

Conductivity of Two-Dimensional Small Gap Semiconductors and Topological Insulators in Strong Coulomb Disorder

Yi Huang^{a,*}, Brian Skinner^b, and B. I. Shklovskii^a

^a School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota, 55455 USA

^b Department of Physics, Ohio State University, Columbus, Ohio, 43202 USA

*e-mail: huan1756@umn.edu

Received May 29, 2022; revised May 29, 2022; accepted June 1, 2022

Abstract—We are honored to dedicate this article to Emmanuel Rashba on the occasion of his 95 birthday. In the ideal disorder-free situation, a two-dimensional bandgap insulator has an activation energy for conductivity equal to half the bandgap, Δ . But transport experiments usually exhibit a much smaller activation energy at low temperature, and the relation between this activation energy and Δ is unclear. Here we consider the temperature-dependent conductivity of a two-dimensional narrow gap semiconductor on a substrate containing Coulomb impurities, mostly focusing on the case when amplitude of the random potential $\Gamma \gg \Delta$. We show that the conductivity generically exhibits three regimes and only the highest temperature regime exhibits an activation energy that reflects the band gap. At lower temperatures, the conduction proceeds through nearest-neighbor or variable-range hopping between electron and hole puddles created by the disorder. We show that the activation energy and characteristic temperature associated with these processes steeply collapse near a critical impurity concentration. Larger concentrations lead to an exponentially small activation energy and exponentially long localization length, which in mesoscopic samples can appear as a disorder-induced insulator-to-metal transition. We arrive at a similar disorder driven steep insulator–metal transition in thin films of three-dimensional topological insulators with very large dielectric constant, where due to confinement of electric field internal Coulomb impurities create larger disorder potential. Away from neutrality point this unconventional insulator-to-metal transition is augmented by conventional metal–insulator transition at small impurity concentrations, so that we arrive at disorder-driven re-entrant metal–insulator–metal transition. We also apply this theory to three-dimensional narrow gap Dirac materials.

DOI: 10.1134/S1063776122100065

1. INTRODUCTION

In a band gap insulator, charged impurities often play a decisive role in determining the properties of the insulating state. Due to the long-ranged nature of the Coulomb potential that they create, such impurities produce large band bending that changes qualitatively the nature of electron conduction relative to the ideal disorder-free situation. An illustrative case is that of a three-dimensional completely-compensated semiconductor, for which positively-charged donors and negatively-charged acceptors are equally abundant and randomly distributed in space. In this case, the impurity potential has large random fluctuations, which can be screened only when the amplitude of this potential reaches Δ , where 2Δ is the band gap. This screening is produced by sparse electron and hole droplets, concentrated in spatially alternating electron and hole clouds (puddles) [1–3] (see Fig. 1). At high enough temperatures the electrical conductivity is due to activation of electrons and holes from the Fermi level to the energy associated with classical percolation across the sample. At lower temperatures the conduc-

tivity is due to hopping between nearest neighbor puddles (NNH). At even smaller temperatures it is due to variable range hopping (VRH) between puddles. Crucially, in each of these temperature regimes the naive relation $E_a = \Delta$ is lost, where E_a is the activation energy for conductivity. Only in the highest temperature regime is there a direct proportionality between E_a and Δ (with a nontrivial small numeric prefactor) [3, 4]; at lower temperatures the observed activation energy is nonuniversal and disorder-dependent [1, 2].

In this paper we consider a similar problem in two dimensions. Specifically, we consider a two-dimensional small band gap semiconductor resting on a thick substrate with a three dimensional concentration of randomly-positioned impurities and focus on the case when $\Gamma \gg \Delta$ (see Fig. 2). We derive the temperature dependence of the electrical conductivity across all temperature regimes and show that the observed activation energy of the conductivity can be very small.

Understanding the relation between the energy gap and the observed activation energy for transport is of crucial importance for studying a variety of new 2D

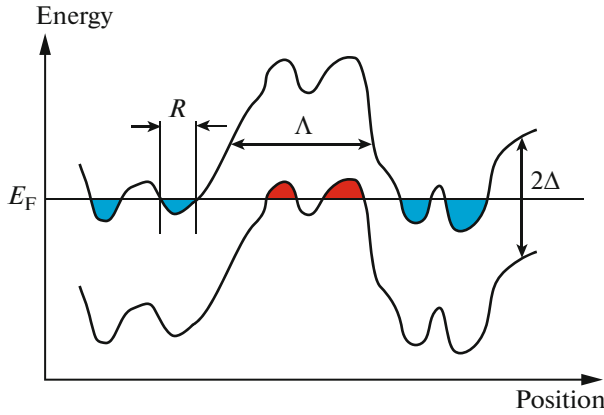


Fig. 1. Schematic energy diagram of a completely compensated semiconductor with relatively weak disorder. The wavy lines show the conduction band bottom and the valence band ceiling separated by the gap 2Δ . Droplets of holes are shaded by red, while electron droplets are shaded by blue. Here R is the size of a droplet, and Λ is the size of a droplet cloud (puddle), which contains several droplets.

electron systems. For example, recent studies of 2D topological insulators (TIs) [5–7], films of 3D TIs [8–24], bilayer graphene (BLG) with an orthogonal electric field [25, 26] and twisted bilayer graphene (TBG) [27–31] use the transport activation energy as a way of characterizing small energy gaps. In all these cases the observed activation energy is much smaller than the energy gap that is expected theoretically or measured through local probes like optical absorption or scanning tunneling microscopy.

Here, we show that there is indeed no simple proportionality between the energy gap and the activation energy except at the highest temperature regime, which is likely irrelevant for many experimental contexts. Instead, we find a wide regime of temperature and disorder strength for which the activation energy is dramatically smaller than the energy gap. At the lowest temperatures the conductivity follows the Efros-Shklovskii (ES) law rather than an Arrhenius law, and this dependence can give the appearance of a small activation energy.

Let us dwell on two likely applications of our theory. First, our results may be especially relevant for ongoing efforts to understand the energy gaps arising in TBG at certain commensurate fillings of the moiré superlattice [27–31]. Such gaps apparently arise from electron–electron interactions, but the observed activation energies of the maximally-insulating state are typically an order of magnitude smaller than the naive interaction scale (see, e.g., [28, 29]), and they vary significantly from one sample to another. Scanning tunneling microscopy studies also suggest a gap on the order of ten times larger than the observed activation energy [32, 33]. The theory we present here offers a natural way to interpret this discrepancy.

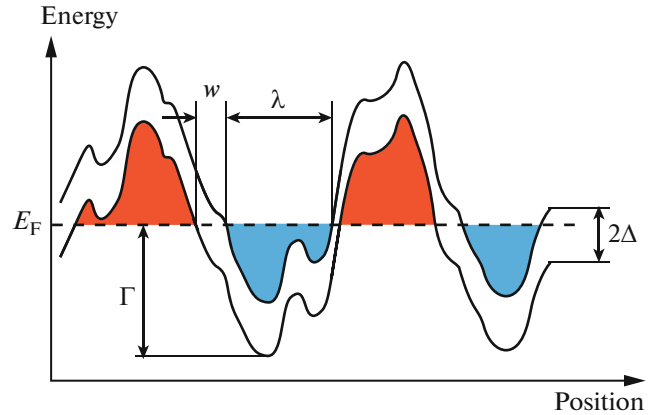


Fig. 2. Schematic picture of a cross section of puddles for the case of strong disorder, $\Gamma \gg \Delta$. The wavy lines show the conduction band bottom and the valence band ceiling separated by the gap 2Δ . The red shaded region above the Fermi level $E_F = 0$ represents a hole puddle, while the blue shaded region below E_F represents an electron puddle. λ is the amplitude of the disorder potential, Γ is the screening length, and w is the width of the barrier between neighboring puddles.

Second, our theory can be applied to the huge body of experimental work on thin films of 3D TI, where the surface electrons have a small gap 2Δ due to hybridization of the surface states of two surfaces [8, 9], or due to intentionally introduced magnetic impurities [10–24]. Understanding the origin of the small apparent activation energy $E_a \ll \Delta$ is crucial for achieving metrological precision of the quantum anomalous Hall effect [11, 13, 16, 19–24, 34–36] and the quantum spin Hall effect [9, 37–39].

The model we consider is a two-dimensional semiconductor with band gap 2Δ atop a substrate with a three-dimensional concentration N of random sign charged impurities. We assume that the semiconductor has a gapped Dirac dispersion law

$$\epsilon^2(\mathbf{k}) = (\hbar v k)^2 + \Delta^2. \quad (1)$$

We are mostly interested in the case when the amplitude Γ of spatial fluctuations of the random potential satisfies $\Gamma \gg \Delta$, so that electron and hole puddles occupy almost half of the space each and are separated by a small insulating gap which occupies only a small fraction of the space (see Fig. 2). This system is an insulator because in 2D neither electron nor hole puddles percolate, and they are disconnected from each other. Throughout this paper we mostly focus on the case of zero chemical potential, for which electron and hole puddles are equally abundant and the system achieves its maximally insulating state. We argue that this situation is likely realized in the experiments of [5–33].

The remainder of this paper is organized as follows. In the following section we first summarize our main results for the temperature-dependent conductivity. Sections 3 and 4 concentrate on the case $\Gamma \gg \Delta$ illus-

trated by Fig. 2. In Section 3 we start from reviewing the fractal geometry of two-dimensional puddles and then calculate the action accumulated by electrons tunneling across the gap between two neighboring fractal metallic puddles, the corresponding localization length, and the critical value of the ratio Γ/Δ , at which crossover to weak localization takes place. In Section 4 we calculate the hopping conductivity for the case $\Gamma \gg \Delta$.

In Section 5 we study the illustrated by Fig. 1 case where the impurity concentration N is lower and present the parameters of NNH and VRH as functions of N . Section 6 studies what happens when the Fermi level moves away from the neutrality point. We arrive at the “phase diagram” of the re-entrant metal–insulator–metal transition. Section 7 deals with the generalization of our results to thin TI films. Because of large interest to such films [8–10, 12–24, 34–49], in this section we add a fair amount of numerical estimates. In Section 8 we briefly return to the problem of three-dimensional, completely-compensated semiconductors with a gapped Dirac dispersion, and extend the previous theory [1–3] to the case when disorder potential fluctuations exceed Δ . We again arrive at a re-entrant metal–insulator–metal transition away from the neutrality point. We close in Section 9 with a summary and conclusion. Some results of this paper are published in its shorter version [50].

2. SUMMARY OF RESULTS

Let us start from the strong disorder case $\Gamma \gg \Delta$ illustrated in Fig. 2. When the typical tunneling transparency $P = \exp(-S)$ of the insulating barrier separating neighboring puddles is small (the action S in units of $\hbar$ is large), one can envision a sequence of three mechanisms of activated transport replacing each other with decreasing temperature, as in a lightly doped wide gap semiconductor [2]. This three-mechanism sequence is illustrated in Fig. 3. At relatively large temperature T electrons and holes can be activated from the Fermi level to the percolation level (i.e., the classical mobility edge). Thus, the conductivity at such large temperatures is given [50] by

$$\sigma = \sigma_1 \exp(-\Delta/T), \quad (T_1 \ll T \ll \Delta) \quad (2)$$

with the prefactor $\sigma_1 \sim e^2/\hbar$. Here and everywhere in this paper we use energy units for the temperature T (absorbing k_B in its definition).

At lower temperatures this mechanism yields to the nearest-neighbor hopping (NNH) of electrons between electron and hole puddles near the Fermi level. Similarly to the case of granular metals [51, 52], the activation energy of such hopping is determined by the typical puddle charging energy E_C

$$\sigma = \sigma_2 \exp(-E_C/T), \quad (T_2 \ll T \ll T_1). \quad (3)$$

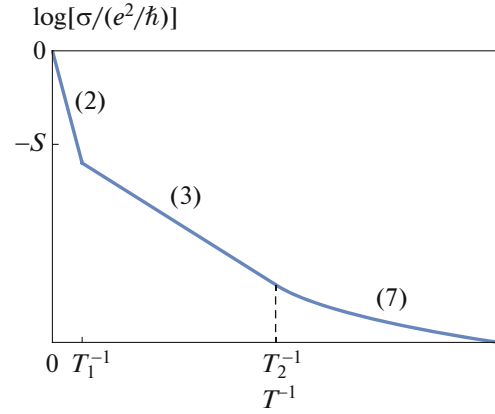


Fig. 3. Logarithm of the dimensionless conductivity $\sigma/(e^2/\hbar)$ as a function of the inverse temperature T^{-1} in the case $1 \ll \Gamma/\Delta \ll (\Gamma/\Delta)_c$. At high temperature $T > T_1$, the conductivity has activation energy Δ . At intermediate temperature $T_2 < T < T_1$, the conductivity is dominated by NNH. At low temperatures $T < T_2$, NNH is replaced by ESVRH. Numbers adjacent to different parts of the line show corresponding equations. Temperatures T_1 and T_2 are given by Eqs. (9) and (10).

