



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Conductivity of two-dimensional narrow gap semiconductors subjected to strong Coulomb disorder

Yi Huang (✉), Yanjun He, Brian Skinner, and B. I. Shklovskii

Phys. Rev. B **105**, 054206 — Published 22 February 2022

DOI: [10.1103/PhysRevB.105.054206](https://doi.org/10.1103/PhysRevB.105.054206)

Conductivity of two-dimensional narrow gap semiconductors subjected to strong Coulomb disorder

Yi Huang (黄奕),¹ Yanjun He,² Brian Skinner,² and B. I. Shklovskii¹

¹*School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455, USA**

²*Department of Physics, The Ohio State University, Columbus, Ohio 43202, USA*

(Dated: February 8, 2022)

In the ideal disorder-free situation, a two-dimensional band gap insulator has an activation energy for conductivity equal to half the band gap 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 insulator on a substrate containing Coulomb impurities, with random potential amplitude $\Gamma \gg \Delta$. We show that the conductivity generically exhibits three regimes of conductivity, and only the highest temperature regime exhibits an activation energy that reflects the band gap. At lower temperatures, the conduction proceeds through activated hopping or Efros-Shklovskii 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 also arrive at a similar steep disorder driven insulator-metal transition in thin films of three-dimensional topological insulators with large dielectric constant, for which Coulomb impurities inside the film create a large disorder potential due to confinement of their electric field inside the film.

I. 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 Γ 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 puddles [1–3]. At high enough temperatures the electrical conductivity is due to activation of electrons and holes from the chemical potential to the energy associated with classical percolation across the sample. At lower temperatures the conductivity is due to hopping between nearest-neighbor puddles. At even smaller temperatures it is due to variable-range hopping 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 non-universal and disorder-dependent [1, 2].

In this paper we consider a similar problem in two dimensions, focusing on the case of strong disorder, $\Gamma \gg \Delta$. Specifically, we consider a two-dimensional small band gap semiconductor resting on a thick substrate with a three-dimensional

concentration of randomly-positioned impurities. We derive the temperature dependence of the electrical conductivity across all temperature regimes and show that observed activation energy 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 2D electron systems. For example, recent studies of 2D topological insulators (TIs) [5–7], films of 3D TIs [8–20], bilayer graphene (BLG) with an orthogonal electric field [21, 22] and twisted bilayer graphene (TBG) [23–27] 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 parametrically smaller than the energy gap. At the lowest temperatures the conductivity follows a Efros-Shklovskii (ES) law [28] 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 [23–27]. 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., Refs. 24 and 25), 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 [29, 30]. The theory

* Corresponding author: huan1756@umn.edu

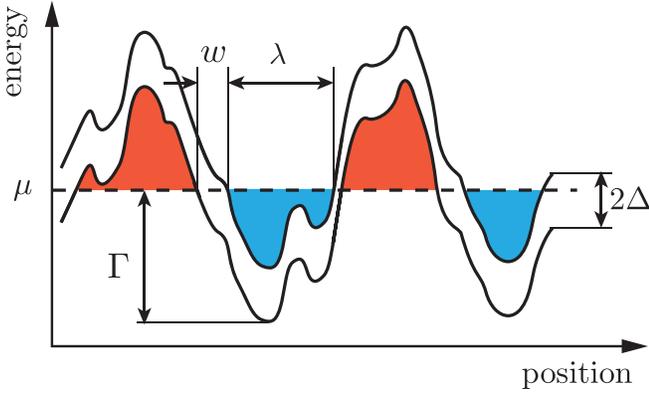


Figure 1. Schematic picture of a cross section of puddles in the case $\Gamma \gg \Delta$. The wavy lines show the conduction band bottom and the valence band ceiling separated by the gap 2Δ . The red shaded region above the chemical potential $\mu = 0$ represents a hole puddle, while the blue shaded region below μ represents an electron puddle, Γ is the amplitude of the disorder potential, λ is the screening length, and w is the width of the barrier between neighboring puddles.

we present here offers a natural way to interpret this discrepancy.

