Hall effect in two-dimensional systems with hopping transport and strong disorder

A. V. Shumilin[1] and N. P. Stepina[2]

[1]*Ioffe Institute, 194021 St.-Petersburg, Russia*
[2]*Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia*

We reconsider the theory of Hall effect in the systems with hopping conduction. The purpose of this study is to compare the percolation approach based on the optimal triad model with numerical simulations and recent experimental results. We show that, in the nearest-neighbor hopping regime, the results of the percolation theory agree to the simulation. However, in the variable range hopping (VRH) regime, the optimal triad model fails to describe the numerical results. It is related to the extremely small probability to find the optimal triad of sites in the percolation cluster in the VRH regime. The contribution of these triads to the Hall effect appears to be small. We describe the Hall mobility in the VRH regime with the empirical law obtained from the numerical results. The law is in agreement with our recent experimental data in two-dimensional quantum dot arrays with the hopping transport.

I. INTRODUCTION

The hopping transport is one of the fundamental kinds of electron transport. It appears in a number of different systems from doped semiconductors to granular metals and organic semiconductors. The Hall effect in metals and semiconductors with free carriers yields important information on their properties allowing to determine carriers concentration and mobility. However, the understanding of the Hall effect in the hopping regime is far from being complete. Moreover, there is still no general agreement on whether the Hall effect can be observed in the hopping regime at all. The problems exist both in the theoretical and experimental approaches to this phenomenon.

The theoretical study started from the work of Holstein [1]. It was shown that although the Hall effect is absent in the model of two-site one-phonon hops that are invoked in most of the hopping transport theories, the Hall effect exists due to many-phonon processes. The magnetic field-dependent contribution to the hopping probability arises from the interference between the amplitude of direct transition between the initial and the final sites of the hop and the amplitude of indirect, second-order transition, involving an intermediate (the third) site. The interference exists and can be important for hops involving any number of phonons. However, for one-phonon hops it leads only to the interference mechanism of magnetoresistance [2–4]. When all the relevant hops are included in the theory, each triad of localization sites starts to act as a source of the Hall current.

The study [1] was focused on the ac current in a system where the number of electrons is small, compared to the number of localized states. The ac current can be described in terms of averaging the Hall current over all the Hall sources. The problem of dc current is more complex and involves the generated Hall current distribution over the network of Miller-Abrahams resistors [5,6]. The most conventional approach to this problem is the percolation theory that treats the system in the limit of strong disorder.

There are two possible reasons for having strong disorder in a system with hopping transport. The first reason, the position disorder, is related to the random positions of localization sites and is controlled by the dimensionless parameter $n^{1/d}a$, where n is the concentration of sites, a is the localization radius, and d is the system dimension. When this disorder is dominant, the system is in nearest-neighbor hopping (NNH) regime. The temperature dependence of conductivity in this regime follows the Arrhenius law. Although the NNH regime can be realized in the experiment, it is not always easy to distinguish it from the transport due to the carrier activation into the conduction band. The percolation theory was applied to the Hall effect in the NNH regime in studies [7–9]. It was shown that the Hall current is determined essentially by the rare optimal triads of the sites which form a junction for the percolation paths. The resulting Hall mobility exponentially decreases in the limit of strong disorder $n^{1/d}a \to 0$.

Another reason of disorder is the random distribution of the localized state energies. The width of this distribution $\Delta\varepsilon$ should be compared to temperature T. The control parameter of the energy disorder is $\Delta\varepsilon/T$. When this disorder is sufficiently strong, the system is in the variable range hopping (VRH) regime. It can be identified in the experiment due to the unique temperature dependence of conductivity, the Mott law [10,11], or the Efros-Shklovskii law [6] in the systems, where the Coulomb gap is essential. The percolation theory for the Hall effect in the VRH regime was discussed in Refs. [7,9,12–15]. Also, the similar theories were developed for the anomalous Hall effect [15,16]. Although the approach used in these studies was more or less the same, the results are surprisingly different. In Refs. [12,15] the Hall mobility μ_H is predicted to have the exponential dependence on temperature $\mu_H \propto \exp[-(T_0'/T)^{1/4}]$, with T_0' smaller than T_0 in the Mott law [12]. In Refs. [9,14] the dependence follows the power law $\mu_H \propto T^\gamma$. The power-law dependence appeared from the contribution of rare optimal triads of localization sites. These triads consist of three sites close to each other, but with the energies that lead to Miller-Abrahams resistances between the

sites of the triad similar to the critical resistance of the percolation cluster. To be effective, such a triad should be positioned in the intersection of the percolation paths [17], as in the case of NNH conductivity. The characteristic correlation distance L_h between such junctions has been evaluated in Ref. [18] and turns out to be much larger than the hopping network correlation length. As a result, strong mesoscopic effects are expected even in relatively large samples [9].

The drawback of the percolation theory is that it is rigorous only in the analytical limit of a small temperature and localization radius. It becomes especially important when the result of the percolation theory is governed by the extremely rare objects such as the discussed optimal triads. It is not always clear at what temperature these rare objects start to dominate over other triads that are not that effective as Hall current sources, but are significantly more numerous. To understand this, the percolation theory can be compared to the numerical simulations based on the direct solution of Kirchhoff equations. Such attempts were made in studies [19,20]. These simulations support expression $\mu_H \propto \exp[-(T_0'/T)^{1/4}]$. However, the computations in Refs. [19,20] were made with only one disorder realization in a three-dimensional (3D) cubic sample with a size $\sim 13 \times 13 \times 13$ sites. It is not clear if this sample can represent a macroscopic disordered system. Also, these results cannot be directly applied to two-dimensional (2D) systems which are the main focus of our study.

The most important problems in the experimental study of hopping Hall effect are the small Hall resistance value and the fact that the Hall effect is masked by magnetoresistance. Nevertheless, a few experiments on the Hall conductivity in the hopping regime were made in Refs. [21–26]. Amitay and Pollak [26] attempted to measure the impurity-hopping Hall effect in germanium and silicon at a sufficiently low temperature and impurity concentration so that any Hall effect from carriers in delocalized states would be negligible. The authors have not succeeded in the observation of Hall effect even though the sensitivity of their measuring system seemed to be sufficient to detect it. The negative result was also obtained in Ref. [27], measured in the system in the deep localization regime. The Hall effect due to hopping conductivity was not detected, and it set an upper bound on the Hall conductivity of 1.7×10^{-13} Ohm cm^{-1} for the given conditions. Most of the experimental observations of Hall effect in the dielectric regime were obtained near the metal-insulation transition (MIT) in the 3D case [21], where a small range of $\sigma_{xx}(T)$ dependence was observed and the absolute value of σ was rather large. It means that the system can not be in a strong localization regime.

The recent advances in the technology allowed us to grow the arrays of Si/Ge quantum dots (QDs) that display the VRH conductivity in the Coulomb gap regime [28]. The interesting property of these arrays is that the localization radius near the Fermi level is much larger than that for impurities in semiconductors. It is comparable or larger than the QD size and interdot distance, and can be controlled by changing the filling factor. The possibility to change the structural parameters of QDs allows the novel way to control the disorder not possible in ordinary doped semiconductors. Recently, we obtained experimental results of the Hall effect in this system [29]. To understand the obtained experimental data, we need to compare our results to the theory. However, the theoretical results themselves are not well established; therefore, we have to reconsider the theory of the Hall effect in the systems with hopping transport before the comparison can be made.

Modern computation potentials allow us to significantly improve the numerical approach [19,20]. As a result, we can verify the results of percolation theory and understand the correct dependence of the Hall mobility on the system parameters for the moderately low temperatures reachable in the experiment. We focus our study on the comparison of the percolation arguments with numeric simulations and with our recent experimental results. Also, we restrict ourselves to the 2D case that was not treated numerically in Refs. [19,20] and is relevant for our experiments.

The paper is organized as follows. In Sec. II we derive the general equations in the form that allows both analytical and numeric treatments. In Sec. III we consider our equations in strongly disordered systems in NNH, Mott law, and Efros-Shklovskii law VRH regimes and compare the percolation arguments with the results of numeric simulation. In Sec. IV we compare our theory with the experimental results. In Sec. V we provide a general discussion of the obtained results.

II. GENERAL EQUATIONS

In this section we extend the approach [9] to the description of the hopping Hall effect to include the triads of sites with arbitrary occupation numbers. First, we consider the interaction of localized electrons with phonons in the density matrix formalism. We derive the rate equations that describe the ordinary two-site one-phonon hops controlling the conductivity and three-site two-phonon hops responsible for the Hall effect. These equations are then linearized to study the linear response to electric and magnetic fields. It leads to the system (9)–(12) of the modified Kirchhoff equations, that is a useful starting point for both numerical simulations and percolation treatment.

In our study we adopt the model when the electrons are localized on pointlike sites. The sites have random energies ε_i that are larger than the overlap integrals t_{ij}. This model is conventional to the hopping transport. In real systems, for example, in quantum dot arrays with hopping conduction, the physics can be more complex. It can include the finite size of a quantum dot and the importance of states of intermediate dots for the long-range hopping. However, the pointlike site model is known to be a good starting point to study the hopping transport. It was assumed in most of the previous studies in the field [7–9,12–15,19,20]. Therefore, we think that it is instructive to achieve a reliable understanding of the hopping Hall effect in the pointlike site model before starting to consider peculiarities of complex systems. We also do not consider the electron spin to make our model as simple as possible.

The starting point of our consideration is the Hamiltonian of a system with hopping transport after the polaron transformation [30]

$$H = \sum_i \varepsilon_i n_i + \widehat{T} + H_{ph}, \quad \widehat{T} = \sum_{ij} t_{ij} \Phi_{ij} a_i^+ a_j. \quad (1)$$

Here, ε_i is the onsite energy, $n_i = a_i^+ a_i$ is the operator of electron density on site i. The overlap integral t_{ij} in the magnetic field can be expressed as $t_{ij} = t_0 \exp(-r_{ij}/a) \exp(i \frac{e}{2c} \mathbf{B} \cdot [\mathbf{r}_i \times \mathbf{r}_j])$, where $\mathbf{B}$ is the applied magnetic field. $\mathbf{r}_i$ is the position of site i, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between sites i and j, a is the localization length. In ordinary doped semiconductors far from the metal-insulator transition a corresponds to the localization radius of a single impurity state. In semiconductors close to the metal-insulator transition and in more complex systems, a can be strongly renormalized by cotunneling processes. It is especially important for granular metals and quantum dot arrays [31]. This renormalization can easily make a larger than the intersite distance [31]. $H_{ph} = \sum_q \omega_q b_q^+ b_q$ is the phonon Hamiltonian. There is no onsite electron-phonon interaction due to the polaron transformation, however, the interaction is included in the transition elements that are proportional to

$$\Phi_{ij} = \exp\left\{ \sum_{\mathbf{q}} b_{\mathbf{q}}^+ [u_j^*(\mathbf{q}) - u_i^*(\mathbf{q})] - b_{\mathbf{q}}[u_j(\mathbf{q}) - u_i(\mathbf{q})] \right\}. \tag{2}$$

Here, $u_i(\mathbf{q}) = (2\mathcal{N})^{1/2} \gamma(\mathbf{q}) \exp(-i\mathbf{q}\mathbf{r}_i)$, $\mathcal{N}$ is the number of atoms in the lattice, and $\gamma(\mathbf{q})$ describes the electron-phonon interaction [30].