Here the prefactor $\sigma_2 \sim (e^2/\hbar) \exp(-S) \ll (e^2/\hbar)$. We show below that

$$E_C = \alpha^2 \Delta (\Delta/\Gamma)^{4/3} = \alpha^2 \Delta (N_0/N)^{4/9} \ll \Delta. \quad (4)$$

Here $\alpha = e^2/(\kappa \hbar v)$ is the analog of the fine structure constant and κ is the dielectric constant of the substrate. With the standard semiconductor value $v \sim 10^6$ m s⁻¹; and with $\kappa = 4$ for SiO₂, 11 for insulating GaAs, 20 for HfO₂ and 1000 for PbTe; α can vary from 1 to 10⁻³. Below in our theory we use α as a small parameter, $\alpha \ll 1$, but our results are semiquantitatively correct at $\alpha = 1$. In Eq. (4) we also used derived below equation

$$\Gamma/\Delta = (N/N_0)^{1/3}, \quad (5)$$

where characteristic concentration

$$N_0(\Delta) = \alpha^2 \kappa^3 \Delta^3 e^{-6}. \quad (6)$$

In Eqs. (4)–(6) and everywhere below we use a scaling approach and omit all numerical coefficients.

At even lower temperatures NNH crosses over to VRH obeying the Efros-Shklovskii (ES) law

$$\sigma = \sigma_3 \exp[-(T_{ES}/T)^{1/2}], \quad (T \ll T_2), \quad (7)$$

with $\sigma_3 \sim e^2/\hbar$. We show below that in this regime

$$T_{ES} = \alpha \Delta (\Delta/\Gamma)^{37/9} = \alpha \Delta (N_0/N)^{37/27} \ll \Delta, \quad (8)$$

and the temperatures associated with the crossover between different regimes are

$$T_1 = \alpha \Delta (\Gamma/\Delta)^{34/9} = \alpha \Delta (N/N_0)^{34/27}, \quad (9)$$

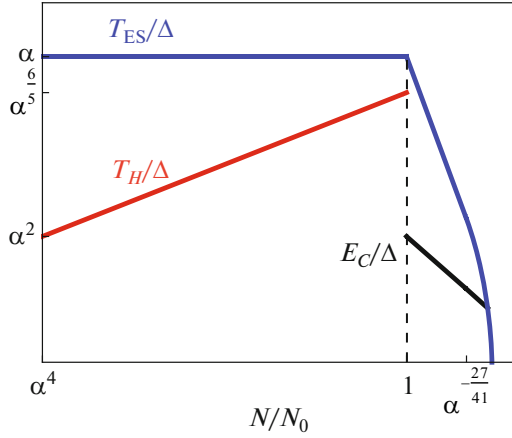


Fig. 4. Schematic log-log plots of characteristic energies of three kinds of hopping conductivity. The characteristic temperature of ES law T_{ES} (blue line), the activation energy of NNH, E_C (black solid line) and the characteristic temperature of hybrid conductivity T_H (red line) are shown as functions of the dimensionless impurity concentration $N/N_0 = (\Gamma/\Delta)^3$. The left part of the plot where $N/N_0 < 1$ corresponds to Eqs. (41) and (12), while the right part at $N/N_0 > 1$ corresponds to Eqs. (8) and (4). In the horizontal axis $N/N_0 = N_c/N_0 = \alpha^{-27/41}$ corresponds to $\Gamma/\Delta = (\Gamma/\Delta)_c$ given by Eq. (28). At this point $T_{\text{ES}} = \alpha^{87/41}\Delta$. When $\Gamma/\Delta > (\Gamma/\Delta)_c$ the localization length ξ increases exponentially and T_{ES} decreases exponentially.

$$T_2 = \alpha^3 \Delta (\Gamma/\Delta)^{13/9} = \alpha^3 \Delta (N/N_0)^{13/27}. \quad (10)$$

Above we dealt with large impurity concentration $N > N_0$, which corresponds to $\Gamma \gg \Delta$. In Section 5 of this paper we study the case $N < N_0$, for which small and sparse electron and hole droplets are able to screen the random potential of impurities, as in the three-dimensional case studied in [1–3] (see Fig. 1). We briefly review our results here. In this case, conductivity is also due to the three-mechanism sequence. The high temperature mechanism is due to free electrons activated by energy Δ and the low temperature mechanism is ES VRH with new $T_{\text{ES}} = \alpha\Delta$. However, activated NNH of the intermediate temperature regime is replaced by the new hopping mechanism, which we call NNH-VRH hybrid (H) mechanism. It works in the new temperature interval $T_2' \ll T \ll T_1'$. We study this hybrid mechanism in Section 5. Here we only want to give a hint to its physics and origin of the term “hybrid.” Indeed, if we focus on puddles this is hopping between nearest neighbor puddles. However, each puddle has many droplets. Therefore, focusing on droplets we deal with VRH.

Optimization of Miller–Abrahams resistor network [2] of all available pairs of droplets of two adja-

cent puddles leads to the new H-mechanism conductivity:

$$\sigma = \sigma_2 \exp[-(T_H/T)^{5/9}], \quad (11)$$

where

$$T_H = \alpha^{6/5} (N/N_0)^{1/5} \Delta. \quad (12)$$

Our results for T_{ES} , E_C , and T_H as a function of the dimensionless impurity concentration $N/N_0 = (\Gamma/\Delta)^3$ are summarized in Fig. 4. So far in this section we did not touch the specifics of TI films discussed below, but qualitatively the results for both cases are similar.

However, such three-mechanisms sequences are not observed in most experiments [5–33]. Instead, experiments tend to report an activated conductivity with activation energy much smaller than Δ .

Here we suggest a possible explanation for such low activation energies. We show below that at $\Gamma/\Delta > \alpha^{-9/41}$ electrons are not localized in single puddles and the first two regimes of conductivity are absent. The only remaining mechanism is the ES VRH with very small T_{ES} . This means that the low temperature “local activation energy” is much smaller than Δ .

An alternative explanation involves the intermediate temperature regimes at $1 < N/N_0 < \alpha^{-27/41}$ or at $\alpha^4 < N/N_0 < 1$. In this case all three mechanisms are present in principle, but at small enough α the intervals $T_2 < T < T_1$ and $T_2' < T < T_1'$ can be large and the observed activation energy can be very small. However, the theoretical prefactor of the NNH conductivity and hybrid hopping conductivity $\sigma_2 \ll e^2/\hbar$. This agrees with some experiments [18, 21], but contradicts to other ones, where prefactor close to $e^2/\hbar$ was observed [16, 20]. For such experiments the ES mechanism seems to provide a better explanation.

3. FRACTAL GEOMETRY OF PUDDLES AND TUNNELING ACTION

Let us start from the brief review geometrical fractal properties of 2D puddles at $\Gamma \gg \Delta$ [53]. The characteristic size (diameter) of a puddle is given by

$$a = \lambda (\Gamma/\Delta)^v, \quad (13)$$

where $v = 4/3$ and λ is the electron screening radius. The perimeter of a puddle reads

$$L = a(\Gamma/\Delta) = \lambda (\Gamma/\Delta)^{v+1}. \quad (14)$$

The perimeter L is parametrically longer than the diameter a because puddles have many “fingers,” which are interlocked with other fingers of neighboring puddles (see Fig. 5). The area of a puddle is given by

$$A = \lambda^2 (\Gamma/\Delta)^{2v-\beta}, \quad (15)$$

where $\beta = 5/36$. The separation distance between nearest-neighbor electron and hole puddles is

$$w = \lambda\Delta/\Gamma. \quad (16)$$

In order to estimate Γ and λ we can use the self-consistent theory of [54], which dealt with the disorder potential at the surface of a bulk TI created by charged impurities with three-dimensional concentration N . In our case the substrate plays the role of the TI bulk and the two-dimensional semiconductor plays the role of the TI surface. The band gap Δ that exists in our case is not important when $\Gamma \gg \Delta$. To begin, we relate Γ to λ as the typical Coulomb energy created by charge fluctuations in a volume λ^3 :

$$\Gamma = \frac{e^2}{\kappa\lambda} (N\lambda^3)^{1/2}. \quad (17)$$

This relation leads to a typical 2D density of states (DOS)

$$g = \kappa^2 \alpha^2 \Gamma / e^4, \quad (18)$$

which in turn leads to the screening radius

$$\lambda = \frac{\kappa}{e^2 g} = \frac{e^2}{\alpha^2 \kappa \Gamma}. \quad (19)$$

Solving Eqs.(17) and (19) for Γ and λ we get [54]

$$\Gamma = \frac{e^2 N^{1/3}}{\kappa \alpha^{2/3}}, \quad (20)$$

$$\lambda = \alpha^{-4/3} N^{-1/3}. \quad (21)$$

Let us now estimate the dimensionless action S (the action in units of $\hbar$) electron accumulate tunneling between nearest neighbor fractal metallic electron and hole puddles separated by narrow insulating gaps (see Fig. 2). The value of S is determined by the tunneling length $r = \Delta/eE$ in the spatially varying electric field E created by impurities:

$$S = \frac{r\Delta}{\hbar v} = \frac{\Delta^2}{eE\hbar v}. \quad (22)$$

It is tempting to use $\Gamma/e\lambda$ for E and arrive at $S = w\Delta/\hbar v = \alpha^{-1}(\Delta/\Gamma)^2$. However, the electric field has strong fluctuations at short distances, so the typical electric field depends on the tunneling distance r .

Since a cube of size r has a typical excess charge $\sqrt{Nr^3}$, the typical electric field associated with the length scale r is $E(r) = e(Nr^3)^{1/2}/\kappa r^2$, which grows with decreasing r . Also, due to the large perimeter length L of puddles we can find rare places where the random electric field is created by a larger-than-average number of excessive charges, $M \gg (Nr^3)^{1/2}$, leading to even larger electric field $E(r) = eM/\kappa r^2$. Below we find the optimal values of M and r which determine S , and we arrive at a value of S value much smaller than the naive estimate $S = \alpha^{-1}(\Delta/\Gamma)^2$. Our optimization procedure is a mesoscopic version of the optimization used in the

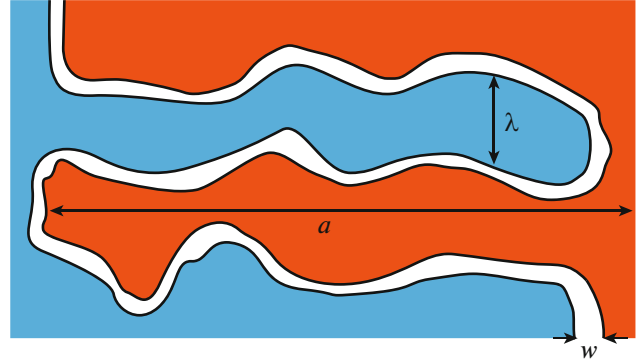


Fig. 5. Schematic picture of interlocked “fingers” of neighboring puddles. Here the length of “fingers” a is of order of the puddle diameter. One can imagine that Fig. 2 shows a vertical cross section of Fig. 5.

theory of the interband absorption of light in compensated three-dimensional semiconductors [2, 55]. It is also similar to the theory of fluctuation-induced excess currents in reverse biased p – n junctions [56].

Below we use S to calculate the localization length ξ that determines hopping transport. Thus, we are interested in fluctuations of electric field which, although rare, happen roughly once at every interface between nearest-neighbor puddles. Thus,

$$(L/\lambda) \exp\left[-\frac{M^2}{Nr^3}\right] = 1. \quad (23)$$

Here we use the Gaussian probability of finding net charge M in a cube of size r . For tunneling across the gap 2Δ we need the potential difference across the cube $Me^2/\kappa r = \Delta$. In other words, $r = r(M) = Me^2/\kappa\Delta$. Substituting $r(M)$ into Eq. (23) and solving for M gives

$$M = \frac{\alpha^{-2}(\Delta/\Gamma)^3}{\ln[(\Gamma/\Delta)^{7/3}]}, \quad (24)$$

which at $\Gamma \gg \Delta$ corresponds to $r(M) \ll w \ll \lambda$.