Second, our theory can be applied to the huge body of experimental work on thin films of a 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 [10–20]. 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, 20, 31, 32] and the quantum spin Hall effect [9, 33–35].

The model we consider of 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

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

where ϵ is the electron energy, \mathbf{k} is the 2D wave vector, v is the Dirac velocity, and \hbar is the reduced Planck constant. We are 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 space (see Fig. 1). This system is an insulator because in 2D neither electron nor hole puddles percolate, and they are disconnected from each other (neglecting, for now, the possibility of quantum tunneling between puddles). Throughout this paper we 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 Refs. 5–27, 29, and 30.

The remainder of this paper is organized as follows. In the following section we first summarize our main results for the temperature-dependent conductivity. In Sec. III we review

the fractal geometry of two-dimensional puddles for the case $\Gamma \gg \Delta$. In Sec. IV we calculate the action accumulated by electrons tunneling across the gap between two neighboring puddles, the corresponding localization length, and the critical value $(\Gamma/\Delta)_c$ of the ratio Γ/Δ , at which crossover to “almost metallic conductivity” takes place. In Sec. V we calculate the hopping conductivity at $1 \ll \Gamma/\Delta \ll (\Gamma/\Delta)_c$. Section VI deals with the generalization of our results to thin TI films. Because of the intense recent interest in such films [8–10, 12–20, 31–45] in this section we add a fair amount of numerical estimates. We close in Sec. VII with a summary and conclusion.

II. SUMMARY OF RESULTS

In situations where 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. 2. At relatively large temperature T electrons and holes can be activated from the chemical potential to the percolation level (i.e., the classical mobility edge). As we show in detail in Appendix A, the activation energy for this process is exactly equal to Δ when the chemical potential $\mu = 0$ (i.e., at the charge neutrality point). Thus, the conductivity at such large temperatures is given 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 hopping of electrons between electron and hole puddles near the chemical potential. Similarly to the case of granular metals [46, 47], the activation energy of such hopping is first determined by the typical puddle charging energy E_C

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

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

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

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 $v \sim 10^6 \text{ m s}^{-1}$; and with $\kappa = 4$ for SiO_2 , 11 for insulating GaAs, 20 for HfO_2 and 1000 for PbTe; α can vary from 1 to 10^{-3} . Below in our theory we use α as a small parameter, $\alpha \ll 1$.

At even lower temperatures the activated hopping (AH) crosses over to the Efros-Shklovskii (ES) law

$$\sigma = \sigma_3 \exp\left[-(T_{\text{ES}}/T)^{1/2}\right], \quad (T \ll T_2), \quad (5)$$

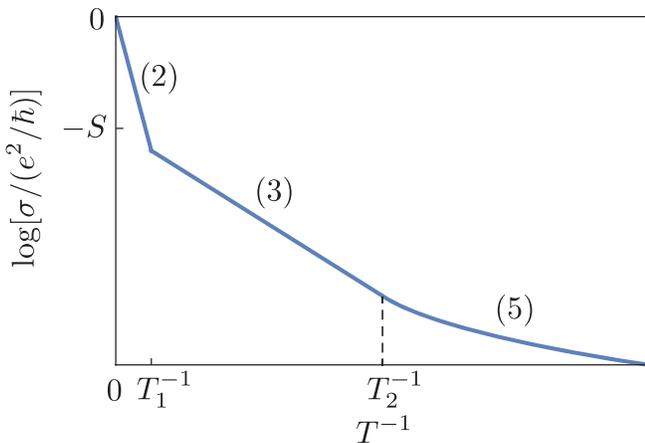


Figure 2. Schematic plot of the logarithm of the dimensionless conductivity $\sigma/(e^2/\hbar)$ as a function of the inverse temperature T^{-1} . At high temperature $T > T_1 = \Delta/S$, the conductivity has activation energy Δ . At intermediate temperature $T_2 < T < T_1$ where $T_2 = E_C^2/T_{ES}$, the conductivity is dominated by activated hopping (AH). At low temperature such that $T < T_2$, AH is replaced by ES law. Numbers adjacent to different parts of the line show corresponding equations.

