden at $T = 0$. In the latter case, $\Delta < 0$, the zero temperature probability is determined by the pole of the function $\sinh^{-2N}(\pi Tt)$ that results in the answer presented in Eq. (22).

At finite temperatures the integral in Eq. (98) can be expressed in terms of the Euler Gamma functions that lead to the general expression (18).

2. Long range Coulomb interaction

In the presence of the long range part of the Coulomb interaction the integrations over time variables on each step

cannot be performed independently. This makes it difficult to derive an exact analytic formula for the tunneling probability. However, as in the case of the elastic cotunneling, one can easily estimate the upper and lower boundaries for this quantity: neglecting the off-diagonal part of the Coulomb interaction, as above, we get the upper boundary. The lower boundary can be obtained by considering the trajectories where all $\Delta\tau_k$ (as well as all $\Delta\tau_k'$) have the same sign. This consideration constraint the values of the coefficient c in the expression for the localization length to the interval $1/4 \lesssim c < 1$. One can also see that only the sites that are geometrically close to each other can interfere. Indeed, let us consider two sites k_1 and k_2 in the tunneling path that are relatively far from each other. Interference effects related to these sites can be included as a small correction. One can easily see that the first order correction vanishes because the average electric field on the site k_1 created by the charges on the site k_2 is zero and vias versa. Thus the contribution of the interference effects falls off with the distance as $1/r^2$ at least. This ensures that there is no dramatic effect related to the long-range part of the Coulomb interaction apart from the numerical correction in the localization length. The same arguments apply to the case of the elastic cotunneling.

X. MOTT TRANSITION IN A PERIODIC GRANULAR ARRAY

In this section we discuss the metal to insulator transition in the regular periodic granular arrays. The Mott gap Δ_M for such a system in the nearest neighbor hopping (quadrupole) approximation considered in Sec. VII becomes exponentially small at large gz, Eq. (3), but remains always finite. In this section we consider how this result is affected by inclusion of high order tunneling processes that are represented by higher order loops in the classical model of Coulomb charges.

In Sec. IX it was shown that at $gz \ll 1$ the contribution of higher order tunnelings processes to the tunneling probability P at zero temperature is suppressed by the factor

$$\left(\frac{g\delta}{E_0^c}\right)^N, \tag{99}$$

where N is the order of the tunneling process. It is clear that the same parameter defines the smallness of the contribution of the high order loops to the thermodynamic potential Ω of the classical gas. If the tunneling conductance is not small, $gz > 1$, the Coulomb interaction of two classical charges is strongly renormalized and instead of the bare Coulomb interaction E_0^c in Eq. (99) one has to take the renormalized one—that is the Mott gap Δ_M given by Eq. (3). We come to the conclusion that the high order loops can be neglected as long as

$$\Delta_M > g\delta. \tag{100}$$

Equation (100) can be rewritten in terms of the intergranular tunneling conductance g in the way

$$g < g_c, \tag{101}$$

where the critical conductance g_c is defined by Eq. (4). The samples characterized by the tunneling conductance $g < g_c$ at zero temperatures are insulators with the Mott gap Δ_M given by Eq. (3). The result (4) for the critical conductance g_c coincides with the value of the stability of the metallic state obtained via the diagrammatic technique from the metallic side of the transition $g > g_c$.[5] This fact confirms that the value (4) indeed separates the metallic and insulating states at $T=0$.

Note that the accuracy of the above considerations is not sufficient to establish whether the transition between the insulating and metallic states is continuous or discontinuous. The details of the behavior of the Mott gap Δ_M in the immediate vicinity of the critical conductance g_c are not known yet.

So far we considered the limit of the zero temperature. Physics of the Mott transition at finite temperature depends crucially on whether the transition is continuous at $T=0$ or it is discontinuous. In the case of continuous transition at $T=0$, at finite temperatures the metallic and insulating phases, strictly speaking, cannot be distinguished. Indeed, the Mott transition does not break any symmetry and the conductivity at finite temperatures is finite even in the "insulating" phase where it acquires the activation form, Eq. (5). However, in the case of the first order phase transition at $T=0$, due to arguments of physical continuity, the transition has to remain of the first order in some interval of finite temperatures. Such a transition would result in the jump of the physical observables even at finite temperatures.