Substituting the electric field $E = Me/\kappa r^2(M)$ into the tunneling action Eq. (22) we have

$$S = \frac{\alpha^{-1}(\Delta/\Gamma)^3}{\ln[(\Gamma/\Delta)^{7/3}]} \approx \alpha^{-1}(\Delta/\Gamma)^{34/9}. \quad (25)$$

In the last step we used the power-law approximation $\ln x = x^{1/3}$ valid for $x \in (3, 100)$ with accuracy better than 30%.

Now we can calculate the electron localization length, ξ , which we need below to calculate the hopping conductivity. After each tunneling through the gap, electron spreads by distance a . This means that at a large distance x the electron accumulates an action $Sx/a = x/\xi$ where

$$\xi = a/S = \alpha a(\Gamma/\Delta)^{34/9}. \quad (26)$$

The fast decrease of S with growing Γ/Δ leads to growth of dimensionless conductance G between two neighboring puddles

$$G = (L/\lambda) \exp(-S), \quad (27)$$

so that we get $G=1$ at some critical value $(\Gamma/\Delta)_c$.

Substituting Eq. (25) into Eq. (27) and setting $G=1$, in terms of power law we arrive at the critical point¹

$$(\Gamma/\Delta)_c = \alpha^{-9/41}, \quad (28)$$

valid for $\alpha \in (1.2 \times 10^{-4}, 0.12)$. This range of α is obtained by substituting Eq. (28) into the requirement for the argument of the logarithm $(\Gamma/\Delta)^{7/3} \in (3, 100)$.²

At larger Γ/Δ the localization length grows exponentially as $\xi = ae^G$. This leads to dramatic growth of the conductivity, namely to insulator–almost metal transition if the sample size is much larger than ξ . For a very small sample, this $(\Gamma/\Delta)_c$ effectively plays the role of the critical value of insulator–metal transition.

4. HOPPING CONDUCTIVITY

The character of the conductivity of our system apparently changes at $\Gamma/\Delta = (\Gamma/\Delta)_c$. At moderate disorder when $1 < \Gamma/\Delta < (\Gamma/\Delta)_c$ electrons are well localized within a puddle and the temperature dependence of the conductivity follows the three-mechanism sequence discussed in the Introduction. In strong disorder case $\Gamma/\Delta > (\Gamma/\Delta)_c$, the localization length $\xi \gg a$ and at all temperatures the conductivity is due to ES VRH with very small T_{ES} and the prefactor $\sigma_0 \sim e^2/\hbar$. In the limited temperature range it can look like activated transport with very small activation energy.

Below we concentrate on the three-mechanism sequence case, when with decreasing temperature the activated conductivity with activation energy Δ is replaced first by NNH and then by ES VRH. This case reminds systems of densely packed metallic granules separated by a thin insulator with Coulomb impurities and we can follow the calculation of their conductivity [51, 52].

Let us start from the discussion of NNH conductivity. By decreasing the temperature such that $\Delta/T \gg S$ or $T \ll T_1 = \Delta/S$, NNH starts playing role and replacing the Δ activation energy by the charging energy of a puddle E_C . In the case of large Γ/Δ we study the fractal structure of puddles which leads to a

peculiar expression for E_C , smaller than standard $E_C = e^2/\kappa a$. Namely we are going to show that

$$E_C = \frac{e^2}{\kappa L} = \frac{e^2 \Delta}{\kappa a \Gamma}. \quad (29)$$

Substituting Eqs. (13) and (19) into Eq. (29) one arrives at Eq. (4). Let us illustrate how this happens comparing the self-capacitance of an isolated puddle $C_0 \sim \kappa a$ with the capacitance of the same puddle surrounded by other puddles, C . In the latter case, extra electron charge of a puddle e is located at the distance of screening radius λ from its border of the length L , while neighboring metallic puddles provides opposite charge on the other side of the border. Thus, all electric field is concentrated at the border between two puddles, mostly between long fingers of electron and hole puddles shown in Fig. 5.

In our system it means that $C \sim \kappa L$ and leads to Eq. (29). The role of fingers interaction in creating large capacitance was also recognized by electrical engineering community [57].

The use of the activation energy E_C is justified when it is larger than the energy level spacing in a puddle. The level spacing is given by

$$\delta = (gA)^{-1} = \alpha^2 (\Delta/\Gamma)^{55/36} \Delta, \quad (30)$$

where g is the 2D DOS given by Eq. (18) and A is the area of a puddle given by Eq. (15). Therefore the ratio $\delta/E_C = (\Delta/\Gamma)^{7/36} \ll 1$ and our use of E_C is legitimate.

Let us now switch to VRH conductivity which replaces NNH one at low enough temperature enough temperature. In the ground state, each puddle i of our system is charged by a random fractional charge $|q_i| \leq e/2$. This happens because some impurities contribute their potential to neighboring puddles effectively by sharing their charge between neighboring puddles, so that each puddle effectively gets a fraction of impurity charge e . On the other hand, electrons contribute their integer charge e to their puddles. Fractional charging provides background disorder and creates random potential resulting in background density of localized states in which Coulomb attraction between excited electrons and remaining at its initial place hole produces the Coulomb gap around the Fermi level [51, 52]. This leads again to ES law in the low temperature limit.

We can calculate T_{ES} in ES law starting from the standard expression $T_{ES} = e^2/\kappa \xi$ [2, 58]. Using $\xi = a/S$ and Eq. (25) we arrive at Eq. (8). We see now that $T_{ES} \ll \Delta$. Equating $(T_{ES}/T)^{1/2}$ to E_C/T with help of Eqs. (8) and (29) we arrive at T_2 given by Eq. (10).

Note that at $T = T_2$ the typical hop length of ES VRH is $\xi(T_{ES}/T)^{1/2} = L \gg a$ so that the range of applicability of Eq. (29) goes beyond the range a of NNH. The reason for this is that in the case $\Gamma \gg \Delta$ the energy of the Coulomb interaction between electron and hole

¹ In the limit of $\alpha \rightarrow 0$, the asymptotic expression to first order reads $(\Gamma/\Delta)_c = \alpha^{-1/3} [\ln(\alpha^{-1})]^{-2/3}$.

² Note that this mesoscopic optimization method based on Eq. (23) is selfconsistent if $\Gamma/\Delta < (\Gamma/\Delta)_c$ (or $G < 1$), so that $e^S > L/\lambda$.

at distance r determining ES VRH $V(r)$ has a peculiar form: $V(r) = e^2/\kappa r$ at $r \gg L$ and $V(r) = e^2/L$ at $r \ll L$.³

For such a potential ES law crosses over to activated behavior Eq. (29) when hop length r becomes smaller than L [59]. Thus, NNH is responsible only for the high temperature part of the temperature range of validity of Eq. (29), the second low temperature one is ES VRH corresponding to $V(r) = e^2/\kappa L$.

Above we studied the simplest case of the gapped Dirac spectrum Eq. (1). In an important case of BLG gapped by perpendicular external electric field,⁴ the spectrum is somewhat different, namely, it has the “mexican hat” shape, where the energy minimum $|\epsilon(\mathbf{k})| = \Delta$ is degenerate and located along a ring $|\mathbf{k}| = k_0$ in 2D $\mathbf{k}$ -space [60]. Nevertheless, at $|\epsilon(\mathbf{k})| \gg \Delta$ the spectrum returns to the same Dirac cone as Eq. (1) and $|\epsilon(\mathbf{k})| \sim \Delta$ is the only characteristic low energy scale. This is why Eqs. (20) and (21) are still valid and the order of magnitude of S is not changed. Thus, all our results for the most interesting case. $\Gamma \gg \Delta$ are still the same as for the gapped Dirac spectrum Eq. (1).

5. MODEST CONCENTRATION OF IMPURITIES

Above we assumed the gap $\Gamma \gg \Delta$ so that we used the Dirac dispersion $\epsilon(k) \approx \hbar v k$ to calculate energies E_C and T_{ES} . In this section we study the opposite case when $\Gamma \ll \Delta$ or $N \ll N_0$, where N_0 is given by Eq. (6). In this case, Γ given by Eq. (20) does not describe the potential fluctuation amplitude and Γ can be considered only as a measure of N . Indeed, when the Fermi level is within the gap there is no screening, unless the Coulomb potential bends the conduction band bottom and the valence band ceiling by an energy slightly larger than Δ and creates small electron and hole droplets.

³ Here we would like to compare energies of two likely configurations of electric field produced by electron and hole located in the plane of a 2D semiconductor at distance r from each other. First configuration corresponds to electric field lines connecting two charges through 3D space with dielectric constant κ . In this configuration the energy stored in the electric field is $e^2/\kappa r$. In the second configuration electric field stays in the plane of the semiconductor with dielectric constant κ . It runs through metallic puddles and connects their perimeters of length $\sim L$, so that puddles form capacitors with capacitance $C \sim \kappa L$. In a typical linear cross-section of the area r^2 there are r/a parallel capacitors with charge ea/r in each of them. Here a is the diameter of a puddle. Each capacitor therefore carries an energy $(ea/r)^2/C$. The total number of involved capacitors is $(r/a)^2$. Thus, total energy of this capacitor network is $e^2/\kappa L$. It is clear now that at $r > L$ electric fields prefer to stay in 3D space leading to $V(r) = -e^2/\kappa r$ for electron–hole interaction energy, while at $r < L$ electric fields stay inside the plane and $V(r) = -e^2/\kappa L$ is independent on r . (This argument ignores logarithmic factors.)

⁴ Graphene has a Fermi velocity $v = 1 \times 10^6$ m/s, and in order to apply our theory we need $\alpha = e^2/\kappa \hbar v = 2.2/\kappa \ll 1$. Namely, the dielectric constant of the environment surrounding the BLG should be $\kappa \gg 2.2$.

Electron and hole droplets form alternating in space fractal electron and hole clouds (puddles) of the size

$$\Lambda = \kappa^2 \Delta^2 / e^4 N, \quad (31)$$

obtained by equating Δ and the random potential amplitude $(e^2/\kappa)(N\Lambda^3)^{1/2}/\Lambda$ inside a cube of size Λ . Such a system of droplets and puddles is the two-dimensional analog of the three-dimensional completely compensated semiconductor [1, 2] schematically shown in Fig. 1.

If the kinetic energy of degenerate electrons in droplets satisfies $\epsilon(k) \ll \Delta$, then one can use the parabolic dispersion law for them

$$\epsilon(k) = \hbar^2 k^2 / 2m, \quad (32)$$

with $m = \Delta/v^2$. To show that indeed $\epsilon(k) \ll \Delta$ we first find the size of a typical droplet R_q , following [1, 2]. Namely, we equate the depth of the potential well $(e^2/\kappa)(NR_q^3)^{1/2}/R_q$ created by a typical fluctuations of charge in a cube of size R_q to the kinetic energy of $(NR_q^3)^{1/2}$ electrons $\epsilon = (NR_q^3)^{1/2} \hbar^2 / m R_q^2$ in the disk of radius R_q and arrive at $R_q = a_B$, where $a_B = \hbar^2 \kappa / m e^2$ is the semiconductor Bohr radius and $m = \Delta/v^2$ is the effective mass.⁵ When deriving $R_q = a_B$ we assumed $N a_B^3 = (\Gamma/\Delta)^3 \alpha^{-4} \gg 1$, or equivalently $1 \gg \Gamma/\Delta \gg \alpha^{4/3}$. Substituting $R_q = a_B$ back to $\epsilon = \hbar^2 (NR_q^3)^{1/2} / m R_q^2$ we get

$$\frac{\epsilon}{\Delta} = \left(\frac{e^2 N^{1/3}}{\alpha^{2/3} \kappa \Delta} \right)^{3/2} = \left(\frac{\Gamma}{\Delta} \right)^{3/2} = \left(\frac{N}{N_0} \right)^{1/2} \ll 1. \quad (33)$$

Next we discuss the conductivity when $N \ll N_0$ or $\Gamma \ll \Delta$. Similarly to the case $\Gamma/\Delta \gg 1$ there are three mechanisms of the conductivity. Activation of free electrons by energy Δ is the same as at $N > N_0$. The tunneling between droplets inside the same puddle is faster than the tunneling between neighbor puddles. For the latter we can find the action S substituting Λ for r into Eq. (22):

$$S = \alpha^{-1} (\Delta/\Gamma)^3 = \alpha^{-1} N_0 / N. \quad (34)$$

This leads to the crossover from activation to hopping at $T_1' = \Delta/S$ or

$$T_1' = \alpha (N/N_0) \Delta. \quad (35)$$

In the adjacent interval of lower temperatures $T_1' \gg T \gg T_2'$ where

⁵ The same result can be obtained by equating the total number of electron states $g\gamma(R_q)R_q^2$ to the excess number of impurity charges $\sqrt{NR_q^3}$, where $\gamma(R_q) = (e^2/\kappa)(NR_q^3)^{1/2}/R_q$ is the potential depth and $g = m/\hbar^2$ is the 2D DOS.