The electron-electron Coulomb interaction does not contribute explicitly to the Hamiltonian (1). We assume that it can be added to the energy $\varepsilon_i \to \varepsilon_i + \sum_j e^2 n_j / r_{ij}$ to include the effects of the Coulomb gap or the Coulomb glass, however, the Coulomb energy does not interfere with the hopping process itself. In the theory of hopping transport it is assumed that the hopping rates are small compared to frequencies $|\varepsilon_i - \varepsilon_j|/\hbar$. Accordingly, we expand the electron density matrix $\widehat{\rho}$ over small hopping rates. With the Hartree decoupling [30] we assume that the zero-order density matrix can be expanded as a product

$$\widehat{\rho}^{(0)} = \prod_i \widehat{\rho}_i^{(0)}, \quad \widehat{\rho}_i^{(0)} = f_i |1\rangle\langle 1| + (1 - f_i)|0\rangle\langle 0|. \tag{3}$$

Here, $\widehat{\rho}_i^{(0)}$ is the density matrix on site i. It corresponds to some probability f_i for site i to have an electron. In equilibrium, f_i is the Fermi function $f_i = 1/(e^{(\varepsilon_i - \mu)/T} + 1)$. The zero-order density matrix corresponds to the situation when the electrons rest on their localization sites.

The dynamics of density matrix $\widehat{\rho}$ can be described with the series of perturbation equation

$$\widehat{\rho}^{(n+1)}(t) = -i \int_{-\infty}^t [\widehat{T}, \widehat{\rho}^{(n)}(t')] dt'. \tag{4}$$

The structure of matrix $\widehat{\rho}^{(n)}$ that appears due to the dynamics (4) is more complex than the structure of $\widehat{\rho}^{(0)}$. The operator $\widehat{T}$ mixes different localization states and, therefore, $\widehat{\rho}^{(n)}$ includes the elements that correspond to electron transition from one localization site to another.

We apply the reduction procedure to separate the small transition elements from the actual hopping process that changes the filling numbers of the sites. To calculate the addition to the density matrix $\widehat{\rho}_i^{(0)}$ due to the hopping we take the trace over phonon states and all the other sites,

i.e., $\delta\widehat{\rho}_i = \text{Tr}_{\text{ph}, j \neq i}(\widehat{\rho} - \widehat{\rho}^{(0)})$. Finally, to describe the electron dynamics as hopping we should assume that the dynamics of the reduced density matrices $\widehat{\rho}_i^{(0)}$ is much slower than the oscillations of perturbations to this matrix that occur at frequencies $\sim (\varepsilon_i - \varepsilon_j)/\hbar$. It allows us to substitute $\widehat{\rho}_i^{(0)}(t')$ with $\widehat{\rho}_i^{(0)}(t)$ in the expression (4).

The effect of ordinary two-site one-phonon hops is then expressed as a reduction of the second-order density matrix $\widehat{\rho}^{(2)}$ where the phonon exponents in (2) are expanded up to the second order of phonon creation and annihilation operators. It leads to the conventional equations for the hopping rates. To consider the Hall effect, one should include three-site, two-phonon hops that are described by the reduced third-order density matrices $\widehat{\rho}^{(3)}$, where the phonon exponents are expanded up to the fourth order over the phonon creation and annihilation operators.

With the assumptions mentioned above we obtain the hopping transport equations that include three-site two-phonon hops:

$$\frac{df_i}{dt}$$

$$= \sum_{j \neq i} f_j (1 - f_i) \left(W_{ij} + \sum_{k \neq i, j} W_{ikj}^{(0)}(1 - f_k) + W_{ikj}^{(1)} f_k \right)$$

$$- f_i (1 - f_j) \left(W_{ji} + \sum_{k \neq i, j} W_{jki}^{(0)}(1 - f_k) + W_{jki}^{(1)} f_k \right). \tag{5}$$

Here, W_{ij}, $W_{ikj}^{(0)}$, and $W_{ikj}^{(1)}$ describe the rates of $j \to i$ hops and yield the contributions to df_i/dt proportional to $f_j(1 - f_i)$. The rate W_{ij} stands for the ordinary two-site hop. The rate $W_{ikj}^{(0)}$ stands for the two-phonon hop involving the intermediate site k that is assumed to be free. Therefore, its contribution is proportional to $(1 - f_k)$. In the similar way, $W_{ikj}^{(1)}$ describes the three-site hop involving the filled site k. The corresponding contribution to df_i/dt includes the term f_k. For a given pair ij, the role of the intermediate site can be played by any site of the system other than i and j.

The rates W_{ij}, $W_{ikj}^{(0)}$, and $W_{ikj}^{(1)}$ are expressed as follows:

$$W_{ij} = \frac{1}{\tau_0} \exp(-2r_{ij}/a) \mathcal{N}(\varepsilon_j - \varepsilon_i), \tag{6}$$

$$W_{ikj}^{(0)} = \frac{1}{4} |t_{ij} t_{jk} t_{ki}| \frac{\mathbf{B} S_{ikj}}{2\Phi_0}$$

$$\times \left(\frac{W_{kj} W_{ik}}{|t_{kj}^2 t_{ik}^2|} + \frac{W_{ij} W_{kj}}{|t_{ij}^2 t_{kj}^2|} + \frac{W_{ij} W_{ki}}{|t_{ij}^2 t_{ik}^2|} \right), \tag{7}$$

$$W_{ikj}^{(1)} = -\frac{1}{4} |t_{ij} t_{jk} t_{ki}| \frac{\mathbf{B} S_{ikj}}{2\Phi_0}$$

$$\times \left(\frac{W_{kj} W_{ik}}{|t_{kj}^2 t_{ik}^2|} + \frac{W_{ij} W_{jk}}{|t_{ij}^2 t_{kj}^2|} + \frac{W_{ij} W_{ik}}{|t_{ij}^2 t_{ik}^2|} \right). \tag{8}$$

Here, τ_0 is the constant describing the characteristic (ordinary) hopping time between close neighbors without a large energy exponent. The similar time for three-site hops is $t_0 \tau_0^2/\hbar$. It

is assumed to be small, compared to τ_0. $\mathcal{N}(\varepsilon_j - \varepsilon_i)$ is the effective probability to find a phonon for the hop. We consider $\mathcal{N}(\varepsilon_j - \varepsilon_i) = 1$ when $\varepsilon_j > \varepsilon_i$ and $\mathcal{N}(\varepsilon_j - \varepsilon_i) = \exp[(\varepsilon_j - \varepsilon_i)/T]$ otherwise. $\mathbf{S}_{ikj}$ is the "vector area" of the triangle ikj, Φ_0 is the flux quanta. We focus on small magnetic fields $\mathbf{B}\mathbf{S}_{ikj} \ll \Phi_0$. In the general case, the linear dependence on the magnetic field should be substituted with oscillating dependence $\sin(\mathbf{B}\mathbf{S}_{ikj}/2\Phi_0)$. The expressions (6)–(8) are derived from (4) in Appendix. These expressions agree to [1,9] when the triangle ikj is considered to have small occupation numbers f_i, f_j, and f_k. However, the expressions (6)–(8) allow the description of the general case of the arbitrary occupation numbers.

We discuss the system in the Ohmic regime. It corresponds to the small perturbations of occupation probabilities f_i. This case allows using the linearized version of the general equations. We consider the dc current and stationary equations $df_i/dt = 0$:

$$\sum_j J_{ij} = 0, \qquad (9)$$

$$J_{ij} = \frac{\varphi_i - \varphi_j}{R_{ij}} + \sum_{k \neq i,j} \frac{e^2 \varphi_k}{T} \frac{\mathbf{S}_{ikj}\mathbf{B}}{2\Phi_0} \Gamma_{ikj}, \qquad (10)$$

$$R_{ij} = \frac{T}{e^2 \Gamma_{ij}}, \quad \Gamma_{ij} = \frac{1}{\tau_0} \exp\left(-\frac{2r_{ij}}{a} - \frac{\varepsilon_{ij}}{T}\right). \qquad (11)$$

Here, R_{ij} is the Miller-Abrahams resistor between sites i and j, which has an exponentially broad distribution $R_{ij} \approx R_0 \exp(\xi_{ij})$, $\xi_{ij} = 2r_{ij}/a + \varepsilon_{ij}/T$ in a material with a strong disorder. $\varepsilon_{ij} = (|\varepsilon_i - \varepsilon_F| + |\varepsilon_j - \varepsilon_F| + |\varepsilon_i - \varepsilon_j|)/2$ is the energy term in the Miller-Abrahams resistance expression. ε_F is the Fermi energy. φ_i is the addition to the electrochemical potential of site i due to the applied current. Γ_{ikj} is the rate of the three-site hop. In a strongly disordered system, it can be estimated as

$$\Gamma_{ikj} = \frac{1}{4t_0\tau_0^2} \exp\left(-\frac{r_{ij} + r_{ik} + r_{kj}}{a}\right)$$
$$\times (e^{(|\varepsilon_i - \varepsilon_F| - \varepsilon_{ij} - \varepsilon_{ik})/T} + e^{(|\varepsilon_j - \varepsilon_F| - \varepsilon_{ij} - \varepsilon_{jk})/T}$$
$$+ e^{(|\varepsilon_k - \varepsilon_F| - \varepsilon_{ik} - \varepsilon_{jk})/T}). \qquad (12)$$

In our expressions for the hopping rates Γ_{ij} and Γ_{ijk}, we keep only the exponential terms in the dependencies on r_{ij} and ε_{ij}. This approximation can be applied in strongly disordered systems when the power-law terms are small compared to the exponential ones.

III. HALL CURRENT IN SYSTEMS WITH STRONG DISORDER

In this section we study the hopping Hall effect in systems with strong disorder. We apply two methods to this problem. The first one is the direct numerical solution of the system (9)–(12) of modified Kirchhoff equations. The second is the analytical approach based on the percolation theory. In the frame of the analytical method, we evaluate the contribution of different triads of sites to the Hall effect. Assuming that the effect is dominated by a small number of the so-called optimal triads, we derive the analytical expression for Hall mobility.

The main purpose of this section is to compare the results of numerical and percolation approaches. We provide this comparison for the three important cases: nearest-neighbor hopping (Sec. III A), variable range hopping with a constant density of states (Sec. III B), and variable range hopping in the Coulomb gap (Sec. III C).

In the numerical simulation we consider a square 2D numerical sample with size L and $N = L^2$ localization sites with random positions. The positions are not correlated. Each site i is ascribed with some energy ε_i. The localization distance a, temperature T, and the distribution of the site energies control the degree of disorder and the hopping regime (NNH or VRH). The boundary conditions are periodical. It means that our simulation represents the infinite system composed of $L \times L$ supercells.

The localization sites distribution determines the system of linear equations (9)–(12). The system is solved numerically without the magnetic field and with small magnetic field B when the corrections to the currents are linear with respect to the field B. We find the normal current J at $B = 0$ directed along the electric field that gives us conductivity σ and the current $J_{\text{Hall}} \propto B$ that is perpendicular to the electric field. The Hall mobility μ_H is proportional to the ratio of these currents $\mu_H = J_{\text{Hall}}/JB$. The results are then averaged over disordered configurations. With numeric simulations we are able to find the μ_H dependence on localization radius a and temperature. The absolute value of μ_H is governed by the parameter $t_0\tau_0/\hbar$ that is not discussed in this study.

