Why Is the Bulk Resistivity of Topological Insulators So Small?

Brian Skinner, Tianran Chen, and B. I. Shklovskii
Fine Theoretical Physics Institute, University of Minnesota, Minneapolis, Minnesota 55455, USA
(Received 27 August 2012; published 23 October 2012)

As-grown topological insulators (TIs) are typically heavily doped n-type crystals. Compensation by acceptors is used to move the Fermi level to the middle of the band gap, but even then TIs have a frustratingly small bulk resistivity. We show that this small resistivity is the result of band bending by poorly screened fluctuations in the random Coulomb potential. Using numerical simulations of a completely compensated TI, we find that the bulk resistivity has an activation energy of just 0.15 times the band gap, in good agreement with experimental data. At lower temperatures activated transport crosses over to variable range hopping with a relatively large localization length.

DOI: 10.1103/PhysRevLett.109.176801 PACS numbers: 73.63.−b, 61.72.uf, 72.20.−i

The three-dimensional (3D) topological insulator (TI) [1–5] has gapless surface states that are expected to exhibit a range of interesting quantum phenomena [6,7]. While a number of 3D TIs have been identified, most of these are poor insulators in the bulk, so that the properties of the surface are obscured in transport measurements. For this reason achieving a bulk-insulating state remains an active topic of research [8–17].

Typically, as-grown TI crystals are heavily doped n-type semiconductors, and, correspondingly, they exhibit metallic conduction. In order to make them insulating these TIs are compensated by acceptors. With increasing compensation $K = N_A/N_D$, where N_D and N_A are the concentrations of monovalent donors and acceptors, respectively, the Fermi level shifts from the conduction band to inside the band gap and then at $K > 1$ into the valence band. When compensation of donors is complete, $K = 1$, the Fermi level is in the middle of the gap and the most insulating state of the TI is reached. The hope is that at $K = 1$ the bulk resistivity ρ should obey the activation law,

$$\rho = \rho_0 \exp(\Delta/k_B T), \tag{1}$$

with an activation energy Δ that is equal to half the band gap E_g. Here, ρ_0 is a constant and $k_B T$ is the thermal energy. Since typically $E_g \sim 0.3$ eV, this expectation would imply a well-insulating bulk at room temperatures and below.

The typical experimental situation at $K = 1$, however, is frustrating [16]. In the range of temperatures between 100 and 300 K the resistivity is activated, but with an activation energy that is three times smaller than expected, $\Delta \sim 50$ meV. At $T \lesssim 100$ K the activated transport is replaced by variable range hopping (VRH), characterized by $\rho \propto \exp[(T_0/T)^x]$ with $x < 1$, and the resistivity grows even more slowly with decreasing T. In Ref. [16] the authors show that Mott VRH ($x = 1/4$) provides a reasonable fit to their data at 50 K $\lesssim T \lesssim$ 100 K. Definite characterization of the temperature exponent x is difficult, however, due to the relatively narrow window of temperature and to variations between samples. At $T \lesssim 50$ K the resistivity saturates due to the contribution of the surface states.

In this Letter we suggest an explanation for the unexpectedly small bulk resistivity of TIs. We assume that both donors and acceptors are shallow and we use the theory of completely compensated semiconductors (CCS) [18,19]. This theory is based on the idea that near $K = 1$, when almost all donors and acceptors are charged, random fluctuations in the local concentration of impurities result in large fluctuations of charge. The resulting Coulomb potential is poorly screened because of the vanishing average concentration $n = N_D - N_A$ of screening electrons. Huge fluctuations in the random potential bend the conduction and valence bands edges and in some places bring them to the Fermi level, thereby creating electron and hole puddles that nonlinearly screen the random potential. Thus, the amplitude of fluctuations is limited only by the semiconductor gap E_g. As a result the ground state of a CCS, shown in Fig. 1, is similar to a network of p-n junctions [18,19]. The characteristic size of these p-n junctions, also called the nonlinear screening radius, is given by

$$R_g = \frac{E_g^2 \kappa^2}{8\pi N e^4}, \tag{2}$$

where κ is the dielectric constant, e is the electron charge, and $N = N_D = N_A$. For $N = 10^{19}$ cm^{-3} and $\kappa = 20$, $R_g \approx 70$ nm $\gg N^{-1/3} \approx 4.6$ nm, so that we deal with a very long range potential. As a result, the resistivity can be dramatically different from the expectation outlined above, which assumed flat bands. First, at relatively high temperatures conduction is due to electrons and holes being activated from the Fermi level to their corresponding classical percolation levels (classical mobility edges), E_e and E_h, in the conduction and the valence bands. These may be substantially closer to the Fermi level μ than $E_g/2$, but so far the resulting value of Δ has not been studied theoretically. Second, at sufficiently low temperatures electrons and holes can hop (tunnel) between distant puddles, so that variable range hopping replaces activated