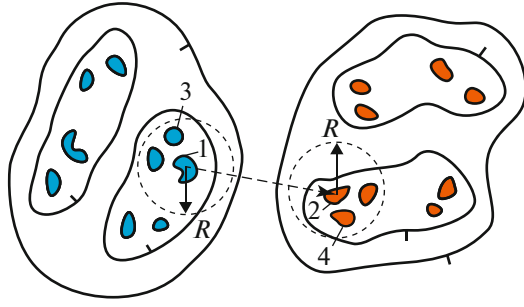


Fig. 6. Schematic map of nearest neighbor electron and hole puddles containing many electron (blue) and hole (red) puddles. The continuous lines are equipotential contours of the electron energy. As in geographical maps the direction of descent is indicated by a short stroke. The smallest contours represent boundaries of droplets at the chemical potential. The dashed arrow shows the shortest hop between the two puddles. At $T \ll T_1'$ electron searches in dashed circles of radius R for droplets 3 and 4 with closer to the chemical potential energies, which provide a smaller inter-puddle hop resistance.

$$T_2' = \alpha^3 (N/N_0)^2 \Delta, \quad (36)$$

we deal with the hopping conductivity of electrons between nearest puddles, but each puddle has many droplets and we should explore hopping between all pairs of droplets of two neighboring puddles searching for the smallest Miller-Abrahams resistors [2]. Thus, focusing on puddles we deal with NNH and focusing on droplets we deal with VRH. Therefore, we call this mechanism of the hybrid hopping conductivity and illustrate it in Fig. 6. At $T_3 \ll T \ll T_1'$, where as we show below

$$T_3 = \alpha^3 (N_0/N)^{1/4} \Delta \gg T_2' \quad (37)$$

an electron hops between two closest droplets 1 and 2 of the nearest neighbor electron and hole puddles. When temperature gets lower it chooses slightly more distant droplets of the same two puddles with energies closer to the Fermi level. At low enough temperature, $T = T_2'$, it chooses the optimal in energy two droplets of the two puddles. At even lower $T < T_2'$ it hops to the second nearest puddle and we arrive at ES VRH. Such a theory of hybrid hopping conductivity was proposed for strongly compensated semiconductors [61] before the discovery of the Coulomb gap and ES law [58]. Here we revise this theory by adding the Coulomb gap to its argumentation.

To find the optimal path at $T_2' \ll T \ll T_3$, an electron searches for more distant droplets 3 and 4 located in two discs of radius R around droplets 1 and 2. This leads to the increase $\delta S = R/\xi(R)$ of the tunneling action S , where $\xi(R) = \hbar/(m\gamma(R))^{1/2}$ and $\gamma(R) = (e^2/\kappa R)(NR^3)^{1/2}$ is height of typical barrier in the spatial scale R . This increase can be overcompensated by

the reduction of activation energy necessary for the hop between new pair of droplets. Using the two-dimensional Coulomb gap density of states $g(\epsilon) = \kappa^2 \epsilon / e^4$ we find that minimum energy $\epsilon_{\min}$ of a droplet in a disc of area R^2 is determined by the condition $\epsilon = [g(\epsilon)R^2]^{-1}$, which leads to $\epsilon_{\min} = e^2/\kappa R$.

Thus, the logarithm of conductivity generated by optimization in a given scale R is

$$\ln \sigma(R) = -S - R[m\gamma(R)]^{1/2}/\hbar - e^2/\kappa RT. \quad (38)$$

It has maximum at

$$R_{\text{opt}} = e^2 \kappa^{-1} T_H^{-5/9} T^{-4/9}, \quad (39)$$

where $T_H = \alpha^{6/5} (N/N_0)^{1/5} \Delta$ as shown in Eq. (12). Substituting R_{opt} back to Eq. (38) we arrive at Eq. (11). Equating R_{opt} to the maximum distance Λ between droplets inside a puddle we arrive at T_2' in Eq. (36). On other hand, equating R_{opt} to the minimum distance a_B between droplets inside a puddle we arrive at the high temperature limit of the hybrid hopping conductivity T_3 given by Eq.(37).

When N tends to N_0 the temperature T_3 tends to T_2' , so that the range of validity of Eq. (11) vanishes. This allows the conductivity we calculated in Section 5 for $N > N_0$ to match at $N = N_0$ the conductivity we found for $N < N_0$ in Section 4.⁶

At lower temperatures $T < T_2'$, the conductivity is dominated by ES VRH. The localization length is given by

$$\xi = \Lambda/S, \quad (40)$$

with Λ given by Eq. (31) and S given by Eq. (34). Substituting Eq. (40) into $T_{\text{ES}} = e^2/\kappa \xi$, we get the characteristic temperature for ES VRH

$$T_{\text{ES}} = \alpha \Delta, \quad (41)$$

which is valid if $\alpha^4 < N/N_0 < 1$ and matches Eq. (8) at $\Gamma = \Delta$ and $N = N_0$.

Our results for T_{ES} and E_C as a function of dimensionless impurity concentration $N/N_0 = (\Gamma/\Delta)^3$ obtained in Sections 4 and 5 are summarized in Fig. 4.

So far we assumed that $\alpha \ll 1$. Let us explore what happens when $\alpha \sim 1$ as, for example, in twisted bilayer graphene. In this case at $N_0 = \kappa^3 \Delta^3 e^{-6} = a_B^{-3}$ and at $N = N_0$ we have $N a_B^3 = 1$. Both sides of $N/N_0 = 1$ on Fig. 4 now shrink. At $N/N_0 \gg 1$ or $N a_B^3 \gg 1$ we get almost metallic conductivity, in spite of complete compensa-

⁶ Above for simplicity we concentrated on the case of $N > \alpha^{8/5} N_0$, when $T_3 > T_1'$. In this case, in the range $T_1' > T > T_3$ conductivity has a constant activation energy. At $N < \alpha^{8/5} N_0$ the high temperature border of the hybrid hopping conductivity Eq. (11) is given by T_1' .

tion. On the other side, $N/N_0 \ll 1$, the conductivity obeys ES law with $T_{ES} = \Delta$ at low temperature $T \ll (N/N_0)^2 \Delta$. At higher temperatures interval $(N/N_0)^2 \Delta \ll T \ll (N/N_0) \Delta$ the conductivity is due to the hybrid hopping described by Eq. (11) with $T_H = \Delta(Na_B^3)^{1/5}$. At $T \gg (N/N_0) \Delta$ the conductivity is dominated by free electrons activated by energy Δ . As a result, at low enough temperature the local apparent activation energy is of the order of $(T\Delta)^{1/2} \ll \Delta$. Note also that at $\alpha \sim 1$ the Bohr energy of hydrogen like impurities $E_B = e^2/\kappa a_B = \alpha^2 \Delta$ becomes of the order of Δ so that localized donor and acceptor states play the role of metallic droplets in the lightly doped limit $N \ll N_0$ (see Chapter 3 of [2]).

6. “PHASE DIAGRAM” OF TRANSITION BETWEEN INSULATOR AND ALMOST METAL

Above we focused on the charge neutrality point where $E_F = 0$. However, for 2D devices one can easily move away from the neutrality point by applying a gate voltage, for example, making $E_F > 0$ and inducing a nonzero net 2D concentration n of electrons. As a result one can study the whole n – N phase diagram. We show below that such a phase diagram has an interesting re-entrant metal–insulator transition (MIT) as a function of increasing N (see Fig. 7).

We already know from Eq. (28) that if $n = 0$, there is an insulator–metal transition (IMT) at $\Gamma/\Delta = (\Gamma/\Delta)_c$, or $N = N_c = N_0 \alpha^{-27/41}$, where the disorder tunneling plays decisive role. On the other hand, for modest impurity concentrations $1 \ll Na_B^3 \ll N_0 a_B^3 = \alpha^{-4}$, there is also an MIT induced at some $n = n_c$. This MIT happens because by increasing n the electron kinetic energy becomes larger than the random Coulomb potential energy, and electron puddles become connected with each other and turn the system from an insulator to an almost metal. We find the net electron concentration n_c associated with the MIT by equating the Fermi energy $E_F = \hbar^2 n/m$ with the Coulomb potential energy fluctuations [62]

$$\frac{\hbar^2 n}{m} = \frac{e^2 \sqrt{Na_B^3}}{\kappa a_B}, \quad (42)$$

where the Bohr radius a_B plays the role of the linear screening length r_s for nonrelativistic 2D electron gas. Solving Eq. (42) one gets the critical percolation threshold concentration

$$n = n_c = (N/a_B)^{1/2}. \quad (43)$$

In the left side of the phase diagram Fig. 7, the MIT border line follows Eq. (43) shown by the dash line. It continues till the maximum at $N \sim N_0$, where $E_F \sim \Gamma \sim \Delta$. Therefore, for the whole phase diagram $E_F \leq \Delta$ and

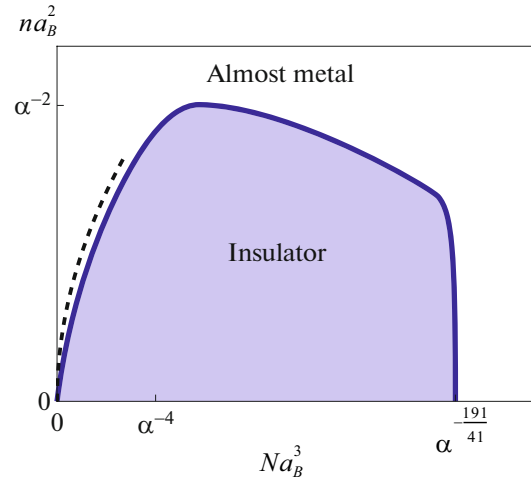


Fig. 7. Schematic n – N phase diagram in a two-dimensional semiconductor. The shaded blue domain is the insulator phase while the white domain is the almost metal phase. On the left (small N) side of the diagram the phase boundary follows Eq. (43) (dashed line) and reaches the maximum near $Na_B^3 = N_0 a_B^3 = \alpha^{-4}$. On the right side the maximum of the phase boundary is determined by criterion $G(n, N) = 1$ for tunneling between electron puddles. When with decreasing n this tunneling rate yields to the tunneling between electron and hole puddles, the boundary becomes vertical, i.e., sticks to the critical point $Na_B^3 = N_c a_B^3 = \alpha^{-191/41}$ all the way till $n = 0$ (c.f. Eq. (28)). We use $\alpha = 0.12$ for this plot.

our use of the nonrelativistic expressions for E_F and r_s in Eq. (42) is justified.

Now we turn to the right side of Fig. 7 and find the IMT border line, which starts at $n = 0$ and $N = N_c$ and matches Eq. (43) near $N = N_0$. Apparently, this part of the border is determined by tunneling between puddles. This tunneling can be illustrated by the small square section of the “chessboard” potential $u(x, y) = u_0 \cos(2\pi x/b) \cos(2\pi y/b)$ which at neutrality point has four identical electron and hole puddles (see the $b \times b$ square in Fig. 8). There are two kinds of electron tunneling events contributing in the conductivity of such a chess-board, the side-to-side tunneling between electron and hole puddles and the diagonal tunneling through the saddle point between two neighboring electron puddles.