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

$$T_{ES} = \alpha\Delta(\Delta/\Gamma)^{37/9} \ll \Delta, \quad (6)$$

and the temperatures associated with the crossover between the different regimes are

$$T_1 = \alpha\Delta(\Gamma/\Delta)^{34/9}, \quad (7)$$

$$T_2 = \alpha^3\Delta(\Gamma/\Delta)^{13/9}. \quad (8)$$

Above we assumed that metallic gates are far enough from the semiconductor so that there is no screening of electron-hole Coulomb interaction leading to a crossover between ES and Mott law of variable range hopping. Such crossover happens when the hop length becomes larger than the distance to the gate [48, 49].

Similar results for all three mechanisms of conductivity are obtained below for thin films of 3D topological insulators. However, such a three-mechanism sequence is not observed in most experiments [5–27, 29, 30]. 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 that at $\Gamma/\Delta > (\Gamma/\Delta)_c$, electrons are not localized in single puddles and the first two regimes of conductivity are absent. For presented above general case $(\Gamma/\Delta)_c = \alpha^{-9/41}$. The only remaining mechanism is the ES law with very small T_{ES} . This means that the low temperature “local activation energy” is much smaller than Δ . Such a theory predicts that the prefactor should be close to e^2/\hbar . This prediction agrees with some experiments [16, 19], but disagrees with others [18, 20].

III. FRACTAL GEOMETRY OF PUDDLES

In this section we briefly review some geometrical fractal properties of 2D puddles at $\Gamma \gg \Delta$ [50]. The characteristic size (diameter) of a puddle is given by

$$a = \lambda(\Gamma/\Delta)^\nu, \quad (9)$$

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

$$L = a(\Gamma/\Delta) = \lambda(\Gamma/\Delta)^{\nu+1}. \quad (10)$$

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. 3]. The area of a puddle is given by

$$A = \lambda^2(\Gamma/\Delta)^{2\nu-\beta}, \quad (11)$$

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

$$w = \lambda\Delta/\Gamma. \quad (12)$$

In order to estimate Γ and λ we can use the self-consistent theory of Ref. 51, which dealt with the disorder potential at the surface of a bulk TI created by charged impurities with 3D 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 for determining the values of Γ and λ 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 :

$$\Gamma = \frac{e^2}{\kappa\lambda}(N\lambda^3)^{1/2}. \quad (13)$$

This relation leads to a typical 2D density of states

$$g = \kappa^2\alpha^2\Gamma/e^4, \quad (14)$$

which in turn leads to the screening radius

$$\lambda = \frac{\kappa}{e^2g} = \frac{e^2}{\alpha^2\kappa\Gamma}. \quad (15)$$

Solving Eqs. (13) and (15) for Γ and λ we get [51]

$$\Gamma = \frac{e^2N^{1/3}}{\kappa\alpha^{2/3}}, \quad (16)$$

$$\lambda = \alpha^{-4/3}N^{-1/3}. \quad (17)$$

IV. TUNNELING ACTION, LOCALIZATION LENGTH AND CONDUCTANCE

Let us now calculate the hopping conductivity of the system of fractal metallic puddles separated by narrow insulating gaps shown in Fig. 1. In this section we estimate the dimensionless tunneling action S . The value of S is determined by

the tunneling length $r = \Delta/eE$ in the spatially-varying electric field E created by impurities:

$$S = \frac{r\Delta}{\hbar v} = \frac{\Delta^2}{eE\hbar v}. \quad (18)$$

It is tempting to use $\Gamma/e\lambda$ for E and arrive at $S = \alpha^{-1}(\Delta/\Gamma)^2$. 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 impurity 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 much smaller than the naive estimate quoted above. Our optimization procedure is a mesoscopic version of the optimization used in the theory of the interband absorption of light in compensated three-dimensional semiconductors [2, 52]. It is also similar to the theory of fluctuation-induced excess currents in reverse biased p - n junctions [53].