Our analytical treatment of Eq. (10) relies on the exponentially broad distribution of coefficients Γ_{ikj} in a strongly disordered system. As soon as we discuss the linear effect over magnetic field $\mathbf{B}$, we can consider potentials φ_k in the last term in (10) to be independent from the magnetic field. Values φ_k are proportional to the applied electric field and are determined by the disorder configuration. In this case, each triad of sites ikj acts as an independent source of the Hall current. The total Hall current J_{Hall} flowing through some cross section of the sample (perpendicular to the electric field direction) can be expressed as the sum of contributions $J_{\text{Hall}}^{(ikj)}$ related to triads ikj:

$$J_{\text{Hall}} = \sum_{ikj} J_{\text{Hall}}^{(ikj)}. \qquad (13)$$

The formal definition of contribution $J_{\text{Hall}}^{(ikj)}$ is as follows. We consider the system where the magnetic flux exists only in the triad ikj, but the other properties are the same as in initial system. The Hall current in this modified system is equal to $J_{\text{Hall}}^{(ikj)}$. Note that contributions $J_{\text{Hall}}^{(ikj)}$ depend not only on the properties of the triad itself, but also on its position, with respect to the percolation cluster. Contributions $J_{\text{Hall}}^{(ikj)}$ have an exponentially broad distribution and their sum is assumed to be controlled by a small number of the largest $J_{\text{Hall}}^{(ikj)}$. The idea of the percolation analysis is to identify these largest contributions and neglect all other ones that are exponentially small compared to $\max(J_{\text{Hall}}^{(ikj)})$.

Let us start from the properties of the triad itself. The currents that flow through the triad can be described as three currents $\mathcal{J}_{ij}$, $\mathcal{J}_{ik}$, and $\mathcal{J}_{kj}$ [Fig. 1(a)]. The effects of the currents are additive. Let us consider one of them $\mathcal{J}_{ij}$ and
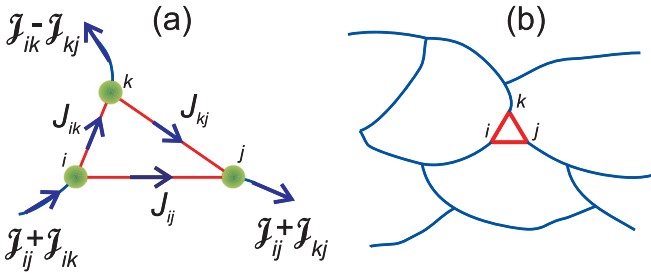
FIG. 1. (a) The currents in triangle ikj, (b) the optimal position of the triangle ikj in the percolation cluster.

assume $\mathcal{J}_{ik} = \mathcal{J}_{kj} = 0$. Current $\mathcal{J}_{ij}$ enters site i, then it is divided between J_{ij} and $J_{ik} = J_{kj}$. Then, it flows through the site j out of the triad.

According to Eq. (10), the site k in the magnetic field acquires a perturbation to its potential that is proportional to $\mathcal{J}_{ij}$:

$$\delta\varphi_k = \mathcal{L}_{kij}\mathcal{J}_{ij}, \quad \mathcal{L}_{kij} = \frac{R_{ij}R_{ik}R_{kj}}{R_{ij} + R_{ik} + R_{kj}}\frac{\mathbf{S}_{ikj}\mathbf{B}}{T}\Gamma_{ikj}. \quad (14)$$

Here, $\mathcal{L}_{kij} = \mathcal{L}_{ijk} = \mathcal{L}_{jki}$. Value $\mathcal{L}_{kij}\mathcal{J}_{ij}$ can be considered as a source of Hall voltage. $\mathcal{L}_{kij}$ values have an exponentially broad distribution

$$\mathcal{L}_{kij} \approx \mathcal{L}_0\frac{\mathbf{S}_{ikj}\mathbf{B}}{T}\exp(\xi_{ikj}), \quad (15)$$

$$\xi_{ikj} = \xi_{ij} + \xi_{ik} + \xi_{kj} - \max(\xi_{ij},\xi_{ik},\xi_{kj}) - \frac{r_{ij} + r_{jk} + r_{ik}}{a} - \min\left(\frac{\varepsilon_{ij} + \varepsilon_{ik} - |\varepsilon_i|}{T}, \frac{\varepsilon_{ij} + \varepsilon_{jk} - |\varepsilon_j|}{T}, \frac{\varepsilon_{ik} + \varepsilon_{jk} - |\varepsilon_k|}{T}\right). \quad (16)$$

Current $\mathcal{J}_{ij}$ is of the order of the "percolation current" $J_{\text{perc}} = L_{\text{cor}}j$ when resistor R_{ij} is included into the percolation cluster. Here, L_{cor} is the correlation length of the percolation cluster, j is the macroscopic current density. Therefore, reasonably large currents are possible when resistor R_{ij} is not larger than the critical resistance of percolation network R_c. Otherwise, current $\mathcal{J}_{ij}$ becomes small $\mathcal{J}_{ij} \sim (R_c/R_{ij})L_{\text{cor}}j$ for $R_{ij} > R_c$. In this case, triangle ikj cannot be an effective Hall current source.

The contribution $J_{\text{Hall}}^{(ikj)}$ of triad ikj depends on its position in the Miller-Abrahams resistor network. If it is shunted by resistances $R_{nm} \ll R_{ij}, R_{ik}, R_{kj}$, its contribution to the Hall current is small. The most effective sources are composed of resistors $R_{ij} \sim R_{ik} \sim R_{kj} \sim R_c$ and are positioned in the junction of three branches of the percolative cluster [Fig. 1(b)]. Hall mobility μ_{Hall} can be, thus, estimated as

$$\mu_{\text{Hall}} \approx p_\triangle \mathcal{L}_{kij}^{(\text{max})}/R_{\text{perc}}, \quad (17)$$

where $\mathcal{L}_{kij}^{(\text{max})}$ is the maximum possible value of $\mathcal{L}_{kij}$ for the triangles that allow the percolation current $\mathcal{J}_{ij} \sim J_{\text{perc}}$. $p_\triangle$ is the probability of three branches of the percolative cluster to be connected by the optimal triad. With the exponential precision, the mobility can be estimated as $\mu_{\text{hall}} \propto \exp[\max(\xi_{ikj}) - \xi_c]$, where ξ_c is the critical exponent of the percolation theory.

A. Nearest-neighbor hopping

In the systems with the nearest-neighbor hopping conduction, the distribution of site energies is not broad compared to the temperature. All the disorder comes from the random positions of localization sites. It is controlled by parameter na^2. We focus on the strong disorder case $na^2 \ll 1$. Therefore, we neglect the energy terms in (16) and get

$$\xi_{ikj} = \frac{r_{ij} + r_{ik} + r_{kj}}{a} - 2\frac{\max(r_{ij},r_{jk},r_{ik})}{a}. \quad (18)$$

The maximum possible value is $\xi_{ikj} = r_{\text{perc}}/a$ where $r_{\text{perc}} = \sqrt{4/\pi}n^{-1/2}$ is the percolation distance of the random site percolation problem [6]. The critical resistance in this system

is equal to $R_c = R_0\exp(2r_{\text{perc}}/a)$. The Hall mobility follows the law

$$\mu_{\text{Hall}} \propto \left(\frac{a}{r_{\text{perc}}}\right)^{\gamma_{nn}}\exp\left(-\frac{r_{\text{perc}}}{a}\right). \quad (19)$$

Here, γ_{nn} describe the power-law dependence of $p_\triangle$ on the localization radius. The conductivity depends on r_{perc} as $\sigma = \sigma_0\exp(-2r_{\text{perc}}/a)$.

For the neighbor hopping regime in the numerical simulation, we consider numerical samples without random energies of localization sites. The site concentration is equal to unity and the positions of sites are random with Poisson distribution. Localization radius a controls the disorder. Equations (10)–(12) are solved numerically for the system to find the dependence of conductivity σ and Hall mobility μ_H on the disorder parameter $n^{1/2}a$. The results are shown in Fig. 2. They are in a good agreement with the analytical predictions for $\gamma_{nn} = 3.3$.

B. Variable range hopping

In the variable range hopping (Mott law) regime, the site energies ε_i are large compared to temperature. The density of states has no peculiarities at the Fermi level and can be described by a constant $g(\varepsilon) \approx g(\varepsilon_F)$. The exponents ξ_{ij} in the resistors contain both the coordinate contribution $2r_{ij}/a$ and the energy contribution ε_{ij}/T. Therefore, the large exponent ξ_c of the critical resistor can reflect the long intersite distance between sites i and j or the large energies of sites.

The long distances between sites in the Hall source triad ikj yield exponentially small (as a function of disorder) Hall mobility μ_H as it was shown in the previous section. The situation with large energies is different. Let us imagine the triangle ikj composed of the three close sites with energies $\varepsilon \approx \xi_c T$. Equation (16) yields the Hall source exponent of this triangle $\xi_{ikj} = \xi_c$. It can be shown that it is the maximum possible value for ξ_{ikj}. It is equal to the critical exponent of the conductivity. It means that if the probability $p_\triangle$ of finding three branches of the percolation cluster connected
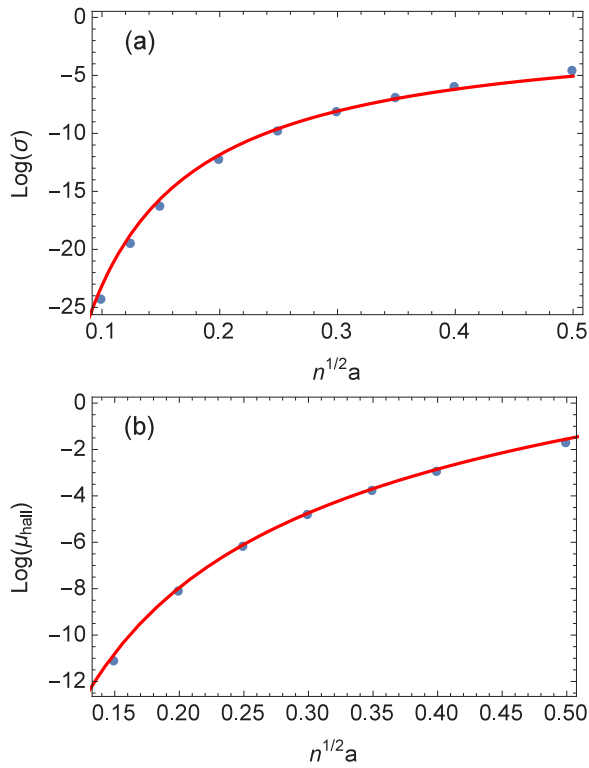
FIG. 2. The results of the numerical simulation of neighbor hopping transport. (a) The simulated conductivity (blue dots) compared to the law $\sigma = \sigma_0 \exp(-2r_{\text{perc}}/a)$ (red curve); (b) the simulated Hall mobility (blue dots) compared to Eq. (19) with $\gamma_{nn} = 3.3$ (red curve).

with such an optimal triangle has a power-law dependence on temperature, the dependence $\mu_{\text{Hall}}(T)$ should follow the same power law. The power law for the $p_\triangle(T)$ dependence is a natural assumption because the distribution of site energies and positions in the percolation cluster follow power laws. However, it will be discussed in some details in Sec. V and the counterarguments for the power law will be provided. Now, we want to note that even if the dependence $p_\triangle(T)$ follows the power law, the dominance of the optimal triangles over the Hall effect is proved only in $T \to 0$ limit. At finite temperatures it should be verified with a numerical simulation.