At $n = 0$ (see Fig. 8a) in both cases the tunneling barrier is of the same height Δ , while the tunneling distance in the side-to-side case, w is substantially shorter than the diagonal distance W . Thus, at $n = 0$ the side-to-side tunneling dominates and so far in this paper we dealt only with this tunneling. The side-to-side tunneling distance w is weakly affected by growing n (see Fig. 8b). As a result the right side of the IMT border at small n is almost a vertical line. On the other hand, we see in Fig. 8b that the distance between electron pud-

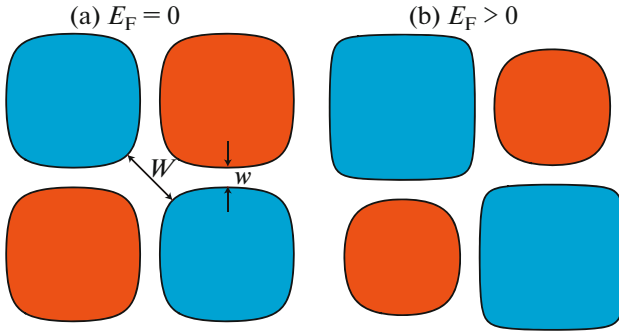


Fig. 8. Illustration of the competition of the two tunneling rates for a chessboard potential. Blue and red domains are the electron and hole puddles separated by insulating gap (white). (a) At the neutrality point, $n = 0$, $E_F = 0$, the shortest tunneling distance between electron puddles W is much larger than the distance w between electron and hole puddles. (b) At $E_F > 0$ and growing n , W decreases. Eventually it becomes smaller than w and vanishes at the percolation transition where all electron puddles merge into the infinite cluster.

dles W steeply decreases with growing n . Eventually, W becomes smaller than w and the diagonal tunneling starts dominating.

Although in Fig. 8 we used a simplified model, the same dynamics is valid for our disordered system. The crossover to the diagonal tunneling leads to a sharp decrease of the slope of the right side border of the insulating phase of Fig. 7, given by the condition $G(n, N) = 1$ for the dimensional conductance. Eventually, with growing n large fraction of diagonal gaps between electron puddles close near $na_B^2 \sim \alpha^{-2}$, where $E_F \sim \Delta \sim \Gamma$ and electron puddles merge at the percolation transition. This means that the right side border merges with the left one near $N \sim N_0$ as shown in Fig. 7.

We see that at $1 < na_B^2 < \alpha^{-2}$ there is an unconventional reentrant MIT. At small $N \ll N_0$ disorder drives a percolation MIT, when Γ exceeds Fermi energy of electrons so that they are forced into isolated puddles. On other hand, at large $N > N_0$ or $\Gamma > \Delta$ the disorder drives electrons back to metal, because of increase of tunneling which becomes dramatic near $\Gamma/\Delta = (\Gamma/\Delta)_c$. Of course, making $n < 0$ one can explore re-entrant MIT phase diagram for holes.

Changing the concentration of impurities N at fixed n and Δ in real experiments is difficult. Instead, one can change the band gap Δ at fixed n and N as this is done, for example, in [63]. However, in this case, with decreasing Δ there is only one MIT at small Δ as in the case of $n = 0$. There is no second transition at large Δ , because as we see from Eq. (43) using $a_B = e^2/(\kappa\alpha^2\Delta)$, the critical concentration of MIT $n_c \propto \Delta^{1/2}$ eventually becomes larger than any fixed n , so that the insulator phase persists. One can return to a re-entrant

MIT transition at fixed N by changing Δ and n simultaneously, for example, by fixing na_B^2 as this is clear from Fig. 7.

We would like to emphasize that predicted in this section reentrant MIM transition is different from MIM transition observed near filling factor $\nu = 1$, where the Mott–Hubbard gap has maximum [64]. Along a horizontal line of Fig. 7, Δ and n are fixed and the concentration of impurities N changes monotonically. Monotonically growing random potential then leads to the classical MIT via creating puddles when N is relatively small, and to the quantum IMT transition via enhanced tunneling when N is big. For the transition discussed in our previous paragraph, N is fixed and both n and Δ change monotonically, such that two transitions are due to the interplay of two different, classical and quantum, disorder effects. Our MIM transition exists when there is no Mott–Hubbard physics.

7. THIN FILM OF THREE-DIMENSIONAL TOPOLOGICAL INSULATOR

In previous sections we dealt with the general model of the trivial 2D semiconductor with gaped Dirac spectrum Eq. (1). In this section we concentrate on special case of a thin film of 3D TI, where the narrow gap 2Δ can be a result of the hybridization of surface states on opposite surfaces of the film [8, 9, 37–43] or created by a concentration of magnetic dopants like Cr [10, 12–24, 34, 35, 44–49]. Because of the promise of such films to achieve a metrological precision of the quantum anomalous Hall effect and the quantum spin Hall effect, in this section we are more specific with material parameters and numerical estimates.

We have in mind TI thin films based on $(\text{Bi}_x\text{Sb}_{1-x})_2\text{Te}_3$, which have very large dielectric constant $\kappa \sim 200$ [65–67]. Using $\kappa \sim 200$ and the Fermi velocity of TI $v \sim 4 \times 10^5$ m/s [68], one gets $\alpha \sim 0.027$. We assume that such a film of width $d \sim 7$ nm is deposited at the substrate with much smaller dielectric constant κ_e , so that electric fields of Coulomb impurities residing inside the film are trapped within the film [69–73]. Namely, electric lines stay inside the film at distances r from impurity if $d < r < d\kappa/\kappa_e$ and in the absence of screening exit the film at $r > d\kappa/\kappa_e$. Thus, the electric field of a charged impurity is $E(r) = 2e/\kappa dr$ at $d < r < d\kappa/\kappa_e$ and $E(r) = e/\kappa_e r^2$ at $r > d\kappa/\kappa_e$. This leads to electrostatic potential $V(r) = (e^2/\kappa d)\ln(\kappa d/r)$ weakly dependent on r in the range $d < r < d\kappa/\kappa_e$.

At $\Gamma > \Delta$ electrons and holes created by band bending self consistently screen electric field at the distance λ , such that $d \ll \lambda \ll (\kappa/\kappa_e)d$. Very slow decay of the electrostatic potential of distant Coulomb impurities allows larger number of them to contribute to Γ . TI films have large concentrations of Coulomb impuri-

ties, $N \sim 10^{19} \text{ cm}^{-3}$, so that we study only their effect and ignore impurities inside the substrate.

Peculiar electrostatics we discussed above leads to new results of self-consistent theory of screening [72] at $E_F = 0$ and $\Gamma \gg \Delta$

$$\Gamma = \frac{e^2 N^{1/3}}{\kappa \alpha^{2/3}}, \quad (44)$$

$$\lambda = \alpha^{-2/3} (Nd^3)^{-1/6} d. \quad (45)$$

Note that the expression for Γ in this case is the same as Eq. (20), i.e. while the expression for λ differs from Eq. (21) and is valid if $\lambda \gg d$. Using above estimates for α , d , and N we get $\lambda \sim 50 \text{ nm}$, so that inequality $\lambda \gg d$ holds and we can use Eq. (45).

Similarly to the Section 4, below we calculate the tunneling action S and the critical $(\Gamma/\Delta)_c$ and describe hopping conductivity of the film. Notice the electric field in the film plane created by charge fluctuations in a disk of radius r and thickness d is given by

$$E = \frac{e\sqrt{Nr^2d}}{\kappa rd} = \frac{e}{\kappa} \frac{\sqrt{N}}{\sqrt{d}}, \quad (46)$$

which turns out to be independent on r . Therefore, there is no, similar to the one in Section 4, enhancement of the electric field at scales shorter than λ . Substituting Eq. (46) or equivalently $E = \Gamma/e\lambda$ into Eq. (22), we arrive at the action

$$S = \frac{w\Delta}{\hbar v} = \frac{\lambda\Delta^2}{\hbar v\Gamma} = \alpha^{-1/3} (Nd^3)^{1/6} (\Delta/\Gamma)^2. \quad (47)$$

However, the electric field $E = eM/\kappa\lambda d$ can still be enhanced by a rare fluctuation of the number of charges $M \gg (N\lambda^2 d)^{1/2}$ with Gaussian probability $\exp(-M^2/N\lambda^2 d)$. This replaces Eq. (23) by

$$(L/\lambda) \exp\left[-\frac{M^2}{N\lambda^2 d}\right] = 1. \quad (48)$$

Solving the above equation we obtain the largest M available in the perimeter

$$M = \{N\lambda^2 d \ln[(\Gamma/\Delta)^{7/3}]\}^{1/2}. \quad (49)$$

Substituting the electric field $E = eM/\kappa\lambda d$ into the action Eq. (22) we obtain

$$S = \frac{\alpha^{-1/3} (Nd^3)^{1/6} (\Delta/\Gamma)^2}{\{\ln[(\Gamma/\Delta)^{7/3}]\}^{1/2}} \approx \alpha^{-1/3} (Nd^3)^{1/6} (\Delta/\Gamma)^{43/18}, \quad (50)$$

which is smaller than the action given by Eq. (47). In the last step, as in the Section 3 we used the power-law approximation $\ln x = x^{1/3}$ valid for $x \in (3, 100)$ with accuracy better than 30%.

Substituting Eq. (50) into the expression of G , Eq. (27), and setting $G = 1$, we arrive at the critical point⁷

⁷ In the limit of $\alpha(Nd^3)^{-1/2} \rightarrow 0$, the asymptotic expression to first order reads $(\Gamma/\Delta)_c = \alpha^{-1/6} (Nd^3)^{1/12} \{\ln[\alpha^{-1} (Nd^3)^{1/2}]\}^{-3/4}$.

$$(\Gamma/\Delta)_c = \alpha^{-2/19} (Nd^3)^{1/19}. \quad (51)$$

Using the estimates $\alpha \sim 0.027$, $N \sim 10^{19} \text{ cm}^{-3}$ and $d = 7 \text{ nm}$ we get $(\Gamma/\Delta)_c = 1.6$.

Let us switch to hopping conductivity of thin TI film and start from NNH where activation energy is determined by the charging energy of a puddle. Similarly to Section 3, capacitance of a puddle within the film is determined by its long border adjacent to neighboring puddles. Near the border there are two stripes with charges $-e$ and e of the length L and width λ . But now electric field at the border is concentrated in the film of width $d \ll \lambda$ with the large dielectric constant κ . This changes the capacitance of the puddle border to $C \sim \kappa L(d/\lambda)$ and leads to

$$E_c = (e^2/\kappa d)(\lambda/L) = (e^2/\kappa d)(\Delta/\Gamma)^{7/3}, \quad (52)$$

at $1 < \Gamma/\Delta < (\Gamma/\Delta)_c$. This charging energy, plays the role of NNH activation energy.

At lower temperatures $T < T_2$ conductivity obeys the ES law with the characteristic temperature $T_{ES} = e^2/\kappa_e \xi$, where $\xi = a/S$. Note that here we use κ_e because at large distances electric field lines leave the film and go through the environment. Using Eqs. (13) and (50) we get

$$T_{ES} = \alpha(\kappa/\kappa_e)\Delta(\Delta/\Gamma)^{49/18}. \quad (53)$$

Crossover between Eq. (2) and Eq. (3) happens at

$$T_4 = \alpha^{1/3} (Nd^3)^{-1/6} (\Gamma/\Delta)^{43/18} \Delta, \quad (54)$$

while crossover between Eq. (3) and Eq. (7) happens at

$$T_5 = E_c^2/T_{ES} = \alpha^{-1}(\kappa_e/\kappa)(\Delta/\Gamma)^{35/18} \Delta^{-1} (e^2/\kappa d)^2. \quad (55)$$

Equations (53) and (51) show that in TI films as in trivial semiconductors (c.f. Eqs. (8) and (28) and Fig. 4) reduction of T_{ES} and crossover from strong localization case to the practically metallic conductivity happens dramatically fast when Γ exceeds Δ .

Now let us study the case of a modest concentration of impurities where $\Gamma < \Delta$ or $N < N_0$ where N_0 is defined by Eq. (6). First we should find the size of a droplet R_q in the cloud, and show that the kinetic energy of a droplet ϵ is smaller than Δ so that the energy dispersion is nonrelativistic at $\Gamma < \Delta$. Equating the depth of the potential well $e^2(NR_q^2 d)^{1/2}/\kappa d$ created by typical fluctuations of charges in a disk of radius R_q and thickness d to the Fermi energy of $(NR_q^2 d)^{1/2}$ electrons in the disc of radius R_q , $\epsilon = \hbar^2(NR_q^2 d)^{1/2}/mR_q^2$, we arrive at $R_q = \sqrt{a_B d}$. This R_q is valid if $a_B > d$ (or $\Delta < \alpha^{-2} e^2/\kappa d$) so that the droplet is still disk-like. Substituting $R_q = \sqrt{a_B d}$ back to the droplet Fermi energy we get the ratio ϵ/Δ given by Eq. (33). Therefore, at

$N < N_0$ the kinetic energy is smaller than Δ and the use of the parabolic dispersion law is justified.