XI. CONCLUSIONS

We have investigated Coulomb interaction effects and transport properties of granular conductors in the insulating (dielectric) regime. We have considered both the periodic and irregular arrays and taken into account the random on-site electrostatic potential. We find the dependence of the Mott gap on the tunneling conductance in the dielectric regime of a regular array. The insulator to metal transition takes place at the critical conductance given by Eq. (4) where the Mott gap becomes of the order of the inverse escape time of the electron from a single grain $g\delta$. The order of the transition remains, however, an open question. For the case of irregular granular arrays we have derived the Efros-Shklovskii law for the conductivity of a granular system. Depending on the temperature the hopping conductivity is dominated by either elastic or inelastic cotunneling processes. In the latter case the localization length weakly (logarithmically) depends on temperature. Our model and the obtained results indicate the origin of the ES conductivity in the spatially periodic arrays of semiconducting quantum dots although the present consideration does not literally apply to semiconductor granules. Extending our approach to semiconductor arrays requires more work; in particular, the generalization of our technique onto the situation of large level spacing δ is needed.

Finally, we would like to discuss the possibility of the observation of the Mott law in granular metals. To recall, in

semiconductors, the Efros-Shklovskii law may turn to the Mott behavior with the increase of temperature. This happens when the typical electron energy ε involved in a hopping process becomes larger than the width of the Coulomb gap Δ_c, i.e., when it falls into the flat region of the density of states where Mott behavior is expected. To estimate the width of the Coulomb gap Δ_c, one compares the ES expression (6) for the density of states with the bare density of states ν_{g_0}, i.e., the DOS in the absences of the long-range part of the Coulomb interactions obtaining

$$\Delta_c = \left(\frac{\nu_{g_0} e^{2d}}{\widetilde{\kappa}^d} \right)^{1/(d-1)}. \tag{102}$$

As we discussed in Sec. II in granular metals the density of states ν_{g_0} that is relevant for hopping conductivity does not coincide with the physical DOS, since only the lowest energy excitations in a single grain must be counted in determination of ν_{g_0}.[4]

Bare density of ground states ν_{g_0} can be easily estimated for the case of strong on-site disorder that we assumed so far as follows: The width of the potential distribution μ_i is approximately given by the average on-site Coulomb energy so that

$$\nu_{g_0} \approx 1/E_0^c a^d. \tag{103}$$

Inserting this value into Eq. (102) for the width of the Coulomb gap we obtain $\Delta_c \sim E_0^c$ meaning that there is no flat region in the density of ground states and, thus, the Mott regime is not possible. However, the Mott regime may appear in the case of weak electrostatic disorder (for a detail discussion of this issue, see Ref. 4). In this case the subgap states will appear very rarely over the granular sample and the bare density of ground states ν_{g_0} is greatly reduced and as a result the width of the Coulomb gap substantially shrinks.

One can easily repeat all the considerations that led us to the expressions for the hopping conductivity for the case of flat density of ground states arriving at the Mott expression

$$\sigma \sim \sigma_0 \exp\left[-(T_0^M/T)^{1/(d+1)} \right], \tag{104}$$

where the temperature T_0^M is given by

$$T_0^M \sim 1/\nu_{g_0} a^d \xi^d, \tag{105}$$

with the dimensionless localization length ξ given by either Eq. (14) or (21) depending on the temperature regime.

The crossover temperature T^* between the two regimes is easily obtained exactly as in the case of semiconductors; by comparing Eqs. (1) and (104) and using the corresponding expressions for T_0 we get

$$T^* \sim e^2 a \xi \left(\frac{e^4 \nu_{g_0}^2}{\widetilde{\kappa}^{d+1}} \right)^{1/(d-1)}. \tag{106}$$

As a final note, we would like to comment on the possibility of fixing the numerical coefficient in the expression for the temperature T_0 that enters exponents of the expressions for the hopping conductivity. In the case of hopping conductivity in semiconductors this coefficient can be fixed by means of mapping the problem of the impurity states in

semiconductors onto a simple classical percolation model that can be investigated completely numerically.[22,39,40] In granular metals, on the qualitative level, the situation is very similar, however, it is not clear that a model of a granular metal can be reduced to a simple unique percolation problem without making additional uncontrolled assumptions. The main complications are that the tunneling process in granular metals (i) in general goes via the curved trajectory, and (ii) depends on the number of the parameters of the intermediate grains in the tunneling path as E_i^c, δ_i, g_{ij}. For this reason, in general, we expect that the coefficient in the expressions for T_0 is not universal and is dependent on the morphology of a granular sample.