For the Mott-law regime in the numerical simulation we add random energies to the sites of our numerical samples. The distribution of energies has the constant density $g(\varepsilon) = n/\Delta\varepsilon$ in some energy range $-\Delta\varepsilon/2 < \varepsilon < \Delta\varepsilon/2$. The Fermi level is assumed to be equal to zero. The parameter $\Delta\varepsilon/T$ controls the energy disorder. The analytical expression for the conductivity in this regime, the Mott law, is

$$\sigma = \sigma_0 \exp[-(T_0/T)^{1/3}]. \tag{20}$$

Here, $T_0 = \beta_{2D}/g(\varepsilon_F)a^2$, $\beta_{2D} = 13.8$ [6], $g(\varepsilon_F)$ is the density of states at Fermi level ε_F. It follows from the percolation theory and is valid in the limit of strong disorder $T_0/T \gg 1$. The prefactor σ_0 can have a power-law dependence on the system parameters that can be derived from the system dimension and the power-law part of the R_{ij} dependence on the system parameters [6]. In our case of a 2D system with $R_{ij} \propto T$, σ_0 is proportional to $1/T$ and is independent of the localization radius.
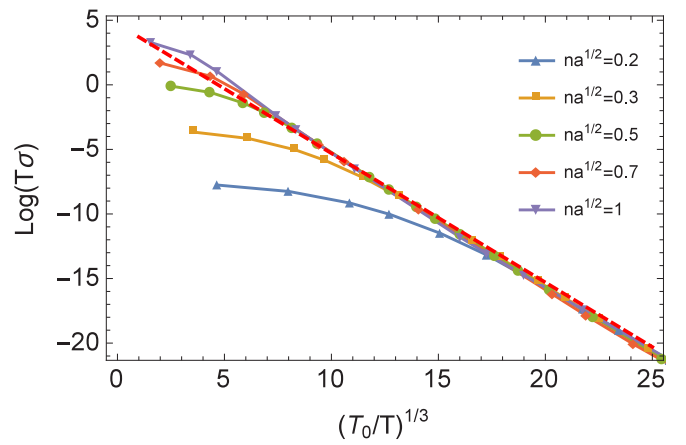

FIG. 3. The simulated temperature dependence of conductivity in the Mott-law VRH regime for values $n^{1/2}a = 0.2$, 0.3, 0.5, 0.7, and 1. The Mott law is shown with a red dashed line.

Therefore, in the VRH regime the product σT should depend only on $\xi_c = (T_0/T)^{1/3}$. The point, when the dependencies of σT on critical exponent ξ_c converge to the single curve, can be considered as a condition for the VRH conductivity. In Fig. 3(a) we show the simulated dependence of $\ln(T\sigma)$ on $(T_{\text{Mott}}/T)^{1/3}$ for different values of the position disorder parameter $n^{1/2}a$. It is compared to the Mott law (red dashed line). It seems that the agreement starts from a relatively small $\xi_c = (T_{\text{Mott}}/T)^{1/3} \approx 5$ for the weak position disorder $n^{1/2}a \lesssim 0.5$. For a stronger position disorder the VRH regime starts from larger ξ_c.

Now, let us discuss the dependencies of Hall mobility on the localization distance and temperature. In the VRH regime these dependencies are related. The transport properties that are determined by dimensionless distances r_{ij}/a and energies ε_i/T can depend only on the combined parameter $\xi_c = [\beta_{2D}/g(\varepsilon_F)a^2T]^{1/3}$ because only the states close to the Fermi energy are important in VRH. It follows from the scaling arguments. The change of the temperature and localization distance $T \to \widetilde{\beta}T$, $a \to \widetilde{\alpha}a$ is equivalent to the change of site density $n \to \widetilde{\alpha}^2 n$ and all site energies $\varepsilon_i \to \widetilde{\beta}^{-1}\varepsilon_i$. When $\widetilde{\alpha}^2\widetilde{\beta} = 1$ (it is the condition for $\xi_c = $ const) the density of states at the Fermi level $g(\varepsilon_F)$ stays constant and the g dependence on energy can be neglected in the Mott-law regime.

However, the Hall sources are proportional to BS_{ikj}/Φ_0. This value is not controlled by the r_{ij}/a ratio but with the squared intersite distance compared to Φ_0/B. The discussed scaling procedure will modify these terms $BS_{ikj}/\Phi_0 \to \widetilde{\alpha}^2 BS_{ikj}/\Phi_0$. Therefore, in the VRH regime μ_{Hall} should depend on the system parameters as follows:

$$\mu_{\text{Hall}} = \frac{a^2}{\Phi_0}f(\xi_c). \tag{21}$$

Here, $f(\xi_c)$ is a function of critical exponent ξ_c. The assumption that the Hall effect is controlled by the optimal triangles yields $f(\xi_c) = \text{const} \times \xi_c^{\gamma_{\text{Mott}}}$ with some power γ_{Mott}.

Our numerical results for Hall mobility are shown in Fig. 4. The dependencies of μ_H/a^2 on $(T_0/T)^{1/3}$ for different $n^{1/2}a$ converge to a single curve at sufficiently small temperatures. It indicates the applicability of the law (21). The curve is
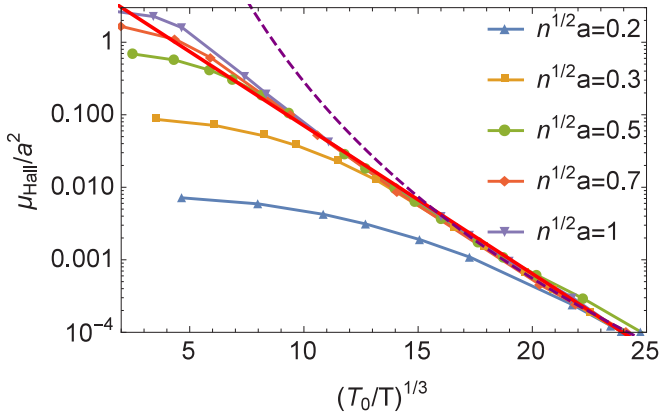
FIG. 4. The numerical results for the Hall mobility. The results are compared to Eq. (22) (red straight line) and the power-law dependence (purple dashed line).

compared to two laws that were discussed in the previous studies. The red straight line corresponds to the exponential dependence

$$\mu_{\text{Hall}} \propto a^2 \exp[-\alpha_{\text{Mott}}(T_0/T)^{1/3}]. \qquad (22)$$

The numerical estimate is $\alpha_{\text{Mott}} = 0.47$. The purple dashed line corresponds to the power-law dependence $\mu_{\text{Hall}}/a^2 \propto (T_0/T)^{\gamma_{\text{Mott}}/3}$. The agreement with the exponential dependence is better. However, at small temperatures, when the curves for small $n^{1/2}a$ converge to the universal dependence, the result can, in principle, be described with the power law, at least for the considered values of T_0/T.

To independently test the optimal triangle assumption, we discuss the average area of the triangles responsible for the Hall effect. The percolation theory predicts that the effect is controlled by the following triangles. At a high temperature and strong position disorder $n^{1/2}a \ll 1$ (NNH regime), the optimal triangle is the equilateral triangle with side r_{perc}. Its area is equal to $(\sqrt{3}/4)r_{\text{perc}}^2 \approx 0.62n^{-1}$. At a small temperature in the VRH regime, the area of the optimal triangle can be estimated as a^2. Note that it is smaller than $(\sqrt{3}/4)r_{\text{perc}}^2$.

The linear nature of Kirchhoff equations allows us to access the area of the optimal triangle in the numerical simulation. Equation (13) states that the Hall current can be described as a sum of the contribution related to each triangle. Although the contributions themselves cannot be easily separated in the final results of simulation, we can make the following numerical experiment. We artificially multiply the value $\mathcal{L}_{ikj}$ for each Hall source to the absolute value of the area of the corresponding triangle $|S_{ikj}|$. Then, we recalculate the Hall current and obtain its new value $J_{\text{Hall}}^{(\text{mod})}$. Finally, we divide the modified Hall current to the original Hall current and obtain the area $\langle S \rangle_{\text{Hall}}$:

$$\langle S \rangle_{\text{Hall}} = \frac{J_{\text{Hall}}^{(\text{mod})}}{J_{\text{Hall}}} = \frac{\sum_{ikj} |S_{ikj}| J_{\text{Hall}}^{(ikj)}}{\sum_{ikj} J_{\text{Hall}}^{(ikj)}}. \qquad (23)$$

It is the area of the triangle averaged with weight $J_{\text{Hall}}^{(ikj)}$, the contribution of the triangle to the total Hall current.

The numerical results for $\langle S \rangle_{\text{Hall}}$ are shown in Fig. 5. At high temperatures $\langle S \rangle_{\text{Hall}}$ is slightly larger than $(\sqrt{3}/4)r_{\text{perc}}^2$
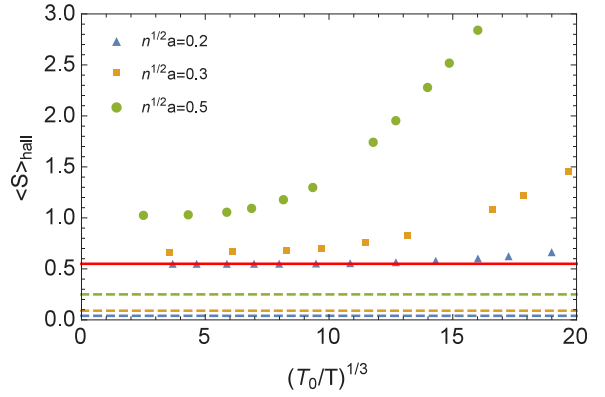

FIG. 5. The averaged area of the triangle responsible for the Hall effect. The red straight line corresponds to value $(\sqrt{3}/4)r_{\text{perc}}^2$. Dashed lines correspond to the percolation theory predictions in the VRH regime.

and tends to this value for a strong position disorder $n^{1/2}a \ll 1$. But, upon decreasing temperature, it increases instead of decreasing to its VRH percolation value. It means that, at the considered parameters, the Hall current is dominated by the triangles that are much larger than the optimal triangle of the VRH percolation theory.

Another result of the percolation theory that we want to test with the numerical simulation is the prediction of very strong mesoscopic effects for the Hall current in VRH [9]. This prediction was based on the concept of optimal triangles. If the Hall effect is controlled by the rare triads of sites, the correlation length of the Hall effect should be proportional to the distance between these triads and be much larger than the correlation length of the percolation cluster. Therefore, it was suggested that even relatively large samples can show strong mesoscopic effects.