Next, we study the three-mechanism conductivity at $N < N_0$. Let us start with the NNH conductivity. Similarly to Section 4, we obtain the nonlinear screening length

$$\Lambda = \frac{\Delta \kappa \sqrt{d}}{e^2 \sqrt{N}} \quad (56)$$

(the size of a puddle) by equating Δ and the random potential of a disk of size Λ .

$$\Delta = \frac{e^2}{\kappa d} \sqrt{Nd\Lambda^2}. \quad (57)$$

As a result, the tunneling action S is again given by Eq. (47) and $S \gg 1$. The activation energy for the NNH conductivity is equal to the charging energy of a droplet. Radial electric field of a charged droplet stays in insulating TI film till distance $(\kappa/\kappa_e)d$ and only then exits to the environment with small dielectric constant κ_e . Therefore, the charging energy determining hopping activation energy at $T_4 > T > T_5$ is

$$E_C = e^2 / [\kappa_e (\kappa/\kappa_e) d] = \frac{e^2}{\kappa d}, \quad (58)$$

which matches Eq. (52) at $\Gamma = \Delta$.⁸ At lower temperature $T < T_5$, using the localization length $\xi = \Lambda/S$ we get the characteristic temperature for ES law

$$T_{ES} = e^2 / \kappa_e \xi = \alpha (\kappa/\kappa_e) \Delta, \quad (59)$$

which matches Eq. (53) at $\Gamma = \Delta$.⁹

Now we can estimate the characteristic energies Γ , Δ , E_C , T_{ES} , T_4 , T_5 and Δ_1 for TI thin films based on $(\text{Bi}_x\text{Sb}_{1-x})_2(\text{Te}_y\text{Se}_{1-y})_3$. Using $\kappa = 200$, $\alpha = 0.027$, $N = 10^{19} \text{ cm}^{-3}$, we have $\Gamma \approx 17 \text{ meV}$. The hybridization gap is related to the thickness by $\Delta = \Delta_0 e^{-d/d_0}$ with $\Delta_0 = 0.5 \text{ eV}$ and $d_0 = 2 \text{ nm}$ [9]. For example, if $d = 7 \text{ nm}$, then $\Gamma \approx 17 \text{ meV}$, $\Delta \approx 15 \text{ meV}$, $S \approx 3$, $E_C \approx 0.8 \text{ meV}$, $T_4 = \Delta/S \approx 60 \text{ K}$, $T_5 = E_C^2/T_{ES} = 0.6 \text{ K}$, $T_{ES} \approx 130 \text{ K}$ and $\Delta_1 = 70 \text{ meV}$ (here assume that the film has hBN on both sides and use $\kappa_e = 5$). In this case, ES conductivity starts when $(T_{ES}/T_5)^{1/2} \sim 15$, so large that ES law is hardly observable because of very large resistance.

⁸ This result of E_C is valid if the size of a droplet R_d is smaller than the electric field confinement distance $(\kappa/\kappa_e)d$, or $\Delta > (\kappa_e/\kappa\alpha)^2 e^2/\kappa d$.

⁹ The above theory assumes that $\Lambda < d\kappa/\kappa_e$, or $\Delta < \Delta_1 = (e^2/\kappa_e)\sqrt{Nd}$. In this case, the Coulomb interaction energy between two droplets inside a puddle is equal to $e^2/\kappa d$ and is independent on the distance between two droplets R . Therefore, there is no optimization with respect to R similar to Eq. (38), and no hybrid mechanism. On the other hand if $\Delta > \Delta_1$, there is intermediate between Eqs. (58) and (59) hybrid regime which we skip here.

Thus, observable activation energy is given by $E_C \sim 0.05\Delta$.

In slightly thicker films with $d \geq 8 \text{ nm}$ the half-gap $\Delta(d) \leq 9 \text{ meV}$ and $\Gamma/\Delta > (\Gamma/\Delta)_c$, so that they are almost metallic and show ES conductivity with much smaller T_{ES} . On the other hand, in slightly thinner films, $d < 7 \text{ nm}$, for which $\Delta > \Gamma$ energies E_C and T_{ES} given by Eqs. (58) and (59) match Eq. (52) and (53) so that practically conductivity is similar to films with $d = 7 \text{ nm}$.

Notice that critical thickness $d = d_c = 7 \text{ nm}$ is very sensitive to values of Δ_0 , N , κ , α , and most importantly d_0 , which are different for different materials. This can explain differences between experimental results in [8, 9].

For the case of magnetically doped TI thin films, the exchange half gap Δ induced by magnetic impurities is of order of 20 meV [20, 21, 49] (which is not directly related to d), so that we have practically the same numbers as in the previous example.

Let us now dwell on the n – N plane phase diagram of the TI film, which as we show below looks almost identical to Fig. 7. At a modest impurity concentration $N < N_0$, we equate the Fermi energy with the Coulomb potential energy fluctuations to find the critical concentration n_c

$$\hbar^2 n/m = e^2 \sqrt{Nr_s^2 d} / \kappa d = e^2 \sqrt{Na_B} / \kappa, \quad (60)$$

where we use $r_s = \sqrt{a_B d}$ for the screening radius r_s of nonrelativistic 2D electron gas in a TI thin film of thickness d . This result can be derived by substituting the nonrelativistic DOS $g = m/\hbar^2$ into $r_s = \sqrt{\kappa d/e^2 g}$ [72]. (Note that the expression of r_s looks like the one in 3D bulk semiconductor, because of peculiar electrostatics of TI film with large dielectric constant.) Equation (60) gives the same expression of n_c as Eq. (43), so the left boundary of the TI phase diagram is exactly the same as in Fig. 7. The only difference between TI diagram and Fig. 7 is that the distance between N_c and N_0 is smaller on the right hand side of the TI phase diagram.

8. NARROW GAP COMPLETELY COMPENSATED THREE-DIMENSIONAL SEMICONDUCTOR

In this section we return to the case of a three-dimensional (3D) completely compensated semiconductor with a gapped Dirac dispersion, from which we started the Introduction section. This problem was studied previously [1–3] for relatively large gap semiconductors, when the random potential is screened by small and well-separated electron and hole droplets, as depicted in Fig. 1 (see also Fig. 13.4 of [2]). But the recent interest in three-dimensional Dirac semimetals has brought renewed emphasis on 3D materials with a

Dirac-like dispersion and very small gap. Such a dispersion arises ubiquitously near the boundary between topological and trivial insulator phases. Examples include ZrTe_5 and HfTe_5 [74], BiTeI [75], BiTeBr [76], and BiSb [77, 78]. In these materials the gap Δ can be controlled by external parameters like pressure or magnetic field.

Below we address both cases $\Gamma_1 \ll \Delta$ and $\Gamma_1 \gg \Delta$. Here our notation Γ_1 replaces Γ of previous sections because in the 3D strong disorder case the self-consistent amplitude of the potential is somewhat smaller than for 2D. According to the self-consistent theory of [79], the amplitude of the disorder potential at $\Delta = 0$ follows

$$\Gamma_1 = \frac{e^2 N^{1/3}}{\kappa \alpha^{1/2}}. \quad (61)$$

As in Section 5, below we use the notation Γ_1 only to represent the impurity concentration N . In the limit $\Gamma_1 \ll \Delta$ the actual amplitude of the random potential is of order Δ . We also define the characteristic concentration $N_1 = \alpha^{3/2} e^{-6} \kappa^3 \Delta^3$, such that $\Gamma_1/\Delta = (N/N_1)^{1/3}$.

Below we briefly discuss the temperature dependence of the conductivity at $\Gamma_1 \ll \Delta$, which is similar to the 2D case with $\Gamma \ll \Delta$, discussed in Section 5. We then consider the conductivity in the case when the amplitude of the random potential $\Gamma_1 \gg \Delta$, and we emphasize the dramatic difference between 2D and 3D in this case.

In the case $\Gamma_1 \ll \Delta$, a 3D completely compensated semiconductor is an insulator with three low temperature mechanisms of conductivity. At relatively high temperature electrons and holes can be activated from the Fermi level to their percolation levels. Because in 3D percolation requires only a small fraction of space, $\sim 17\%$, this activation energy is relatively small [3], $E_a \approx 0.3\Delta$.

At lower temperatures activation to the percolation level is replaced by NNH, with activation energy given by the charging energy of a droplet, $E_c = e^2/\kappa R$, where the droplet radius $R = a_B (Na_B^3)^{-1/9}$ for $Na_B^3 \gg 1$ or $N/N_1 \gg \alpha^{9/2}$ and $R = a_B$ for $Na_B^3 \ll 1$ or $N/N_1 \ll \alpha^{9/2}$. This leads to $E_c = \alpha^{3/2} (N/N_1)^{1/9}$ for $N/N_1 \gg \alpha^{9/2}$ and $E_c = \alpha^2 \Delta$ for $N/N_1 \ll \alpha^{9/2}$.

The prefactor of NNH conductivity is equal to $\exp(-S)$, where using Eq. (31) we get

$$S = \frac{\Lambda \Delta}{\hbar v} = \frac{1}{\alpha^{1/2}} \left(\frac{\Delta}{\Gamma_1} \right)^3 = \frac{N_1}{\alpha^{1/2} N}. \quad (62)$$

At the lowest temperatures the conductivity is dominated by ES VRH. Using Eqs. (31), (40), and (62) we get $T_{\text{ES}} = e^2/\kappa \xi = \alpha \Delta$.

Let us now turn to the conductivity in the case $\Gamma_1 \gg \Delta$, for which the bending of conduction and

valence bands looks similar to Fig. 2. In this case electrons occupy almost half of space, but we know that only 17% of the space is enough to provide percolation. This means that in the 3D case, in contrast with the 2D case discussed in Section 3, at $\Gamma_1 \gg \Delta$ we deal with a good metal. There is therefore a critical disorder strength, such that $(\Gamma_1/\Delta)_c \sim 1$, which produces a percolative insulator-to-metal transition (IMT).

Assuming that the random potential energy $u(\mathbf{r})$ follows the Gaussian distribution, we can estimate $(\Gamma_1/\Delta)_c$ and the critical concentration N_c of IMT by equating the volume fraction of electron (or hole) puddles to 17%. Since the condition to have hole puddles is $u > \Delta$ (c.f. Fig. 1), we have

$$\int_{\Delta}^{\infty} \frac{e^{-u^2/2\Gamma_1^2} du}{\sqrt{2\pi}\Gamma_1} = 0.17. \quad (63)$$

By solving this equation we get the critical $(\Gamma_1/\Delta)_c = 1.05$ and the corresponding impurity concentration $N_c = 1.15N_1$ at the IMT.

One can estimate the small width of this IMT by calculating the hopping conductivity at the vicinity of IMT where $(\Gamma_1/\Delta)_c - \Gamma_1/\Delta \ll 1$ and finding the distance from the IMT where $[(\Gamma_1/\Delta)_c - \Gamma_1/\Delta]$ ceases to be exponentially small. This estimate would remind the calculation of the width of integer QHE steps [80]. Such a theory is beyond the scope of this paper.

So far in this section we dealt with completely compensated semiconductor. Let us now briefly consider strongly compensated n -type semiconductor in which concentrations of donors N_D and acceptors N_A are close, but different. It is convenient to describe such a semiconductor by the total concentration of Coulomb impurities $N = N_D + N_A$ and three dimensional concentration of electrons $n_3 = N_D - N_A \ll N$. We are interested in the phase diagram of IMT in the n_3 - N plane, which is shown in Fig. 9.