After the present study was completed, we became aware of the report[41,42] where the hopping transport in granular metals via elastic and inelastic cotunneling in the linear regime (low applied voltage) was considered; our results, and in particular results for the localization lengths, agree up to the coefficients that cannot be fixed within a generic granular model due to the reasons presented above.

APPENDIX A: MINIMIZATION OF THE PROBABILITY OF THE INELASTIC COTUNNELING

The result for the hopping conductivity in the regime of inelastic cotunneling is obtained via minimization of expression (18) with respect to the hopping distance N under the constraint

$$Na\Delta\widetilde{\kappa}/e^2 = b, \tag{A1}$$

following from the ES expression for the density of states, with b being a constant. As a first step of the minimization procedure, we write the tunneling probability P_{in} in a standard form:

$$P_{in} \sim e^A, \tag{A2}$$

with the action A given by

$$A = -\frac{2N}{\xi_{in}} + \ln \Gamma(N + i\Delta/2\pi T) + \ln \Gamma(N - i\Delta/2\pi T)$$
$$- \ln \Gamma(2N) - \Delta/2T, \tag{A3}$$

where ξ_{in} is the localization length given by (21). Taking into account that according to Eq. (A1) $d\Delta/dN = -\Delta/N$, the minimization of the action A with respect to N in the limit of large N yields
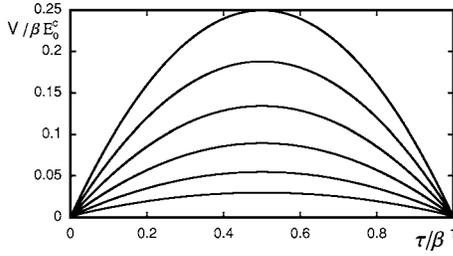
FIG. 10. Dependence of the potential $V(\tau)$ normalized to the value βE_c^0 on imaginary time τ normalized to β obtained via numerical solution of Eqs. (B2) and (B3). The curves are plotted for $E_0^c\beta=40$ and different conductances $gz=0, 0.4, 0.8, 1.2, 1.6,$ and 2.0 (top to bottom). On the scales of β all curves can be fitted by quadratic form B1 with very good ($\sim 10^{-5}$) accuracy. At the same time close to the boundaries there are significant deviations from parabolic law as shown on Fig. 11.

$$\ln 4 + 2/\xi_{in} = \ln(1 + x^2) + \pi x[1 + (2/\pi)\arctan x], \quad \text{(A4)}$$

where the parameter x is defined as

$$x = \Delta/2\pi TN. \quad \text{(A5)}$$

We see that Eq. (A4) fixes the ratio Δ/TN for a given localization length. Using Eq. (A4) and the condition $N \gg 1$, one can simplify the action (A3) and write it in a form

$$A = -\frac{\Delta}{T}\left(1 + \frac{2}{\pi}\arctan x\right), \quad \text{(A6)}$$

which with the help of Eqs. (A1) and (A5) can be further expressed via the parameter x as

$$A = -\sqrt{\frac{2be^2}{aT\tilde{\kappa}}}\sqrt{x\pi}[1 + (2/\pi)\arctan x]. \quad \text{(A7)}$$

Within the logarithmic accuracy the localization length can be viewed as a small parameter; after that one gets from Eq. (A4) that $x \approx 1/\pi\xi$ and Eq. (A7) reduces to the following final expression for the action:

$$A \approx -\sqrt{\frac{8be^2}{\xi_{in}aT\tilde{\kappa}}}. \quad \text{(A8)}$$

We would like to note that, in principle, one should consider an alternative minimization scheme for the tunneling process: The tunneling first occurs with $\Delta=0$ as in the case of the elastic process, and then the electron gets an extra energy due to phonon or electron collision, the contribution from which is not included directly in the expression for the probability of the inelastic process involved. One can easily show that in this case one would get $A=-\sqrt{8be^2/Ta\xi_{in}\tilde{\kappa}}$ that coincides with the action (A8) obtained in the leading order in the parameter $1/\xi_{in}$. One can check, however, that taking large but finite $1/\xi_{in}$ results in the correction to the action (A8) that consists of the reduction of the coefficient $\sqrt{8}$. Thus the process with intrinsic inelasticity that we considered is significantly more favorable. This verifies and validates that the minimization scheme used in our calculations is the appropriate one.
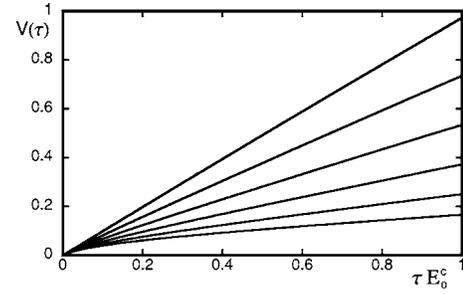