It is also important to study mesoscopic effects to verify the applicability of our numerical results. The real samples are usually larger than our numerical samples. In Fig. 6(a) we compare the results for the numerical samples with size 70×70 considered in the rest of this study with the results for smaller 50×50 numerical samples containing 2500 sites. If the correlation length for the Hall effect is larger than our numerical samples, one should expect significant difference in the Hall mobility calculated for different system size. However, the results for 50×50 and 70×70 systems are in a good agreement up to the smallest considered temperatures.

In Fig. 6(b) we show the standard deviation of the logarithm of normal and Hall currents, $\sigma(\ln(J))$ and $\sigma(\ln(J_{\text{Hall}}))$, correspondingly. When this deviation is small $\sigma(\ln(J)) \ll 1$, the fluctuations of the current are much smaller than the average current and the system is macroscopic. The opposite case $\sigma(\ln(J)) \gg 1$ corresponds to the exponentially broad distribution of currents. This result is expected for the systems with the hopping transport that are smaller than the correlation length. The calculated standard deviation of $\ln(J_{\text{Hall}})$ is slightly larger than the one for the normal current, but the difference is not dramatic. Both standard deviations are less than unity for the considered system parameters. It means that the mesoscopic effects for the Hall current should be only slightly larger than the ones for the ordinary conductivity and

FIG. 7. The density of site energies in one numerical sample obtained with the zero-temperature Monte Carlo simulation.
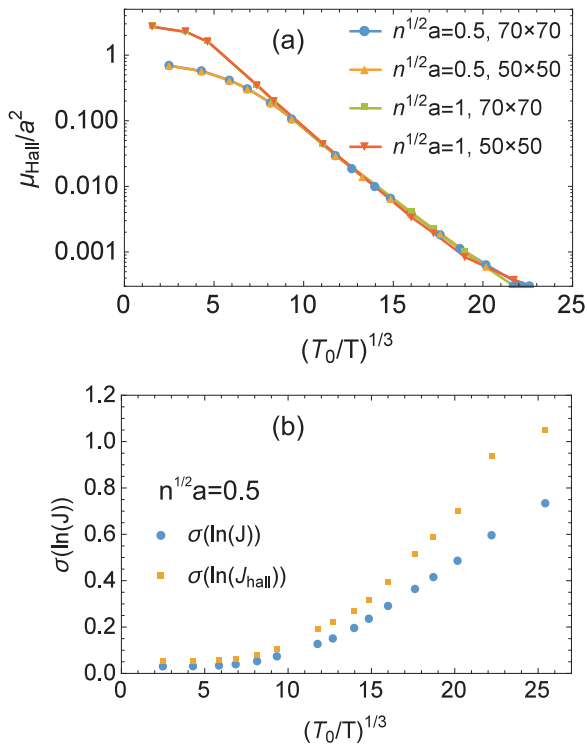
FIG. 6. (a) The comparison of the Hall mobility in the numerical simulation with different sample sizes. (b) The standard deviation of the logarithm of normal and Hall currents.

our numerical samples are larger than the correlation length for Hall effect at least for considered system parameters. Note that the size of our numerical samples (4900 sites) is still small, compared to the most real samples that are studied experimentally.

C. Variable range hopping with the Coulomb gap

Our experimental results are obtained in the samples that demonstrate the variable range hopping conductivity in the Efros-Shklovskii regime $\sigma \propto \exp[-(T_{ES}/T)^{1/2}]$. In this regime the Coulomb interaction between hopping electrons becomes essential. Strictly speaking, the Miller-Abrahams network cannot be rigorously derived for this case. Nevertheless, many important results for this regime are obtained by considering the resistor network where the Coulomb correlations were included as the Coulomb gap. In the present section we include the Coulomb gap to our system of Kirchhoff equations. We suggest, however, that our results can be dependent on the probability to find a triangle of critical resistors with special relations between distances and energies (as it is predicted by the percolation theory). The positions of sites in the Coulomb gap are correlated. Therefore, to keep these correlations in our system, we do not simply ascribe each site a random energy with a distribution that includes the Coulomb gap. We follow a more complex procedure. We start with a numerical sample with random positions of the sites and consider a random half of them to be filled with electrons. Then, we run the zero-temperature Monte Carlo algorithm, i.e., we resolve all one-electron hops that decrease the total energy of the system including the electron-electron
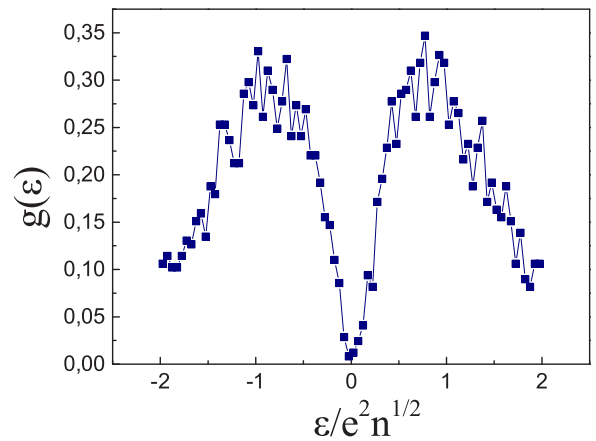
Coulomb repulsion. The details of this algorithm are given in Ref. [32]. It yields a metastable state of the system that naturally includes the Coulomb gap and the correlations in positions of sites with the energies close to the Fermi level. In the obtained state, we find all one-electron energies and substitute them to the expressions for the Miller-Abrahams resistors and Hall sources. The rest of the calculation is the same as in the regime of the Mott law. Our approach allows us to consider the "static" Coulomb correlations, however, it disregards the dynamic correlations, i.e., the modification of energies ε_i due to the electron hops. Nevertheless, it allows the consideration of relatively large numerical samples deep in the VRH regime, which are not easy to access with the finite-temperature Monte Carlo algorithm (that includes all the dynamic Coulomb correlations).

The distribution of the site energies obtained with our method in a single numerical sample is shown on Fig. 7. It shows that, even in a single numerical sample, the Coulomb gap is well defined. The results of simulation in the Efros-Shklovskii regime are shown in Fig. 8. They are qualitatively similar to the results in the Mott-law regime, however, the dependencies of $T\sigma$ and μ_{Hall} on ξ_c converge to a universal curve more slowly than in the regime of the Mott law. We suggest that the reason for it is the double transition: from nearest-neighbor hopping to VRH and from Mott VRH to Efros-Shklovskii VRH. The dependence of Hall mobility on the localization radius and temperature follows the law

$$\mu_{Hall} \propto a^2 \exp\left[-\alpha_{ES}\left(\frac{T_{ES}}{T}\right)^{1/2}\right], \qquad (24)$$

where T_{ES} is the temperature from the Efros-Shklovskii law. It is proportional to a^{-1}. The numeric value for α_{ES} is $\alpha_{ES} = 0.3$.

IV. COMPARISON WITH EXPERIMENT

Our theoretical investigation indicates that Hall effect can be detected more easily in systems with large localization radius. The range of critical exponent values ξ_c accessible in the experiment is limited by the exponential growth of system resistance at large ξ_c. For reasonable values of the critical
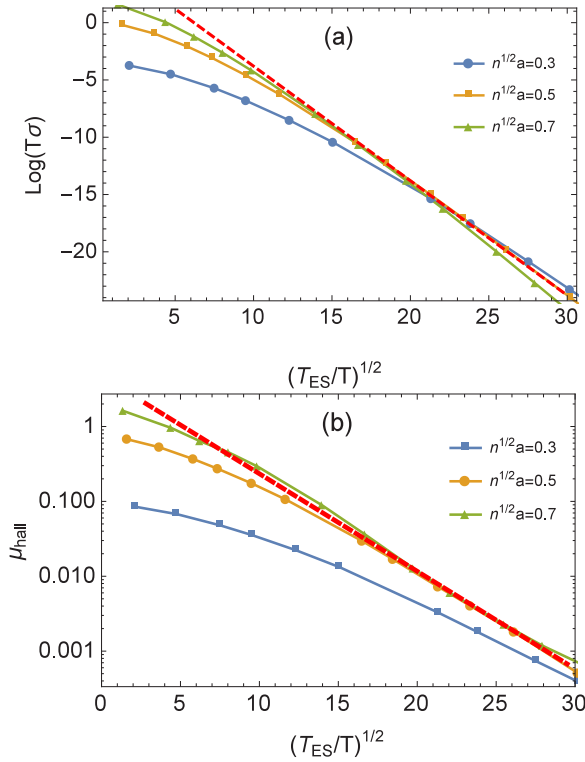
FIG. 8. The results of numerical simulation of variable range hopping transport in the Efros-Shklovskii regime. (a) The conductivity compared to the Efros-Shklovskii law. (b) The Hall mobility compared to the expression (24).

FIG. 9. The energy levels of a single quantum dot (a) and the density of states $g(\varepsilon)$ (b). The maxima of $g(\varepsilon)$ correspond to the quantum dot levels. The levels are broadened due to the random potential.

exponent, the temperature dependence of the Hall mobility converges to the universal VRH curve only for relatively large $n^{1/2}a$. For small $n^{1/2}a$, it converges at very large ξ_c. Finally, at the same values of the critical exponent, the system with large a should display a larger Hall effect $\propto a^2$ due to the scaling arguments.

As we discussed it in Introduction, several experimental measurements of the Hall effect were obtained in 3D systems in the vicinity of the metal-insulation transition where the localization radius diverges. Deep in the strong localization regime, where the localization radius is small, the Hall effect was usually not visible. In this section we compare the theory with our recent experimental measurements of the Hall effect in the p-doped two-dimensional arrays of tunnel-coupled Ge/Si quantum dots (QDs). These arrays display the VRH conductivity with the localization radius much larger than its typical value in doped semiconductors. The main parts of the presented experimental results were preliminarily published in Ref. [29].

The QD arrays were grown with the low-temperature (about of 300 °C) molecular beam epitaxy. The small size of a quantum dot (lateral size 15–20 nm and 1.5–2 nm high) leads to a large energy separation of quantum levels $\gtrsim 10$ meV and to the nonmonotonic density of states (Fig. 9). The dominant mechanism of the transport in the discussed arrays is the variable range hopping between quantum dots. It was shown with the temperature dependencies of conductance to follow the Efros-Shklovskii law $\sigma(T) = \gamma T^m \exp[-(T_{ES}/T)^{1/2}]$. Value m was close to zero [33].
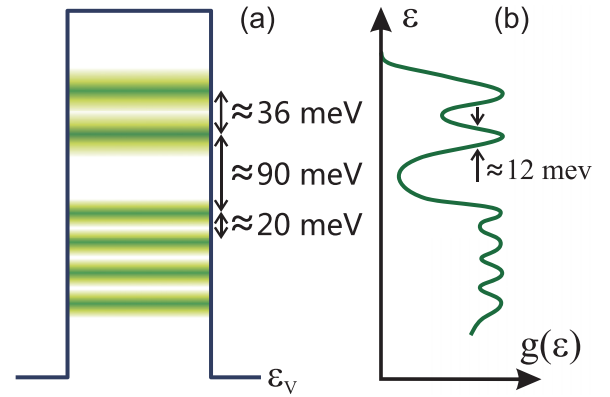
The mechanism of the variable range hopping transport in quantum dot arrays is slightly different from the VRH mechanism in doped semiconductors. The tunneling path to a distant quantum dot inevitably crosses other (intermediate) dots. The hop to a distant dot includes the cotunneling process involving the states in the intermediate dots. This process is schematically shown in Fig. 10. The hole from site i cannot hop to site j because of the large energy difference $|\varepsilon_i - \varepsilon_j| \gg T$. Instead, it hops on the distant site k with $\varepsilon_k \sim \varepsilon_i$. The process involves the state on quantum dot j as the intermediate virtual state. The tunneling amplitude for hop $i \rightarrow k$ can be estimated as $\widetilde{I}_{ik} = I_{ij}I_{jk}/(\varepsilon_i - \varepsilon_j)$. Here, I_{ij} is the overlap integral between the states on quantum dots i and j. The cotunneling can involve any number of intermediate dots.