At modest impurity concentration $1 < Na_B^3 < N_1 a_B^3 = \alpha^{-9/2}$, we equate the Fermi energy with the Coulomb potential energy fluctuations to find the critical concentration n_{3c}

$$\frac{\hbar^2 n_3^{2/3}}{m} = \frac{e^2 \sqrt{N r_s^3}}{\kappa r_s}, \quad (64)$$

where $r_s = a_B^{1/2} n_3^{-1/6}$ is the screening length for nonrelativistic 3D electron gas. By solving Eq. (64) one obtains [2, 81]

$$n_3 = n_{3c} = N^{2/3}/a_B, \quad (65)$$

which is shown by the small N side phase boundary in Fig. 9. Near $N = N_1$ the phase boundary drops almost vertically to $n_3 = 0$ at $N = N_c = 1.15N_1$. Thus, as in 2D case we arrive at disorder driven reentrant metal-insulator-metal transition. Note that for 3D case, we

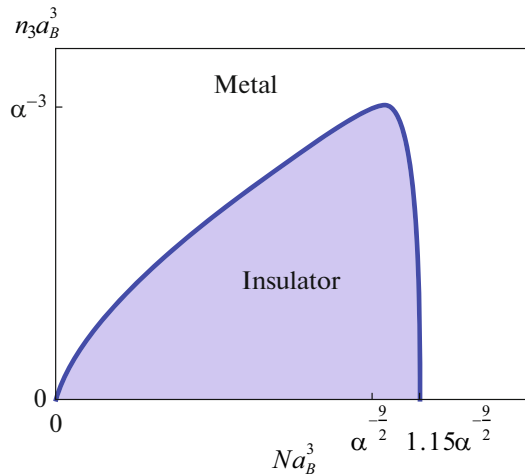


Fig. 9. Schematic n_3 – N plane phase diagram of a 3D narrow gap strongly compensated semiconductor. The shaded blue domain is the insulator phase while the white domain is the metal phase. The phase boundary follows Eq. (65) on the left side, reaches the maximum $n_3 a_B^3 = \alpha^{-3}$ near $N a_B^3 = N_1 a_B^3 = \alpha^{-9/2}$, and then vertically drops at $N = N_c = 1.15 N_1$.

use the word “metal” as opposite to 2D term “almost metal” we used in Fig. 7. This is because real metallic phase is allowed only in 3D.

9. SUMMARY AND CONCLUSIONS

In this paper we have considered the temperature-dependent conductivity of a two-dimensional insulator subjected to disorder by Coulomb impurities in the substrate. Our primary results can be summarized as follows. When the impurity concentration N is below a certain value N_0 (see Eq. (6)), the random potential of charged impurities necessarily produces large band bending, which the amplitude Δ becomes much larger than Δ . Then the system can be described as a network of large and closely-spaced fractal puddles (Fig. 2) separated by narrow insulating barriers (Fig. 5). This disorder landscape implies low-energy pathways for electron conduction, leads to the “three-mechanism sequence” illustrated in Fig. 3. The high temperature regime with $E_a = \Delta$ is relegated to only such high temperatures that T is comparable to Δ . The second regime, the nearest neighbor hopping between puddles (NNH), exhibits a parametrically smaller activation energy, whose value depends on the impurity concentration. At the lowest temperatures the conductivity is due to the Efros–Shklovskii variable range hopping (VRH), which may appear as an even smaller activation energy when measured over a limited temperature range. Experiments are instead more likely to observe NNH or ES VRH, with an activation energy that declines very rapidly with increasing N (Fig. 4).

When the impurity concentration N exceeds another critical value N_c the tunnel barriers between puddles become thin enough to be nearly transparent, and electrons are delocalized across many puddles. In this limit the conductivity follows ES law with the localization length growing exponentially with increased disorder. The corresponding apparent activation energy falls exponentially, so that in mesoscopic samples one effectively has an unconventional disorder-induced insulator-to-metal transition. The analogous problem for three-dimensional insulators (see Section 8) shows a genuine IMT due to percolation of electron and hole puddles separately.

Above we were talking about the neutrality point. When the Fermi level is away from neutrality point and the concentration of impurities is relatively small, there is a conventional metal-insulator transition with increasing disorder. Combining it with insulator–metal transition at large impurity concentrations away from neutrality we arrive at a disorder driven re-entrant metal–insulator–metal transition. (See phase diagrams of such transitions shown in Figs. 7 and 9.)

Our results have implications for a wide variety of experiments on 2D electron systems with a narrow energy gap. Some of these include 2D and thin 3D TIs, Bernal bilayer graphene with a perpendicular displacement field, and twisted bilayer graphene, as mentioned in the Introduction. In such systems the temperature-dependent conductivity is often used as a primary way to diagnose the magnitude of energy gaps. Our results here suggest that such studies suffer an essentially unavoidable limitation, since the apparent activation energy E_a at low temperature has no simple relation to the energy gap, and in general E_a can be taken only as a weak lower bound. No wonder that the transport activation energy in many cases is 10–100 times smaller than the value expected theoretically or measured by probes like optical absorption or tunneling spectroscopy. In this paper we studied in details gapped thin films of 3D topological insulators, which due to the large dielectric constant have peculiar 3D-like electrostatics (see Section 7).

The existence of an apparent disorder-induced IMT in strongly compensated semiconductor is an especially striking result of our analysis. For conventional insulators, this apparent transition cannot be called a true IMT, since in 2D the zero-temperature conductance flows toward zero in the thermodynamic limit for any finite amount of disorder [82]. However, the situation may be different for thin TI films, since the spin–orbit coupling of the TI surface states permits a stable metallic phase [83, 84]. A full theory of this IMT in TI films is beyond the scope of our current analysis.

Above we discussed how conductivity changes with the growth of the concentration of impurities N . However, in a typical experiment N is not well known and remains fixed, while Δ is tuned. For example, in BLG

this tuning is done by changing orthogonal displacement field D . Then our theory implies that measured at high temperature activation energy E_a agrees with the value of the gap $\Delta(D) \propto D$ predicted by [60, 85], if roughly $\Delta(D) > \Gamma = \alpha^{-2/3} e^2 \kappa^{-1} N^{1/3}$. On the other hand, if $\Delta(D) < \Gamma$ then the observed activation energy E_a is much smaller than the gap $\Delta(D)$. Thus, measuring minimum value of E_a coinciding with theoretical $\Delta(D)$ one can find N . Let us do this for the studied in [86] sample of BLG separated by thin hexagonal boron nitride (hBN) layers from silicon oxide, assuming that the majority of charged impurities are located in the bulk of SiO_2 . The smallest observed $E_a(D) = 11$ meV still resides on the theoretical line $\Delta(D)$ (c.f. Fig. 3c in [86]). This means that $N \leq N_0(E_a)$. Using the Dirac velocity for graphene $v \sim 10^6$ m/s for bilayer graphene layers with SiO_2 $\kappa = 4$, we get $\alpha \approx 0.5$. For $E_a(D) = 11$ meV we get $N \leq 10^{16} \text{ cm}^{-3}$.

Let us now estimate the concentration of charged impurities for BLG samples sandwiched between two hBN layers and deposited on the graphite gate [86] (they are likely C atoms [87] substituting for B and N). For this sample E_a agrees with theoretical $\Delta(D)$ till $E_a = 0.5$ meV. This means that in hBN N is so small that impurities are distributed in the layer of thickness d much smaller than $N^{-1/3}$. We generalized our theory to this case and arrived at the estimate $N \leq 2 \times 10^{14} \text{ cm}^{-3}$ for the state-of-the-art hBN. Indeed, for impurities located in a thin hBN layer of thickness $d \ll N^{-1/3} < \lambda$ the potential mean-square fluctuation

$$\Gamma^2 = (e^4 N d / \kappa^2) \ln(\lambda \sqrt{N d}).$$

Combining with Eq. (19), we get

$$\lambda = \alpha^{-12/7} (N d)^{-1/2}, \quad \Gamma = \alpha^{-2/7} e^2 \sqrt{N d} / \kappa.$$

Using $E_a \approx 0.5$ meV and $d \approx 20$ nm, we get that in hBN $N \leq N_0(E_a) = \alpha^{4/7} \kappa^2 E_a^2 / e^4 d \approx 2 \times 10^{14} \text{ cm}^{-3}$.

ACKNOWLEDGMENTS

We are grateful to David Goldhaber-Gordon, Ilya Gruzberg, Shahal Ilani, Kin Fai Mak, Koji Muraki, Stevan Nadj-Perge and Christoph Stampfer for helpful discussions.

FUNDING

Y.H. gratefully acknowledges support from Larkin Fellowship at the University of Minnesota. B.S. was partly supported by NSF grant DMR-2045742.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

ADDITIONAL INFORMATION

This article was prepared for the special issue of *Journal of Experimental and Theoretical Physics* dedicated to the 95th birthday of Professor E.I. Rashba.

REFERENCES

1. B. I. Shklovskii and A. L. Efros, *Sov. Phys. JETP* **35**, 610 (1972).
2. B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors*, Vol. 45 of *Springer Series in Solid-State Sciences* (Springer, Berlin, 1984).
3. B. Skinner, T. Chen, and B. I. Shklovskii, *Phys. Rev. Lett.* **109**, 176801 (2012).
4. T. Chen and B. Skinner, *Phys. Rev. B* **94**, 085146 (2016).
5. E. B. Olshanetsky, Z. D. Kvon, G. M. Gusev, A. D. Levin, O. E. Raichev, N. N. Mikhailov, and S. A. Dvoretzky, *Phys. Rev. Lett.* **114**, 126802 (2015).
6. Z. D. Kvon, D. A. Kozlov, E. B. Olshanetsky, G. M. Gusev, N. N. Mikhailov, and S. A. Dvoretzky, *Phys. Usp.* **63**, 629 (2020).
7. L. Pan, X. Liu, Q. L. He, A. Stern, G. Yin, X. Che, Q. Shao, P. Zhang, P. Deng, C.-Y. Yang, B. Casas, E. S. Choi, J. Xia, X. Kou, and K. L. Wang, *Sci. Adv.* **6**, eaaz3595 (2020).
<https://www.science.org/doi/pdf/10.1126/sciadv.aaz3595>
8. D. Nandi, B. Skinner, G. H. Lee, K.-F. Huang, K. Shain, C.-Z. Chang, Y. Ou, S.-P. Lee, J. Ward, J. S. Moodera, P. Kim, B. I. Halperin, and A. Yacoby, *Phys. Rev. B* **98**, 214203 (2018).
9. S. K. Chong, L. Liu, K. Watanabe, T. Taniguchi, T. Sparks, F. Liu, and V. V. Deshpande, *Nat. Portfolio* (2021).
<https://doi.org/10.21203/rs.3.rs-519444/v1>
10. J. G. Checkelsky, J. Ye, Y. Onose, Y. Iwasa, and Y. Tokura, *Nat. Phys.* **8**, 729 (2012).
11. C.-Z. Chang, J. Zhang, X. Feng, J. Shen, Z. Zhang, M. Guo, K. Li, Y. Ou, P. Wei, L.-L. Wang, Z.-Q. Ji, Y. Feng, S. Ji, X. Chen, J. Jia, X. Dai, Z. Fang, S.-C. Zhang, K. He, Y. Wang, L. Lu, X.-C. Ma, and Q.-K. Xue, *Science* (Washington, DC, U. S.) **340**, 167 (2013).
12. K. He, X.-C. Ma, X. Chen, L. Lü, Y.-Y. Wang, and Q.-K. Xue, *Chin. Phys. B* **22**, 067305 (2013).
13. M. Mogi, R. Yoshimi, A. Tsukazaki, K. Yasuda, Y. Kozuka, K. S. Takahashi, M. Kawasaki, and Y. Tokura, *Appl. Phys. Lett.* **107**, 182401 (2015).
<https://doi.org/10.1063/1.4935075>
14. L. Zhang, D. Zhao, Y. Zang, Y. Yuan, G. Jiang, M. Liao, D. Zhang, K. He, X. Ma, and Q. Xue, *APL Mater.* **5**, 076106 (2017).
<https://doi.org/10.1063/1.4990548>
15. W. Wang, Y. Ou, C. Liu, Y. Wang, K. He, Q.-K. Xue, and W. Wu, *Nat. Phys.* **14**, 791 (2018).
16. E. J. Fox, I. T. Rosen, Y. Yang, G. R. Jones, R. E. Elmquist, X. Kou, L. Pan, K. L. Wang, and D. Goldhaber-Gordon, *Phys. Rev. B* **98**, 075145 (2018).
17. J. Moon, J. Kim, N. Koirala, M. Salehi, D. Vanderbilt, and S. Oh, *Nano Lett.* **19**, 3409 (2019).
<https://doi.org/10.1021/acs.nanolett.8b03745>