FIG. 11. Dependence of the potential $V(\tau)$ on imaginary time τ measured in units $1/E_0^c$ in the vicinity of the charge located at $\tau = 0$.

APPENDIX B: NUMERICAL SOLUTION OF THE MEAN FIELD EQUATIONS (64) AND (68)

In order to solve Eqs. (64) and (68) numerically one has to choose some limits for the time variable τ. This can be done by considering the finite temperature case and, accordingly, using the Matsubara technique. At finite temperatures the time variable τ is constrained to the interval $(0, \beta = 1/T)$ and all functions as $V(\tau), \rho(\tau)$ have to be periodic with respect to the shift over β. The expected insulating behavior of the potential (69) at finite temperatures is expressed via the relation

$$V(\tau) \rightarrow E_0^c\kappa_\tau^{-1}(\tau - \tau^2/\beta), \quad \text{(B1)}$$

which for $\tau \ll \beta$ coincides with Eq. (69). The above result (B1) can be obtained directly as the Matsubara Fourier transform of the function $V_\omega = 2E_0^c\kappa_\tau^{-1}/\omega^2$.

The kernel of the integral equation (68) also has to be modified as $1/\tau^2 \rightarrow (\pi T)^2/\sin^2(\pi T\tau)$, and Eq. (68) assumes the form

$$q(\tau) = 2\pi gzT^2\int_0^\beta d\tau' \frac{e^{-2V(\tau-\tau')}\sinh[V(\tau') - V(\tau)]}{\sin^2 \pi T(\tau - \tau')}. \quad \text{(B2)}$$

Periodic boundary conditions also account for the periodic location of the external charges, such that within the interval $(0, \beta)$ the Poison equation becomes

$$\partial_\tau^2 V(\tau) = -2E_0^c[q(\tau) - \delta(\tau) - \delta(\tau - \beta) + 1/\beta]. \quad \text{(B3)}$$

Note that we introduced a homogeneous background charge with the charge density $1/\beta$ that on average compensates the external point charges. Without this compensating charge a periodic solution for the potential would not be possible. This term disappears in the limit $\beta \rightarrow \infty$ and thus does not modify the results in the limit $T \rightarrow 0$ that we are interested in.

Equations (B2) and (B3) can be easily solved by iterations. The results for the potential dependencies for conductances gz ranging from 0 to 2 are shown on Fig. 10. On the scales of β all curves follow the law (B1) that assumes insulating behavior. On the scales of the order E_c^{-1}, as one expects, there are deviations from the asymptotic behavior as shown on Fig. 11.

[1] A. Gerber, A. Milner, G. Deutscher, M. Karpovsky, and A. Gladkikh, Phys. Rev. Lett. **78**, 4277 (1997).

[2] R. Parthasarathy, Xiao-Min Lin, K. Elteto, T. F. Rosenbaum, and H. M. Jaeger, Phys. Rev. Lett. **92**, 076801 (2004).

[3] K. B. Efetov and A. Tschersich, Europhys. Lett. **59**, 114 (2002); Phys. Rev. B **67**, 174205 (2003).

[4] J. Zhang and B. I. Shklovskii, Phys. Rev. B **70**, 115317 (2004).

[5] I. S. Beloborodov, K. B. Efetov, A. V. Lopatin, and V. M. Vinokur, Phys. Rev. Lett. **91**, 246801 (2003).

[6] I. S. Beloborodov, A. V. Lopatin, and V. M. Vinokur, Phys. Rev. B **70**, 205120 (2004).

[7] M. V. Feigel'man, A. S. Ioselevich, and M. A. Skvortsov, Phys. Rev. Lett. **93**, 136403 (2004).

[8] V. I. Kozub, V. M. Kozhevin, D. A. Yavsin, and S. A. Gurevich, Pis'ma Zh. Eksp. Teor. Fiz. **80**, 5 (2005).