The transport in QD arrays due to the cotunneling processes can be described with the conventional variable range hopping theory [31]. However, the localization radius a in the theory is not the localization radius of a single QD, but
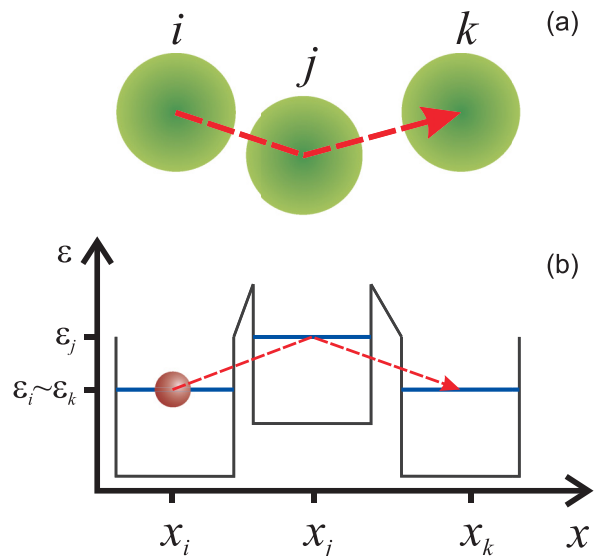

FIG. 10. Cotunneling between distant dots i and k. (a) The real-space positions of dots i, j, and k; (b) the energy diagram where x is the generalized coordinate.
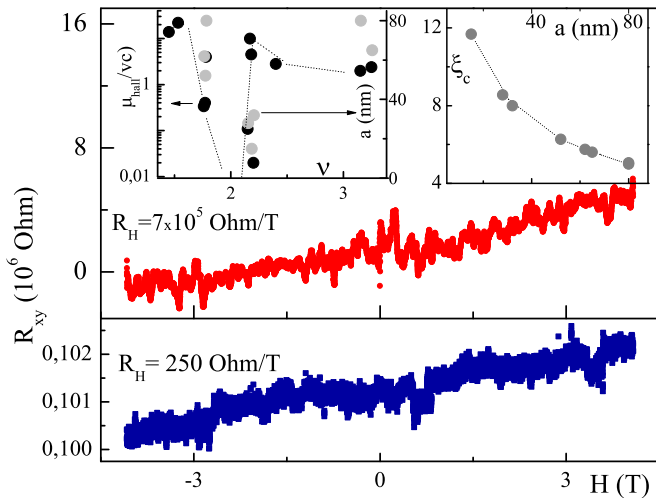
FIG. 11. $R_{xy}(H)$ dependence for two high-resistance samples. Left inset: the dependence of mobility (black symbols) and localization radius (gray symbols) on filling factor. Right inset: $\xi_c(a)$ dependence.

FIG. 12. The comparison of experimental results with Eq. (24). The red curve corresponds to α_{ES} treated as the fitting parameter, while for the blue curve we adopt the value $\alpha_{ES} = 0.3$ obtained from the numerical simulation.

it is strongly modified by the virtual states on the intermediate dots. This modified localization radius can easily be comparable to or larger than the size of a quantum dot. It depends on the density of states at the Fermi level [29,34,35]. It allows us to control a by changing the QD filling factor ν and the Fermi level position with respect to the density of states.

We measured the Hall effect in the structures with different dot filling factors, which were varied by changing the boron concentration in the δ-boron-doped silicon layer which is 5 nm below the quantum dot layer. The conductance values for the samples under study were shown to be in the range $\sim 10^{-5}$–10^{-11} Ohm/□ at 4.2 K. From the comparison of the measured $\sigma(T)$ dependence with the Efros-Shklovskii law we determined the critical exponent $\xi_c = (T_{ES}/T)^{1/2}$ and localization radius $a = Ce^2/\epsilon k_B T_{ES}$. Here, ϵ is the permittivity, k_B is the Boltzmann constant, and C is a numerical coefficient that, according to Ref. [36], is equal to 6.2. It was shown that the localization radius changes from ~ 25 to ~ 80 nm depending on the filling factor and, correspondingly, the Fermi level position.

In Fig. 11 we show R_{xy} for two high-resistance samples. The Hall coefficients R_H were determined from the slopes of $R_{xy}(H)$ lines as $R_H = [R_{xy}(H) - R_{xy}(-H)]/2H$. It allows avoiding the symmetrical contribution of magnetoresistance due to the asymmetry of the contacts and a possible shift of the amplifier's zero. We observed a strong correlation between $\mu(\nu)$ and $a(\nu)$ nonmonotonic dependencies that are obviously observed in the left inset to Fig. 11.

In Fig. 12 we compare the experimental results to Eq. (24). The blue curve corresponds to the parameter $\alpha_{ES} = 0.3$ obtained from the numerical simulation. The red curve corresponds to α_{ES} treated as the fitting parameter. Expression (24) agrees to the experimental data. The best agreement is achieved for $\alpha_{ES} = 0.56$. However, the precision of experimental measurements is insufficient to reliably prove this value. Let us note that there are physical reasons for the measured Hall conductivity to be different from the predictions
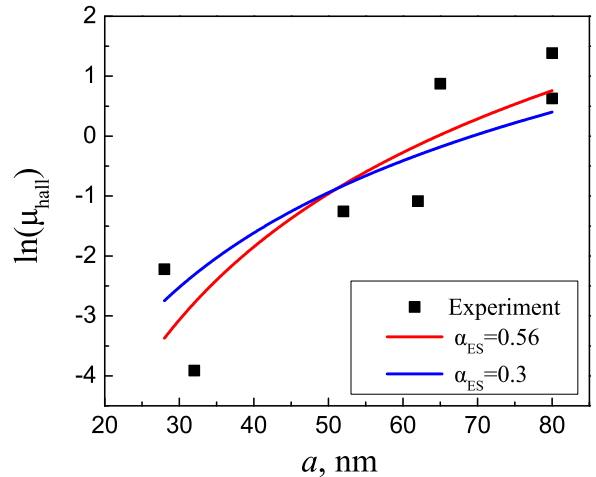
of the theory. The complex nature of long-range hopping in QD arrays can modify the dependence $\mu_{Hall}(a)$ obtained from the pointlike site model. However, further experimental investigations are needed to understand if it is the case.

V. DISCUSSION

The conventional approach to the hopping transport in strongly disordered systems is the percolation theory that is based on the assumption that all the exponentially small terms can be neglected at a sufficiently strong disorder. This assumption leads to the result that the Hall effect is controlled by the rare optimal triads of sites. In the NNH regime, they are the triads that form equilateral triangles with side r_{perc} corresponding to the percolation resistance. The triangles should be positioned in the intersections of three branches of the percolation cluster. In the VRH regime, the restrictions to the optimal triangles are even more solid. They, still, are the triangles consisting of critical resistors. However, now they should include only the critical resistors of the specific kind: the ones with a small length and large energies.

Our numerical results show that the model of optimal triads works well in the NNH regime. The dependence (19) obtained from the optimal triangle model agrees to the numerical results. Moreover, the dominant area of the triangle $\langle S_{ikj} \rangle_{Hall}$ calculated at a high temperature is in a quantitative agreement with the prediction of the NNH percolation theory $(\sqrt{3}/4)r_{perc}^2$. However, the optimal triangle model fails to describe the VRH case. The calculated temperature dependence of μ_{Hall} follows the exponential laws (22) and (24) instead of the power law predicted by the percolation theory. The clearest evidence of the failure of optimal triangle model is the temperature dependence of $\langle S_{ikj} \rangle_{Hall}$. The area of the optimal triangle is smaller in the VRH regime than in NNH. However, $\langle S_{ikj} \rangle_{Hall}$ grows with the decreasing temperature, indicating that the Hall effect is controlled by the triangles that are larger even than the optimal triangle of the NNH regime. To our opinion, it indicates that the Hall effect in the

VRH regime is not dominated by rare optimal triangles, but by more numerous "typical" triangles. The area of the typical triangle increases with the decreasing temperature because more distant hops become important in the VRH regime.

Why does the optimal triangle model work well in one regime and fail in the other? Here, we argue that the reason is the extremely low probability $p_\triangle$ to find the optimal triangle of the VRH regime. First, the probability for three branches of the percolation cluster to be connected with a triangle of critical resistors is already small. Each branch of the percolation cluster contains at least one critical resistor. However, it can be positioned in any place of the branch, not necessarily at its edge. The branch of the percolation cluster contains $\propto \xi_c$ resistors. Therefore, we estimate the probability to find a critical resistor at the edge of the branch as $1/\xi_c$. The triangle of critical resistors at the intersection of three branches can appear when all the three branches contain critical resistors on their edges, therefore, its probability can be estimated as $1/\xi_c^3$. This estimate can be compared to the numerical result for the preexponential part of the dependence (19), $(a/r_{\mathrm{perc}})^{\gamma_{nn}} \propto \xi_c^{-\gamma_{nn}}$. Note that the critical exponent in the NNH regime is equal to $2r_{\mathrm{perc}}/a$. The numerical value $\gamma_{nn} = 3.3$ is close to our simplified estimate 3.

In the VRH regime, the optimization of the exponent (16) leads to the following restriction to the energies of the sites composing the "optimal triad." Two energies should lie in the interval $(\varepsilon_F + (\xi_c - 1)T, \varepsilon_F + \xi_c T)$, while the third one can have any energy larger than ε_F [there is the second possible option, when the three sites have the energies lower than ε_F and the energies of two sites are in the interval $(\varepsilon_F - \xi_c T, \varepsilon_F - (\xi_c - 1)T)$; this option leads to the similar results]. If we presume that the distribution of site energies in the critical resistor is flat, it will add an additional factor ξ_c^{-2} to $p_\triangle$ leading to the estimate $p_\triangle \propto \xi_c^{-5}$. We argue, however, that this expression overestimates $p_\triangle$. There are two reasons for it. The first reason is discussed in Ref. [13] and is related to the small probability of finding a site with a maximum possible energy in the percolation cluster.