18. I. T. Rosen, I. Yudhistira, G. Sharma, M. Salehi, M. A. Kastner, S. Oh, S. Adam, and D. Goldhaber-Gordon, *Phys. Rev. B* **99**, 201101 (2019).
19. Y. Okazaki, T. Oe, M. Kawamura, R. Yoshimi, S. Nakamura, S. Takada, M. Mogi, K. S. Takahashi, A. Tsukazaki, M. Kawasaki, Y. Tokura, and N.-H. Kaneko, *Appl. Phys. Lett.* **116**, 143101 (2020). <https://doi.org/10.1063/1.5145172>
20. L. K. Rodenbach, I. T. Rosen, E. J. Fox, P. Zhang, L. Pan, K. L. Wang, M. A. Kastner, and D. Goldhaber-Gordon, *APL Mater.* **9**, 081116 (2021). <https://doi.org/10.1063/5.0056796>
21. K. M. Fijalkowski, N. Liu, P. Mandal, S. Schreyeck, K. Brunner, C. Gould, and L. W. Molenkamp, *Nat. Commun.* **12**, 1 (2021).
22. G. M. Ferguson, R. Xiao, A. R. Richardella, D. Low, N. Samarth, and K. C. Nowack, *arXiv: 2112.13122 [cond-mat.mes-hall]* (2021).
23. I. T. Rosen, M. P. Andersen, L. K. Rodenbach, L. Tai, P. Zhang, K. L. Wang, M. A. Kastner, and D. Goldhaber-Gordon, *arXiv: 2112.13123 [cond-mat.mes-hall]* (2021).
24. Y. Okazaki, T. Oe, M. Kawamura, R. Yoshimi, S. Nakamura, S. Takada, M. Mogi, K. S. Takahashi, A. Tsukazaki, M. Kawasaki, Y. Tokura, and N.-H. Kaneko, *Nat. Phys.* **18**, 25 (2022).
25. K. Zou and J. Zhu, *Phys. Rev. B* **82**, 081407 (2010).
26. T. Taychatanapat and P. Jarillo-Herrero, *Phys. Rev. Lett.* **105**, 166601 (2010).
27. M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, *Science (Washington, DC, U. S.)* **367**, 900 (2020).
28. P. Stepanov, I. Das, X. Lu, A. Fahimniya, K. Watanabe, T. Taniguchi, F. H. L. Koppens, J. Lischner, L. Levitov, and D. K. Efetov, *Nature (London, U.K.)* **583**, 375 (2020).
29. J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, *Nature (London, U.K.)* **592**, 43 (2021).
30. Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, *Nature (London, U.K.)* **556**, 80 (2018).
31. Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Nature (London, U.K.)* **556**, 43 (2018).
32. Y. Xie, B. Lian, B. Jäck, X. Liu, C.-L. Chiu, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, *Nature (London, U.K.)* **572**, 101 (2019).
33. Y. Choi, H. Kim, C. Lewandowski, Y. Peng, A. Thomson, R. Polski, Y. Zhang, K. Watanabe, T. Taniguchi, J. Alicea, and S. Nadj-Perge, *Nat. Phys.* **17**, 1375 (2021).
34. R. Yu, W. Zhang, H.-J. Zhang, S.-C. Zhang, X. Dai, and Z. Fang, *Science (Washington, DC, U. S.)* **329**, 61 (2010).
35. J. Zhang, C.-Z. Chang, P. Tang, Z. Zhang, X. Feng, K. Li, L. Li Wang, X. Chen, C. Liu, W. Duan, K. He, Q.-K. Xue, X. Ma, and Y. Wang, *Science (Washington, DC, U. S.)* **339**, 1582 (2013). <https://www.science.org/doi/pdf/10.1126/science.1230905>
36. C.-Z. Chang, C.-X. Liu, and A. H. MacDonald, *arXiv: 2202.13902 [condmat.mes-hall]* (2022).
37. C.-X. Liu, H. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, *Phys. Rev. B* **81**, 041307(R) (2010).
38. H.-Z. Lu, W.-Y. Shan, W. Yao, Q. Niu, and S.-Q. Shen, *Phys. Rev. B* **81**, 115407 (2010).
39. J. Linder, T. Yokoyama, and A. Sudbo, *Phys. Rev. B* **80**, 205401 (2009).
40. Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, X. Dai, W.-Y. Shan, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, *Nat. Phys.* **6**, 584 (2010).
41. Y. Sakamoto, T. Hirahara, H. Miyazaki, S.-I. Kimura, and S. Hasegawa, *Phys. Rev. B* **81**, 165432 (2010).
42. T. Zhang, J. Ha, N. Levy, Y. Kuk, and J. Strosio, *Phys. Rev. Lett.* **111**, 056803 (2013).
43. D. Kim, P. Syers, N. P. Butch, J. Paglione, and M. S. Fuhrer, *Nat. Commun.* **4**, 2040 (2013).
44. Y. L. Chen, J.-H. Chu, J. G. Analytis, Z. K. Liu, K. Igarashi, H.-H. Kuo, X. L. Qi, S. K. Mo, R. G. Moore, D. H. Lu, M. Hashimoto, T. Sasagawa, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z. X. Shen, *Science (Washington, DC, U. S.)* **329**, 659 (2010).
45. S.-Y. Xu, M. Neupane, C. Liu, D. Zhang, A. Richardella, L. Andrew Wray, N. Alidoust, M. Leandersson, T. Balasubramanian, J. Sánchez-Barriga, O. Rader, G. Landolt, B. Slomski, J. Hugo Dil, J. Osterwalder, T.-R. Chang, H.-T. Jeng, H. Lin, A. Bansil, N. Samarth, and M. Zahid Hasan, *Nat. Phys.* **8**, 616 (2012).
46. M. Ye, W. Li, S. Zhu, Y. Takeda, Y. Saitoh, J. Wang, H. Pan, M. Nurmamat, K. Sumida, F. Ji, Z. Liu, H. Yang, Z. Liu, D. Shen, A. Kimura, S. Qiao, and X. Xie, *Nat. Commun.* **6**, 8913 (2015).
47. Y. Tokura, K. Yasuda, and A. Tsukazaki, *Nat. Rev. Phys.* **1**, 126 (2019).
48. Y. Deng, Y. Yu, M. Z. Shi, Z. Guo, Z. Xu, J. Wang, X. H. Chen, and Y. Zhang, *Science (Washington, DC, U. S.)* **367**, 895 (2020).
49. R. Lu, H. Sun, S. Kumar, Y. Wang, M. Gu, M. Zeng, Y.-J. Hao, J. Li, J. Shao, X.-M. Ma, Z. Hao, K. Zhang, W. Mansuer, J. Mei, Y. Zhao, et al., *Phys. Rev. X* **11**, 011039 (2021).
50. Y. Huang, Y. He, B. Skinner, and B. I. Shklovskii, *Phys. Rev. B* **105**, 054206 (2022).
51. T. Chen, B. Skinner, and B. I. Shklovskii, *Phys. Rev. Lett.* **109**, 126805 (2012).
52. J. Zhang and B. I. Shklovskii, *Phys. Rev. B* **70**, 115317 (2004).
53. M. B. Isichenko, *Rev. Mod. Phys.* **64**, 961 (1992).
54. B. Skinner and B. I. Shklovskii, *Phys. Rev. B* **87**, 075454 (2013).
55. B. I. Shklovskii and A. L. Efros, *Sov. Phys. JETP* **32**, 733 (1971).
56. M. E. Raikh and I. M. Ruzin, *Sov. Phys. Semicond.* **19**, 745 (1985).

57. H. Samavati, A. Hajimiri, A. Shahani, G. Nasserbakht, and T. Lee, *IEEE J. Solid-State Circuits* **33**, 2035 (1998).
58. A. L. Efros and B. I. Shklovskii, *J. Phys. C: Solid State Phys.* **8**, L49 (1975).
59. B. I. Shklovskii, *Low Temp. Phys.* **43**, 699 (2017).
60. E. McCann and M. Koshino, *Rep. Prog. Phys.* **76**, 056503 (2013).
61. B. I. Shklovskii, *Sov. Phys. Semicond.* **7**, 77 (1973).
62. Y. Huang, Y. Ayino, and B. I. Shklovskii, *Phys. Rev. Mater.* **5**, 044606 (2021).
63. T. Li, S. Jiang, L. Li, Y. Zhang, K. Kang, J. Zhu, K. Watanabe, T. Taniguchi, D. Chowdhury, L. Fu, J. Shan, and K. F. Mak, *Nature (London, U.K.)* **597**, 350 (2021).
64. A. Ghiotto, E.-M. Shih, G. S. S. G. Pereira, D. A. Rhodes, B. Kim, J. Zang, A. J. Millis, K. Watanabe, T. Taniguchi, J. C. Hone, L. Wang, C. R. Dean, and A. N. Pasupathy, *Nature (London, U.K.)* **597**, 345 (2021).
65. W. Richter and C. R. Becker, *Phys. Status Solidi B* **84**, 619 (1977).
66. N. Borgwardt, J. Lux, I. Vergara, Z. Wang, A. A. Taskin, K. Segawa, P. H. M. van Loosdrecht, Y. Ando, A. Rosch, and M. Grüninger, *Phys. Rev. B* **93**, 245149 (2016).
67. T. Bömerich, J. Lux, Q. T. Feng, and A. Rosch, *Phys. Rev. B* **96**, 075204 (2017).
68. J. Zhang, C.-Z. Chang, Z. Zhang, J. Wen, X. Feng, K. Li, M. Liu, K. He, L. Wang, X. Chen, Q.-K. Xue, X. Ma, and Y. Wang, *Nat. Commun.* **2**, 574 (2011).
69. N. S. Rytova, *Moscow Univ. Phys. Bull.* **3**, 18 (1967).
70. A. Chaplik and M. Entin, *J. Exp. Theor. Phys.* **34**, 1335 (1971).
71. L. V. Keldysh, *JETP Lett.* **29**, 658 (1979).
72. Y. Huang and B. I. Shklovskii, *Phys. Rev. B* **103**, 165409 (2021).
73. Y. Huang and B. I. Shklovskii, *Phys. Rev. B* **104**, 054205 (2021).
74. H. Weng, X. Dai, and Z. Fang, *Phys. Rev. X* **4**, 011002 (2014).
75. M. K. Tran, J. Levallois, P. Lerch, J. Teyssier, A. B. Kuzmenko, G. Autès, O. V. Yazyev, A. Ubaldini, E. Giannini, D. van der Marel, and A. Akrap, *Phys. Rev. Lett.* **112**, 047402 (2014).
76. A. Ohmura, Y. Higuchi, T. Ochiai, M. Kanou, F. Ishikawa, S. Nakano, A. Nakayama, Y. Yamada, and T. Sasagawa, *Phys. Rev. B* **95**, 125203 (2017).
77. S. Singh, A. C. Garcia-Castro, I. Valencia-Jaime, F. Muñoz, and A. H. Romero, *Phys. Rev. B* **94**, 161116 (2016).
78. D. Vu, W. Zhang, C. Şahin, M. E. Flatté, N. Trivedi, and J. P. Heremans, *Nat. Mater.* **20**, 1525 (2021).
79. B. Skinner, *Phys. Rev. B* **90**, 060202 (2014).
80. D. G. Polyakov and B. I. Shklovskii, *Phys. Rev. B* **48**, 11167 (1993).
81. B. I. Shklovskii and A. L. Efros, *Sov. Phys. JETP* **33**, 468 (1971).
82. E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, *Phys. Rev. Lett.* **42**, 673 (1979).
83. S. Hikami, A. I. Larkin, and Y. Nagaoka, *Prog. Theor. Phys.* **63**, 707 (1980).
<https://academic.oup.com/ptp/articlepdf/63/2/707/5336056/63-2-707.pdf>.
84. R. S. K. Mong, J. H. Bardarson, and J. E. Moore, *Phys. Rev. Lett.* **108**, 076804 (2012).
85. S. Slizovskiy, A. Garcia-Ruiz, A. I. Berdyugin, N. Xin, T. Taniguchi, K. Watanabe, A. K. Geim, N. D. Drummond, and V. I. Fal'ko, *Nano Lett.* **21**, 6678 (2021).
<https://doi.org/10.1021/acs.nanolett.1c02211>
86. E. Icking, L. Banszerus, F. Wörtche, F. Volmer, P. Schmidt, C. Steiner, S. Engels, J. Hesselmann, M. Goldsche, K. Watanabe, T. Taniguchi, C. Volk, B. Beschoten, and C. Stampfer, *arXiv: 2206.02057 [cond-mat.mes-hall, cond-mat.mtrlsci]* (2022).
87. M. Onodera, M. Isayama, T. Taniguchi, K. Watanabe, S. Masubuchi, R. Moriya, T. Haga, Y. Fujimoto, S. Saito, and T. Machida, *Carbon* **167**, 785 (2020).