 © 2012 American Physical Society

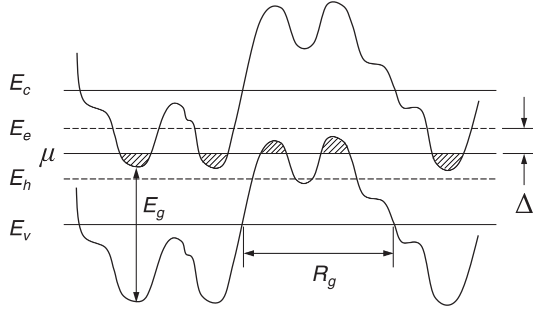

FIG. 1. Energy diagram of a completely compensated TI with band gap E_g. The upper and the lower straight lines (E_c and E_v) indicate the unperturbed positions of the bottom of the conduction band and the ceiling of the valence band; the middle line (μ) corresponds to the Fermi level. Meandering lines represent the band edges, which are modulated by the fluctuating potential of charged impurities; R_g is the characteristic size of these potential fluctuations. The percolation levels for electrons, E_e, and holes, E_h, are shown by dashed lines; the activation energy Δ corresponds to the difference $E_e - \mu$ (or $\mu - E_h$). Puddles occupied by carriers are shaded. Shallow impurity levels are not shown because they merge with the band edges.

transport. In the low temperature limit $\rho(T)$ should obey the Efros-Shklovskii (ES) law of VRH [20],

$$\rho = \rho_0 \exp[(T_{\mathrm{ES}}/T)^{1/2}], \tag{3}$$

where $T_{\mathrm{ES}} = Ce^2/\kappa\xi$, ξ is the localization length of states with energy close to the Fermi level, and C is a numerical coefficient. So far the magnitude of T_{ES} and the nature of the crossover between activated and VRH conduction have not been studied.

In this Letter, motivated by the TI resistivity puzzle, we return to CCS and model numerically the $K = 1$ case. For moderately large T we find that $\Delta = 0.15E_g$. For a TI with $E_g = 0.3$ eV this implies $\Delta = 45$ meV, in agreement with observed values [16]. We also calculate the single-particle density of states (DOS) of impurity states, and we find that the DOS has a Coulomb gap at the Fermi level [20]. We show from our simulation that the resistivity is described by Eq. (3) at low temperatures and crosses over to Eq. (1) at higher T. We present a crude estimate of the localization length ξ which suggests that $T_{\mathrm{ES}} \sim 900$ K and that the crossover between activation and ES VRH occurs at $T \sim 40$ K. Together our results for the activated and VRH resistivity establish a universal upper limit for the resistivity $\rho(T)$ that one can achieve for a 3D TI compensated by shallow inpurities.

In order to model the CCS numerically, we simulate a cube filled by an equal number of randomly positioned donors and acceptors (20 000 of each). We numerate all donors and acceptors by the index i and we define $n_i = 0, 1$ as the number of electrons residing at impurity i and the variable f_i to discriminate between donors ($f_i = 1$) and acceptors ($f_i = -1$). The resulting Hamiltonian is

$$H = \frac{E_g}{2}\sum_i f_i n_i + \sum_{\langle ij \rangle} V(r_{ij}) q_i q_j, \tag{4}$$