[9] Y. L. Loh V. Tripathi, and M. Turlakov, cond-mat/0501749.

[10] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interaction in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).

[11] L. P. Gorkov, A. L. Larkin, and D. E. Khmelnitskii, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 248 (1979) [Sov. Phys. JETP **30**, 228 (1979)].

[12] P. Sheng, B. Abeles, and Y. Arie, Phys. Rev. Lett. **31**, 44 (1973); B. Abeles, P. Sheng, M. D. Coutts, and Y. Arie, Adv. Phys. **24**, 407 (1975).

[13] M. Pollak and C. J. Adkins, Philos. Mag. B **65**, 855 (1992).

[14] N. F. Mott, Adv. Phys. **16**, 49 (1967).

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

[16] I. S. Beloborodov, A. V. Lopatin, V. M. Vinokur, and V. I. Kozub, cond-mat/0501094.

[17] V. Ambegaokar, U. Eckern, and G. Schön, Phys. Rev. Lett. **48**, 1745 (1982); for review see Ref. 33.

[18] A. Schmid, Phys. Rev. Lett. **51**, 1506 (1983).

[19] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[20] A. I. Yakimov, A. V. Dvurechenskii, A. I. Nikiforov, and A. A. Bloshkin, Pis'ma Zh. Eksp. Teor. Fiz. **77**, 445 (2003) [JETP Lett. **77**, 376 (2003)].

[21] T. Tran, I. S. Beloborodov, X. M. Lin, V. M. Vinokur, and H. M. Jaeger, Phys. Rev. Lett. **95**, 076806 (2005).

[22] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, New York, 1988).

[23] We consider the system periodic if (i) all the granules are of the same size and shape and arranged in a regular structure; (ii) all the tunneling conductances between the nearest neighbors are equal; and (iii) there is no disorder in the dielectric matrix that causes local fluctuations in granules' chemical potential.

[24] D. V. Averin and Yu. V. Nazarov, Phys. Rev. Lett. **65**, 2446 (1990).

[25] J. E. Mooij, B. J. van Wees, L. J. Geerligs, M. Peters, R. Fazio, and G. Schön, Phys. Rev. Lett. **65**, 645 (1990).

[26] J. S. Meyer, A. Kamenev, and L. I. Glazman, Phys. Rev. B **70**, 045310 (2004).

[27] A. Altland, L. I. Glazman, and A. Kamenev, Phys. Rev. Lett. **92**, 026801 (2004).

[28] M. Muller and L. B. Ioffe, Phys. Rev. Lett. **93**, 256403 (2004).

[29] S. Pankov and V. Dobrosavljevic, Phys. Rev. Lett. **94**, 046402 (2005).

[30] B. I. Shklovskii, Fiz. Tekh. Poluprovodn. (S.-Peterburg) **6**, 2335 (1973) [Sov. Phys. Semicond. **6**, 1964 (1973)].

[31] Equation (18) gives the probability of spin unresolved process, i.e., it includes the sum over spin components of the final state.

[32] L. D. Landau and I. M. Lifshitz, *Quantum Mechnics*, 3rd ed. (Elsevier Science, Oxford, 1977).

[33] G. Schön and A. D. Zaikin, Phys. Rep. **198**, 237 (1990).

[34] We use the dimensionless units for energy of the classical model assuming that the effective temperature $T=1$.

[35] See, for example, L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Course of Theoretical Physics Vol. 5 (Pergamon Press, Oxford, New York, 1980).

[36] D. P. Arovas, F. Guinea, C. P. Herrero, and P. San Jose, Phys. Rev. B **68**, 085306 (2003).

[37] I. S. Beloborodov, K. B. Efetov, A. Altland, and F. W. J. Hekking, Phys. Rev. B **63**, 115109 (2001).

[38] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover Publications, Inc., New York, 1965).

[39] V. Ambegaokar, B. I. Halperin, and J. S. Langer, Phys. Rev. B **4**, 2612 (1971).

[40] A. Miller and E. Abrahams, Phys. Rev. **120**, 745 (1960).

[41] M. V. Feigel'man and A. S. Ioselevich, cond-mat/0502481. During the processing of our manuscript this paper was published in Ref. 42.

[42] M. V. Feigel'man and A. S. Ioselevich, Pis'ma Zh. Eksp. Teor. Fiz. **81**, 341 (2005) [JETP Lett. **81**, 227 (2005)].