Here, we want to discuss the second reason that has, to the best of our knowledge, been never discussed. It is related to the fact that the "optimal triangle" should be connected to the rest of the percolation cluster, and it is not easy for the considered energies of its sites. In Fig. 13 we show the optimal triangle with the nearest sites of three branches of the percolation cluster. For the sake of the qualitative estimate we consider the branches to be straight lines with angle $2\pi/3$ between them. We assume that the two sites of the triangle with energies $\sim \xi_c T$ are i and j. The resistors $i - i_1$ and $j - j_1$ are part of the percolation cluster. Therefore, $\varepsilon_{i,i1}/T + 2r_{i,i1}/a \lesssim \xi_c$ and $\varepsilon_{j,j1}/T + 2r_{j,j1}/a \lesssim \xi_c$. The energies in these relations are large $\varepsilon_{i,i1} \gtrsim \varepsilon_i - \varepsilon_F \sim \xi_c T$, $\varepsilon_{j,j1} \gtrsim \varepsilon_j - \varepsilon_F \sim \xi_c T$. It means that distances $r_{i,i1}$ and $r_{j,j1}$ should be small, of the order of $a/2$. However, resistors $i_1 - j_1$, $i_1 - k$, and $j_1 - k$ should not shunt resistors $i - j$, $i - k$, and $j - k$. It imposes serious restrictions on energies ε_{i1} and ε_{j1}:

$$(\sqrt{3} + 1) + \max(\varepsilon_{i1} - \varepsilon_F, \varepsilon_{j1} - \varepsilon_F)/T \gtrsim \xi_c,$$

$$\sqrt{3} + \max(\varepsilon_{i1} - \varepsilon_F, \varepsilon_k - \varepsilon_F)/T \gtrsim \xi_c, \quad (25)$$

$$\sqrt{3} + \max(\varepsilon_{j1} - \varepsilon_F, \varepsilon_k - \varepsilon_F)/T \gtrsim \xi_c.$$
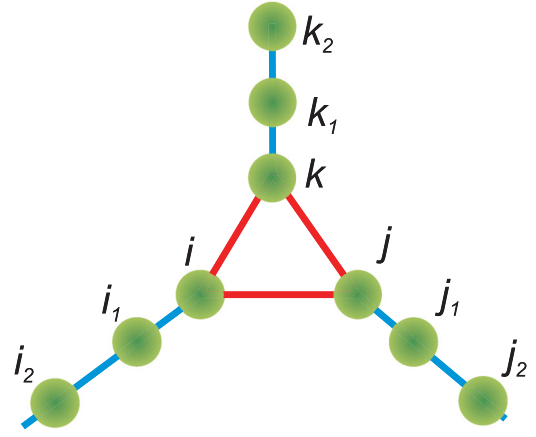

FIG. 13. The "optimal triad" with its connection to the nearest sites of the percolation cluster. The critical resistors composing the triad are shown with red. Other resistors of the percolation cluster are shown with blue.

Here, we assumed that the sides of triangle ijk are equal to $a/2$. These restrictions mean that, in addition to energies ε_i, ε_j, two other energies from $\varepsilon_{i1}, \varepsilon_{j1}, \varepsilon_k$ should also be close to $\varepsilon_F + T\xi_c$. However the range of possible energies $\sim(\varepsilon_F + \xi_c T - (\sqrt{3} + 1)T, \varepsilon_F + \xi_c T)$ is larger than the one for ε_i, ε_j. If the critical exponent ξ_c is large compared to $\sqrt{3} + 1$, another similar argument leads to the restrictions for the energies of other sites ε_{k1}, ε_{i2}, ε_{j2}, etc. It means that the "optimal triangle" of the VRH percolation theory is actually quite a sophisticated and improbable complex of sites that allow the connection of the triangle to the rest of the percolation cluster. The number of the sites in this complex grows with ξ_c leading to the dependence $p_\triangle(\xi_c)$ that is stronger than any power law.

However, even if we consider only the restrictions (25), the apparent dependence $p_\triangle(\xi_c)$ is $p_\triangle \propto \xi_c^{-7}$. These small probabilities should be compared to the contribution of the nonoptimal triangles that can be evaluated as $\exp(-\alpha_{\mathrm{Mott}}\xi_c)$ from Eq. (22) with $\alpha_{\mathrm{Mott}} = 0.47$. Condition $\exp(-0.47\xi_c) < \xi_c^{-7}$ yields $\xi_c > 60$. These values can hardly be accessible in experiment. Note that the dependence of sample resistance on ξ_c is $R = R_0 \exp(\xi_c)$. If $R_0 \sim 1$ Ohm, $\xi_c = 60$ leads to $R > 10^{26}$ Ohm. It means that at measurable system resistances, the optimal triads should not be important for the Hall effect in the VRH regime.

In a wide range of temperatures and localization distances the Hall mobility can be described with the law $\mu_{\mathrm{Hall}} \propto a^2 \exp(-\alpha\xi_c)$. Let us note that the range of possible ξ_c values is limited by the conditions of the VRH regime and the restriction for reasonably large conductivity. It makes the structures with the hopping transport and large localization radius a a good choice to study the Hall effect in the VRH regime. The localization radius in the structures discussed in Sec. IV is ~ 10 times larger than in typical doped semiconductors. It means that the Hall effect in these structures should be ~ 100 times larger than in ordinary semiconductors at the same ξ_c. However, the complex nature of long-range hops in these structures (that inevitably includes cotunneling) can modify the physics of Hall effect. The comparison of the present

theory with experiments in QD ensembles with the VRH transport shows that Eq. (24) describes the experimental data. Nevertheless, the α_{ES} values obtained from the experiment and from numerical simulations of the pointlike site model with the Coulomb gap seem to be different.

Finally, we want to note that the Hall effect is not the only phenomenon in the hopping transport that is related to the two-phonon hopping with interference. Recently, it was shown that the current-induced spin polarization, spin galvanic effect, and spin Hall effect also appear due to the similar processes [37]. The effects were controlled by the complex interplay of the disorder strength and spin relaxation time relation to the hopping time. However, the theory was made only for the case of position disorder, i.e., the NNH regime. We argue that the theory [37] can be reduced to the theory of ordinary Hall effect in the limit of large spin relaxation time. Therefore, our results on the optimal triangles should be important for the theory of spin generation, at least, in some limiting cases.

In conclusion, we revised the theory of the hopping Hall effect in 2D systems. We compared the predictions of the percolation theory to the numerical simulations based on the solution of modified Kirchhoff equations in different regimes. The percolation theory is in agreement with the numerical results in the neighbor hopping regime. However, in the variable hopping regime, it fails to describe the results of simulation. We argue that it is related to the extremely small probability of finding the optimal triad of the sites in the VRH regime due to the complex nature of the triad and its connection to the percolation cluster. The numerical results in the VRH regime can be described by an empirical law that agrees with our recent experimental data.

APPENDIX: DERIVATION OF THE HOPPING RATES

First, we derive the rates of the ordinary two-site hops with one phonon. It is instructive to explain our method for this simple case. The ordinary hops are described by the reduction of second-order density matrix $\langle \widehat{\rho}^{(2)} \rangle_i$. We describe here the hopping between sites i and j. It appears in the second-order perturbation over T_{ij}. It means that the states of all the sites other than i and j are not modified during the hop. Therefore, we can start from the density matrix $\widehat{\rho}^{(0)}$ already reduced over all the sites $k \neq i, j$. We express this density matrix as

$$\widehat{\rho}^{(0)} = f_i f_j a_i^+ a_j^+ |\emptyset\rangle\langle\emptyset| a_j a_i + f_i (1-f_j) a_i^+ |\emptyset\rangle\langle\emptyset| a_i$$
$$+ f_j (1-f_i) a_j^+ |\emptyset\rangle\langle\emptyset| a_j + (1-f_i)(1-f_j)|\emptyset\rangle\langle\emptyset|. \tag{A1}$$

Here, $|\emptyset\rangle\langle\emptyset|$ describes the "vacuum" state when sites i and j are free and phonons are in the equilibrium.

The hopping is possible only for the second and the third terms in (A1), i.e., when the two sites have one electron. In

this case, one can use one-electron notation $|i\rangle = a_i^+ |\emptyset\rangle$. Let us consider the hopping from site j to site i. To find its rate we should consider the term $f_j(1-f_i)|j\rangle\langle j|$ and find the contribution of $\langle\rho^{(2)}\rangle_i$ to the state with filled site i. It allows one to give the rate of hopping $j \to i$ as

$$W_{ij} = -\frac{1}{\mathcal{T}} \text{Tr}_{\text{ph}} \int_{-\mathcal{T}}^{\mathcal{T}} dt_1 \int_{-\mathcal{T}}^{t_1} dt_2 \langle i|$$
$$\times [T_{ij}(t_1), [T_{ij}(t_2), |j\rangle\langle j|]]|i\rangle. \tag{A2}$$

Here, Tr_{ph} is the trace over the final states of the phonon subsystem. $\mathcal{T}$ is some time interval that is large compared to $\varepsilon_i - \varepsilon_j$. The expression under the integral depends only on the difference $t_1 - t_2$. Therefore, the average time $(t_1 + t_2)/2$ can be integrated out and it is canceled with the time interval $\mathcal{T}$. The operator $T_{ij} = t_{ij} a_i^+ a_j + t_{ji} a_j^+ a_i$ corresponds to the transition of the electron between sites i and j with the simultaneous emission or absorbtion of some number of phonons. Let us denote the full energy of the system when the electron is on site j as E_j. Different E_j values are possible due to the fluctuations of phonon numbers.

$$W_{ij} = t_{ij} t_{ji} \text{Tr}_{\text{ph}} \left(\Phi_{ij} \rho_{\text{ph}}^{(0)} \Phi_{ji} \int e^{i(E_i - E_j)t'} dt' \right). \tag{A3}$$

The integral over time $t' = t_1 - t_2$ yields the delta function $\delta(E_i - E_j)$ indicating the energy conservation during the hop. The term $\text{Tr}_{\text{ph}} \Phi_{ij}(t_1) \rho_{\text{ph}}^{(0)} \Phi_{ji}(t_2)$, where $\rho_{\text{ph}}^{(0)}$ is the equilibrium phonon density matrix contains the electron-phonon interaction and the probability of finding the phonons required for the hop. If we consider the electron-phonon interaction to be small, the one-phonon processes will dominate the hop, and we find

$$\text{Tr}_{\text{ph}} \left[\delta(E_i - E_j) \Phi_{ij} \rho_{\text{ph}}^{(0)} \Phi_{ji} \right]$$
$$= 2|\gamma(q)|^2 g^{(\text{ph})}(q) \left(\theta(\varepsilon_j - \varepsilon_i) + \frac{1}{e^{|\varepsilon_j - \varepsilon_i|/T} - 1} \right). \tag{A4}$$

Here, $g^{(\text{ph})}(q)$ is the density of states for the phonons with $q = |\varepsilon_i - \varepsilon_j|/c$, where c is the sound velocity. $|\gamma(q)|^2$ describes the strength of the electron-phonon coupling at the given absolute value of wave vector q. In the case of large energies, $\varepsilon_i, \varepsilon_j \gg T$, Eq. (A4) can be estimated as

$$\text{Tr}_{\text{ph}} \left[\delta(E_i - E_j) \Phi_{ij} \rho_{\text{ph}}^{(0)} \Phi_{ji} \right] \approx \frac{1}{2\pi t_0^2 \tau_1} \mathcal{N}(\varepsilon_j - \varepsilon_i),$$
$$\mathcal{N}(\varepsilon_j - \varepsilon_i) = \begin{cases} 1, & \varepsilon_j > \varepsilon_i \\ e^{(\varepsilon_j - \varepsilon_i)/T}, & \varepsilon_j < \varepsilon_i. \end{cases} \tag{A5}$$

Here, τ_1 is the time constant associated with normal hops. We neglect the possible power-law dependence of τ_1 on the site energies in comparison with the strong dependence $\mathcal{N}(\varepsilon_j - \varepsilon_i)$. The normal hopping rate can be then expressed as

$$W_{ij} = \frac{1}{\tau_1} \exp(-2r_{ij}/a) \mathcal{N}(\varepsilon_j - \varepsilon_i). \tag{A6}$$

The quantum mechanics of the hop is illustrated in the diagram shown in Fig. 14. The two solid lines correspond to the two electron density matrix indexes. The interaction with phonons is shown with the ring. Note that, after the polaron transformation, any number of dashed phonon lines can enter
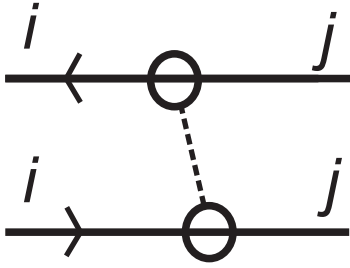
FIG. 14. One-phonon processes that lead to ordinary hops.

the vertex. The energy should be conserved in the vertex. The phonon line corresponds to the factor $\mathcal{N}(\varepsilon_j - \varepsilon_i)$.