where $q_i = (f_i + 1)/2 - n_i$ is the net charge of site i and all energies are defined relative to the middle of the band gap. The first term of Eq. (4) contains the energies of donor and acceptor sites, which for the case of shallow impurities is very close to $\pm E_g/2$. The second term is the total interaction energy of charged impurities. For two impurities at a distance $r \gg a_B$, where a_B is the Bohr radius of impurity states, one can describe their interaction using the normal Coulomb law $V(r) = e^2/\kappa r$. For example, an empty donor shifts the energy of an electron at a distant filled donor by an amount $-e^2/\kappa r$. On the other hand, for a pair of impurities with separation $r < a_B$, quantum mechanical averaging over the electron wave function becomes important (such close impurity pairs are common in heavily doped semiconductors, where $a_B > N^{-1/3}$). A pair of very close donors, for example, cannot create an electron state deeper than that of the heliumlike ion with binding energy $2e^2/\kappa a_B$. In order to capture this quantum phenomenon in an approximate way, we use the classical Hamiltonian of Eq. (4) with a truncated Coulomb potential $V(r) = e^2/\kappa(r^2 + a_B^2)^{1/2}$. The result of this truncation is to eliminate the unphysically deep electron states that would result from very compact impurity pairs with an unmodified $1/r$ interaction. We will show below that our results are mostly insensitive to the details of this truncation. Note that Eq. (4) does not include the kinetic energy of electrons and holes in the conduction and valence bands and, therefore, aims only at describing the low temperature ($k_B T \ll E_g$) physics of CCS.

In all results below we use dimensionless units for r, a_B, ξ, H, E_g, and $k_B T$, measuring all distances in units of $N^{-1/3}$ and all energies in units of $e^2 N^{1/3}/\kappa$. Thus, Eq. (4) can be understood as dimensionless, with $E_g \gg 1$ and $V(r) = (r^2 + a_B^2)^{-1/2}$. For a TI with $E_g = 0.3$ eV, $\kappa = 20$ and $N = 10^{19}$ cm^{-3}, the unit of energy $e^2 N^{1/3}/\kappa \approx 15$ meV, so that the dimensionless gap $E_g \approx 20$. We were unable to directly model $E_g = 20$, since in this case the very large $R_g \approx 16$ leads to large size effects. Instead, we present results for the more modest $E_g = 10$, where $R_g \approx 4$ and size effects are negligible, and for $E_g = 15$, where $R_g \approx 9$ and size effects can be treated using extrapolation. Unless otherwise stated, results below use $a_B = 2$ and are averaged over 100 random initializations of the donor and acceptor positions.

In our simulation, we first search for the set of electron occupation numbers $\{n_i\}$ that minimizes H. We start by assuming that all donors are empty ($n_i = 0$, $q_i = 1$) and that all acceptors are filled ($n_i = 1$, $q_i = -1$). These charged donors and acceptors create a random Coulomb potential whose magnitude exceeds E_g. We then sequentially choose pairs consisting of one filled site and one

empty site and attempt to transfer an electron from the filled site to the empty site. If the proposed move lowers the total system energy H, it is accepted, otherwise it is rejected. To describe the change in H resulting from such a transfer it is convenient to introduce the single-electron energy state, ε_i, at a given impurity i:

$$\varepsilon_i = \frac{E_g}{2} f_i - \sum_{j \neq i} V(r_{ij}) q_j. \quad (5)$$

The process of transferring electrons concludes when all pairs i, j with $n_i = 1$ and $n_j = 0$ satisfy the ES stability criterion:

$$\varepsilon_j - \varepsilon_i - V(r_{ij}) > 0. \quad (6)$$

This final arrangement of electrons can be called a pseudo-ground state, since higher stability criteria of the ground state (involving multiple simultaneous electron transfers) are not checked. Such pseudo-ground states are known to accurately describe the properties of the real ground state at all but extremely small energies [19,21,22].

Once the pseudo-ground state $\{n_i\}$ is known, the DOS $g(\varepsilon)$ is calculated by making a histogram of the single-electron energies $\{\varepsilon_i\}$. The result is shown in Fig. 2, with the DOS in units of $2N/(e^2N^{1/3}/\kappa)$, so that the total area is equal to unity. Occupied and empty states are separated by the Fermi level at $\varepsilon = 0$. The nearly constant DOS between $-E_g$ to E_g reflects a practically uniform distribution of the random potential from $-E_g/2$ to $+E_g/2$. Near the Fermi level one sees the Coulomb gap that is a universal result of the ES stability criterion [20].

Once the energies $\{\varepsilon_i\}$ are calculated, we evaluate the resistivity using the approach of the Miller-Abrahams resistor network [19]. Namely, each pair of impurities i, j is said to be connected by the resistance $R_{ij} = R_0 \exp[2r_{ij}/\xi + \varepsilon_{ij}/k_BT]$, where the activation energy ε_{ij} is defined [19] as follows:

$$\varepsilon_{ij} = \begin{cases} |\varepsilon_j - \varepsilon_i| - V(r_{ij}), & \varepsilon_j\varepsilon_i < 0 \\ \max[|\varepsilon_i|, |\varepsilon_j|], & \varepsilon_j\varepsilon_i > 0. \end{cases} \quad (7)$$

The resistivity of the system as a whole is found using a percolation approach. Specifically, we find the minimum value R_c such that if all resistances R_{ij} with $R_{ij} < R_c$ are left intact, while others are eliminated (replaced with $R = \infty$), then there exists a percolation pathway connecting opposite faces of the simulation volume. The system resistivity $\rho(T)$ is taken to be proportional to R_c, which captures the exponential term while details of the prefactor are ignored [19].

In Fig. 3 we plot the computed resistivity as a function of temperature, using the dimensionless logarithm of the resistance $(\ln\rho)^* = (\xi/2)\ln(R_c/R_0)$ and the dimensionless temperature $T^* = 2k_BT/\xi$. These notations are introduced to exclude any explicit dependence on ξ. Figure 3(a) shows $(\ln\rho)^*$ versus $(T^*)^{-1/2}$ over the huge range of temperatures $0.03 < T^* < 200$. One can see that at low temperatures $T^* < 0.3$ the resistivity is well described by the ES law, Eq. (3), with $C \approx 4.4$. The higher temperature range $1 < T^* < 200$ is plotted separately as a function of $1/T^*$ in Fig. 3(b). Here we find two activated regimes of hopping

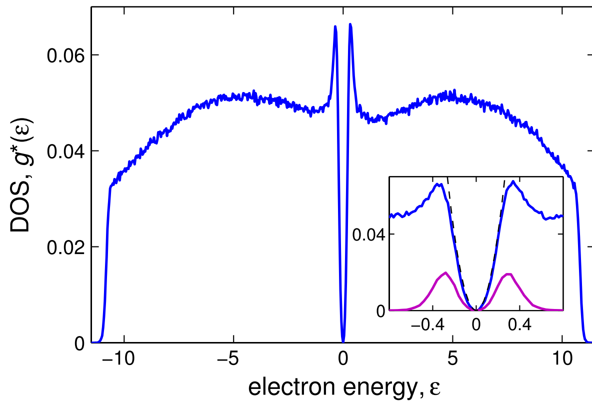

FIG. 2 (color online). Dimensionless single-electron DOS $g^*(\varepsilon) = g(\varepsilon)/[2N/(e^2N^{1/3}/\kappa)]$ for a completely-compensated semiconductor with $a_B = 2$ and $E_g = 10$. The inset shows the DOS near the Fermi level $\varepsilon = 0$ (upper curve, blue). For comparison, the quadratic Coulomb gap $g(\varepsilon) = (3/\pi)\varepsilon^2$ is shown by the dashed line [20,23]. The lower (magenta) line shows separately the DOS of rare filled donors and empty acceptors.

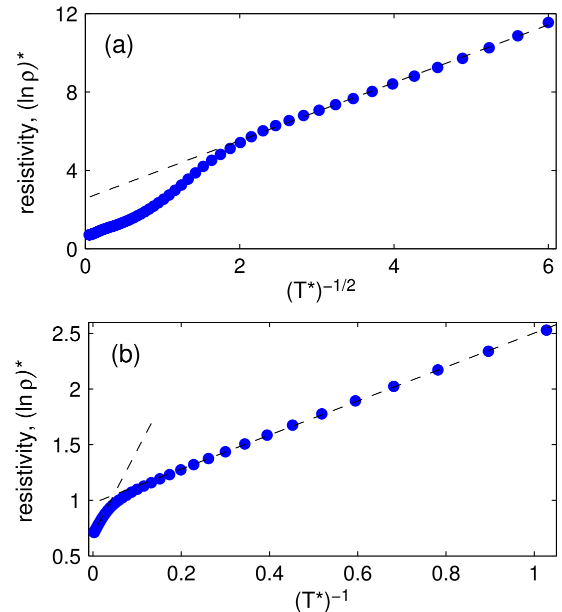

FIG. 3 (color online). The temperature dependence of the resistivity for $E_g = 10$ (blue dots). The dimensionless resistivity $(\ln\rho)^*$ is plotted in (a) against $(T^*)^{-1/2}$ to illustrate that the resistivity follows the ES law at low temperatures, and in (b) against $(T^*)^{-1}$ to show that the resistivity is activated at larger T^*, with two distinct activation energies. The dashed lines (black) are linear best fits.