The Hall effect in the hopping transport cannot be described with the ordinary hops (A6). Its description should include the hops $j \to i$ that occur in the presence of the third site k. Let us first consider the situation when site k is free. The corresponding hopping rate $W_{ikj}^{(0)}$ can be expressed as

$$W_{ikj}^{(0)} = i \frac{1}{\mathcal{T}} \mathrm{Tr}_{\mathrm{ph}} \int_{t_1 > t_2 > t_3} \langle i |$$
$$\times [T_{ikj}(t_1), [T_{ikj}(t_2), [T_{ikj}(t_3), |j\rangle\langle j|]]] |i\rangle. \quad (A7)$$

Here, $T_{ikj} = T_{ik} + T_{ij} + T_{jk}$. Although the commutators in (A7) contain quite a number of terms, only some of them lead to the final state of the electron on site i and yield the nonzero contribution to $W_{ikj}^{(0)}$. We separate the relevant terms

$$W_{ikj}^{(0)} = i \frac{1}{\mathcal{T}} \mathrm{Tr}_{\mathrm{ph}} \int_{t_1, t_2 > t_3} \langle i |(T_{ij}(t_1)|j\rangle\langle j|T_{jk}(t_3)T_{ki}(t_2)$$
$$- T_{ik}(t_2)T_{kj}(t_3)|j\rangle\langle j|T_{ji}(t_1))|i\rangle. \quad (A8)$$

Note the difference in the time integration between expressions (A7) and (A8). In (A8) the relation between t_1 and other times is arbitrary. The time integration in the first term of expression (A8) yields

$$2\pi \delta(E_i - E_j) \left(\frac{-i}{E_j - E_k} + \pi \delta(E_j - E_k) \right). \quad (A9)$$

The result of the time integration in the second term is the complex conjugate of (A9). Let us note that the contribution of the terms including $\delta(E_i - E_j)/(E_j - E_k)$ to the hopping rate is proportional to $t_{ij}t_{jk}t_{ki} + t_{ik}t_{kj}t_{ji} = 2\,\mathrm{Re}(t_{ij}t_{jk}t_{ki})$. This contribution has only the quadratic dependence on the applied magnetic field. Although this contribution determines the interference magnetoresistance, in the theory of the Hall effect that is linear on the applied field, we neglect this term.
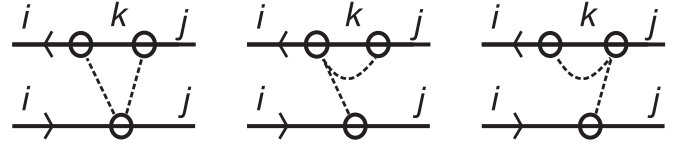

FIG. 15. Two-phonon processes that lead to the Hall effect.

The part of $W_{ikj}^{(0)}$ that is responsible for the Hall effect can be expressed as

$$W_{ikj}^{(0)} = -4\pi^2 \mathrm{Im}(t_{ij}t_{jk}t_{ki})\mathrm{Tr}_{\mathrm{ph}}$$
$$\times \left[\delta(E_i - E_j)\delta(E_j - E_k)\Phi_{ij}\rho_{\mathrm{ph}}^{(0)}\Phi_{jk}\Phi_{ki} \right]. \quad (A10)$$

The two delta functions in (A10) mean that the process should include at least two phonons. There are three two-phonon processes that allow the "energy conservation laws" $\delta(E_i - E_j)\delta(E_j - E_k)$. These processes are shown schematically in Fig. 15. These three processes lead to different phonons participating in the hop and, correspondingly, to different terms $\mathcal{N}$.

The sum over possible processes leads to the following expression for $W_{ikj}^{(0)}$:

$$W_{ikj}^{(0)} = \frac{1}{4}|t_{ij}t_{jk}t_{ki}| \frac{\mathbf{BS}_{ikj}}{2\Phi_0} \left(\frac{W_{kj}W_{ik}}{|t_{kj}^2 t_{ik}^2|} + \frac{W_{ij}W_{kj}}{|t_{ij}^2 t_{kj}^2|} + \frac{W_{ij}W_{ki}}{|t_{ij}^2 t_{ik}^2|} \right). \quad (A11)$$

Now, let us discuss the situation when site k is filled before the hop. In this case, we start from state $|kj\rangle = a_k^+ a_j^+ |\emptyset\rangle$. Let us note that operators a_k^+ and a_j^+ anticommute and, in our notations, $|kj\rangle = -|jk\rangle$. The hopping that includes site k corresponds to the electron transition from site k to site j and the following electron transition of from site j to site k. The hopping rate $W_{ikj}^{(1)}$ can be expressed as

$$W_{ikj}^{(1)} = i \frac{1}{\mathcal{T}} \mathrm{Tr}_{\mathrm{ph}} \int_{t_1, t_2 > t_3} \langle ik |(T_{ij}(t_1)|kj\rangle\langle jk|T_{ki}(t_3)T_{jk}(t_2)$$
$$- T_{kj}(t_2)T_{ik}(t_3)|kj\rangle\langle jk|T_{ji}(t_1))|ki\rangle. \quad (A12)$$

Note the difference in the order of the transitions between expressions (A12) and (A8). In (A12) transition $j \to k$ occurs after $k \to i$. It leads to somewhat different averaged phonon numbers and to the expression

$$W_{ikj}^{(1)} = -\frac{1}{4}|t_{ij}t_{jk}t_{ki}| \frac{\mathbf{BS}_{ikj}}{2\Phi_0} \left(\frac{W_{kj}W_{ik}}{|t_{kj}^2 t_{ik}^2|} + \frac{W_{ij}W_{jk}}{|t_{ij}^2 t_{kj}^2|} + \frac{W_{ij}W_{ik}}{|t_{ij}^2 t_{ik}^2|} \right).$$
$$(A13)$$

[1] T. Holstein, Phys. Rev. **124**, 1329 (1961).

[2] V. L. Nguyen, B. Z. Spivak, and B. I. Shklovskii, Zh. Eksp. Teor. Fiz. **89**, 1770 (1985) [Sov. Phys. JETP **62**, 1021 (1985)].

[3] B. I. Shklovskii and B. Z. Spivak, in *Hopping Transport in Solids*, 1st ed., edited by M. Pollak and B. Shklovskii (Elsevier, Amsterdam, 1991), Vol. 28, p. 271.

[4] A. V. Shumilin and V. I. Kozub, Phys. Rev. B **85**, 115203 (2012).

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

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

[7] H. Bottger and V. V. Bryksin, Phys. Status Solidi B **81**, 433 (1977).

[8] L. Friedman and M. Pollak, Philos. Mag. B **38**, 173 (2006).

[9] Y. M. Galperin, E. P. German, and V. G. Karpov, Zh. Eksp. Teor. Fiz. **99**, 343 (1991) [Sov. Phys. JETP **72**, 193 (1991)].

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

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

[12] M. Gruenewald, H. Mueller, P. Thomas, and D. Wuertz, Solid State Commun. **38**, 1011 (1981).

[13] L. Friedman and M. Pollak, Philos. Mag. B **44**, 487 (2006).

[14] L. Friedman, Phys. Rev. B **25**, 3512 (1982).

[15] A. A. Burkov and L. Balents, Phys. Rev. Lett. **91**, 057202 (2003).

[16] X.-J. Liu, Xin Liu, and J. Sinova, Phys. Rev. B **84**, 165304 (2011).

[17] N. F. Mott, E. A. Davis, and R. A. Street, Philos. Mag. **32**, 961 (1975).

[18] A. S. Skal and B. I. Shklovskii, Fiz. Tekh. Poluprovodn. **8**, 1568 (1974) [Sov. Phys. Semicond. **8**, 1024 (1975)].

[19] P. N. Butcher and J. A. Mcinnes, Philos. Mag. B **44**, 595 (2006).

[20] P. N. Butcher, J. A. Mcinnes, and S. Summerfield, Philos. Mag. B **48**, 551 (2011).

[21] D. W. Koon and T. G. Castner, Solid State Commun. **64**, 11 (1987).

[22] L. Essaleh, S. M. Wasim, and J. Galibert, Mater. Lett. **60**, 1947 (2006).

[23] Y. Zhang, P. Dai, I. Kam, and M. P. Sarachik, Phys. Rev. B **49**, 5032 (1994).

[24] A. Avdonin, P. Skupinski, and K. Grasza, Phys. B (Amsterdam) **483**, 13 (2016).

[25] Y. Kajikawa, Phys. Status Solidi C **14**, 1600129 (2017).

[26] M. Amitay and M. Pollak, J. Phys. Soc. Jpn. **21**, 549 (1966).

[27] R. S. Klein, Phys. Rev. B **31**, 2014 (1985).

[28] A. I. Yakimov, A. V. Dvurechenskii, V. V. Kirienko, Yu. I. Yakovlev, A. I. Nikiforov, and C. J. Adkins, Phys. Rev. B **61**, 10868 (2000).

[29] N. P. Stepina, A. V. Nenashev, and A. V. Dvurechenskii, JETP Lett. **106**, 308 (2017).

[30] H. Bottger and V. V. Bryksin, *Hopping Conduction in Solids* (Akademie-Verlag, Berlin, 1985).

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

[32] A. V. Shumilin, Solid State Commun. **183**, 51 (2014).

[33] A. I. Yakimov, A. V. Dvurechenskii, A. I. Nikiforov, and A. A. Bloshkin, JETP Lett. **77**, 376 (2003).

[34] N. P. Stepina, A. I. Yakimov, A. V. Nenashev, A. V. Dvurechenski, N. A. Sobolev, D. P. Leitao, V. V. Kirienko, A. I. Nikiforov, E. S. Koptev, L. Pereira, and M. S. Karmo, J. Exp. Theor. Phys. **103**, 269 (2006).

[35] A. V. Nenashev, J. O. Oelerich, and S. D. Baranovskii, J. Phys.: Condens. Matter **27**, 093201 (2015).

[36] D. N. Tsigankov and A. L. Efros, Phys. Rev. Lett. **88**, 176602 (2002).

[37] D. S. Smirnov and L. E. Golub, Phys. Rev. Lett. **118**, 116801 (2017).