conductivity. At extremely high temperatures $T^* > 50$ we see the large activation energy $E_a \sim 0.75E_g$ while in the intermediate range $1 < T^* < 10$ we see an activation energy $\Delta = (0.15 \pm 0.01)E_g$. We repeated this analysis for the larger band gap $E_g = 15$ using systems of 10 000, 20 000, and 30 000 donors and by extrapolating to infinite size we find $\Delta = (0.15 \pm 0.02)E_g$. These results for Δ remain unchanged, within our statistical uncertainty, if we use $a_B = 1$ instead of $a_B = 2$.

It should be noted that the large activation energy $E_a \sim 0.75E_g$ observed at $T^* > 50$ does not have any physical meaning for a real CCS, since at such large temperatures the conduction is not due to hopping but rather to free, *hot* carriers far from the conduction and valence band edges. Nonetheless, for our model Hamiltonian this result is consistent with established theories which say that at such large temperatures $E_a = \langle \varepsilon_{ij} \rangle$, where $\langle \ldots \rangle$ denotes averaging over all pairs i, j (see Chap. 8 of Ref. [19]).

On the other hand, the second activation energy $\Delta = 0.15E_g$ makes full physics sense and should be seen in experiment. At $T \ll E_g$ electrons optimize their conductivity by hopping among impurities that are energetically close to the Fermi level. The activation energy Δ can be understood as the resulting percolation level for hopping between nearest-neighboring sites. In other words, if electrons are activated only to those sites with $|\varepsilon| < \varepsilon_p$, then precisely at $\varepsilon_p \geq \Delta = 0.15E_g$ there exists an infinite conduction pathway for electrons comprised of hops of length $\sim N^{-1/3}$ or shorter. In a heavily doped semiconductor this energy is equivalent to the activation energy of electrons from the Fermi level to the conduction band mobility edge E_e. (Of course, holes are activated from the Fermi level to their percolation level E_h as well). For a typical TI $E_g = 0.3$ eV, so that we get $\Delta = 45$ meV, in good agreement with typical experimental data [16]. (We note, however, that recent experiments on Sn-doped Bi_2Te_2Se have achieved $\Delta \sim 125$ meV [17]. Such large activation energies may be associated with deep donor impurity levels, which go beyond our model).

This activation to the percolation level persists until much smaller temperatures, where Δ becomes prohibitively large compared to the thermal energy. At such small T^* conduction proceeds by VRH among electron or hole puddles at the Fermi level and the resistivity is given by Eq. (3).

One can interpret the relatively small numerical factor 0.15 above by recalling that in a typical 3D continuous random potential, ~17% of space has a potential smaller than the percolation level [19]. As we demonstrated above the energy of the conduction band bottom is roughly uniformly distributed in the interval $(0, E_g)$. This means that the percolation level E_e should be close to $0.17E_g$ and makes our result $\Delta = 0.15E_g$ quite reasonable.

So far we have emphasized results that do not explicitly depend on the localization length ξ. In fact, knowledge of ξ is necessary to predict T_{ES} and the transition temperature T_t between Eq. (1) and (3) in real temperature units. [According to Fig. 3(a), the transition happens at $T^* \approx 1/2$, or $T_t \approx \xi/4$.] We argue now that in a TI ξ is quite large, leading to a prominent role for VRH. To see this, consider that if an electron with energy close to the Fermi level is assumed to tunnel from one electron puddle to another distant puddle along the straight line connecting them, then the tunneling path passes through regions where the conduction band bottom is quite high above Fermi level. This implies a small tunneling amplitude, or $\xi \ll a_B$. In fact, however, a tunneling electron can use the same geometrical path as a classical percolating electron with energy Δ above the Fermi level. In order to roughly estimate ξ, we assume that along such a classical percolation path the tunneling barriers V are uniformly distributed in the range $0 \leq V \leq \Delta$ and we neglect the curvature of this path. Integrating the action along this path then gives $\xi \sim \hbar/(m\Delta)^{1/2} = a_B\sqrt{e^2/a_B\Delta}$. For a TI with $E_g = 20$ and $a_B = 2$ this gives $\xi \simeq 0.8$. This crude estimate leads to $T_{ES} \sim 900$ K and $T_t \sim 40$ K, which is similar in magnitude to the experimentally observed $T_t \sim 100$ K where the resistivity crosses over from activated to VRH behavior [16].

We note that if one plots our result for $(\ln\rho)^*$ against $(T^*)^{-1/4}$ in the relatively narrow crossover range 50 K $< T <$ 100 K, one gets a mostly straight line, as seen in Ref. [16]. However, our results suggest that at low temperatures the bulk resistivity follows the ES law of VRH with temperature exponent $x = 1/2$, which should become apparent if the bulk resistivity can be probed to very low temperature. Such measurements are presumably possible in samples that are much thicker than those studied in Ref. [16] (~ 100 μm). For such thick samples conduction through the bulk of the TI crystal dominates over the surface transport until much smaller temperatures.

To conclude, we have studied numerically the bulk resistivity of a TI crystal with band gap E_g as a CCS. We find that at high temperatures $k_BT \gtrsim 0.03E_g$ the resistivity is activated with relatively small activation energy $0.15E_g$, in agreement with experimental data [16]. At lower temperatures the resistivity crosses over to ES VRH, with an estimated characteristic temperature $T_{ES} \sim 900$ K. Thus, Eq. (3) with $T_{ES} \sim 900$ K at $T \lesssim 40$ K and Eq. (1) with $\Delta = 0.15E_g$ at $T \gtrsim 40$ K give the upper limit for resistivity that one can achieve for a heavily doped and completely compensated TI with shallow impurities.

We are grateful to Y. Ando, A. L. Efros, M. S. Fuhrer, Yu. M. Galperin, M. Müller, and N. P. Ong for useful discussions. This work was supported primarily by the MRSEC Program of the National Science Foundation under Grant No. DMR-0819885. T. Chen was partially supported by the FTPI.

[1] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).
[2] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).
[3] R. Roy, Phys. Rev. B **79**, 195322 (2009).
[4] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[5] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Phys. Rev. B **78**, 195424 (2008).
[6] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[7] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
[8] D.-X. Qu, Y. S. Hor, J. Xiong, R. J. Cava, and N. P. Ong, Science **329**, 821 (2010).
[9] J. G. Analytis, R. D. McDonald, S. C. Riggs, J.-H. Chu, G. S. Boebinger, and I. R. Fisher, Nat. Phys. **6**, 960 (2010).
[10] J. Xiong, Y. Luo, Y. H. Khoo, S. Jia, R. J. Cava, and N. P. Ong, Phys. Rev. B **86**, 045314 (2012).
[11] J. G. Checkelsky, Y. S. Hor, M.-H. Liu, D.-X. Qu, R. J. Cava, and N. P. Ong, Phys. Rev. Lett. **103**, 246601 (2009).
[12] N. P. Butch, K. Kirshenbaum, P. Syers, A. B. Sushkov, G. S. Jenkins, H. D. Drew, and J. Paglione, Phys. Rev. B **81**, 241301 (2010).
[13] J. G. Analytis, J.-H. Chu, Y. Chen, F. Corredor, R. D. McDonald, Z. X. Shen, and I. R. Fisher, Phys. Rev. B **81**, 205407 (2010).
[14] K. Eto, Z. Ren, A. A. Taskin, K. Segawa, and Y. Ando, Phys. Rev. B **81**, 195309 (2010).
[15] Z. Ren, A. A. Taskin, S. Sasaki, K. Segawa, and Y. Ando, Phys. Rev. B **84**, 075316 (2011).
[16] Z. Ren, A. A. Taskin, S. Sasaki, K. Segawa, and Y. Ando, Phys. Rev. B **84**, 165311 (2011).
[17] Z. Ren, A. A. Taskin, S. Sasaki, K. Segawa, and Y. Ando, Phys. Rev. B **85**, 155301 (2012).
[18] B. I. Shklovskii and A. L. Efros, Sov. Phys. JETP **35**, 610 (1972).
[19] A. L. Efros and B. I. Shklovskii, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, New York, 1984), http://www.tpi.umn.edu/shklovskii.
[20] A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).
[21] A. Möbius, M. Richter, and B. Drittler, Phys. Rev. B **45**, 11568 (1992).
[22] A. L. Efros, B. Skinner, and B. I. Shklovskii, Phys. Rev. B **84**, 064204 (2011).
[23] A. L. Efros, J. Phys. C **9**, 2021 (1976).