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-neighboring puddles. Thus,

$$(L/\lambda) \exp\left[-\frac{M^2}{Nr^3}\right] = 1. \quad (19)$$

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. (19) and solving for M gives

$$M = \frac{\alpha^{-2}(\Delta/\Gamma)^3}{\ln[(\Gamma/\Delta)^{7/3}]}, \quad (20)$$

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. (18) we have

$$S = \frac{\alpha^{-1}(\Delta/\Gamma)^3}{\ln[(\Gamma/\Delta)^{7/3}]} \simeq \alpha^{-1}(\Delta/\Gamma)^{34/9}. \quad (21)$$

In the last step we used the power-law approximation $\ln x \approx x^{1/3}$, which is 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. Consider a tunneling path with a displacement $x \gg a$. The action associated with this path is dominated by the action for tunneling across the narrow insulating gaps between puddles, which the electron must cross every time it displaces across one puddle diameter ξ . Consequently, the total action of the tunneling path is $Sx/a = x/\xi$ where

$$\xi = a/S = \alpha a(\Gamma/\Delta)^{34/9}. \quad (22)$$

Notice, however, that the fast decrease of S with growing Γ/Δ leads to a fast increase of the dimensionless conductance between two neighboring puddles

$$G = (L/\lambda) \exp(-S), \quad (23)$$

so that we get $G = 1$ at some critical value $(\Gamma/\Delta)_c$. Substituting Eq. (21) into Eq. (23) and setting $G = 1$, we arrive at an estimate for the critical disorder strength [54]

$$(\Gamma/\Delta)_c = \alpha^{-9/41}, \quad (24)$$

valid for $\alpha \in (1.2 \times 10^{-4}, 0.12)$. This range of α is obtained by substituting Eq. (24) into the requirement for the argument of the logarithm $(\Gamma/\Delta)^{7/3} \in (3, 100)$ [55].

At larger Γ/Δ the localization length grows exponentially as $\xi = ae^G$. This growth leads to a dramatic growth of the conductivity, which we dub an ‘‘insulator - almost metal transition’’ (IAMT), if the sample size is much larger than ξ . For a sufficiently small sample, $(\Gamma/\Delta)_c$ effectively plays the role of disorder-induced insulator-metal transition.

V. HOPPING CONDUCTIVITY

At moderate disorder when $1 < \Gamma/\Delta < (\Gamma/\Delta)_c$ electrons are well localized within a single puddle and the temperature dependence of the conductivity follows the three-mechanism sequence described above, for which with decreasing temperature the activated conductivity with activation energy Δ is replaced first by AH and then by the ES law. In this case the system is similar to a network of densely packed metallic granules separated by a thin insulating matrix with Coulomb impurities, and we can follow the calculation of their conductivity [46, 47].

Let us start from the discussion of AH conductivity at $\Delta/T \gg S$ or $T \ll T_1 = \Delta/S$. [One arrives at Eq. (7) for T_1 with the help of Eq. (21)]. In this case, the charging energy of a puddle E_C replaces Δ as the activation energy for conductivity. 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 the standard expression $E_C = e^2/\kappa a$. Namely we are going to show that

$$E_C = \frac{e^2\Delta}{\kappa a\Gamma}. \quad (25)$$

Substituting Eqs. (9) and (15) into Eq. (25) one arrives at Eq. (4). Let us illustrate how this happens by comparing the self-capacitance of an isolated puddle, $C_0 \sim \kappa a$, with the capacitance C of the same puddle surrounded by other puddles. In the latter case, because our puddle has metallic properties, an excess electron charge e spreads to the border (perimeter), while neighboring metallic puddles provide 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. 3.

This mechanism of enhanced capacitance was recognized by the electrical engineering community [56]. In our system it means that $C \sim \kappa L$, and $E_C = e^2/C$ leads to Eq. (25).

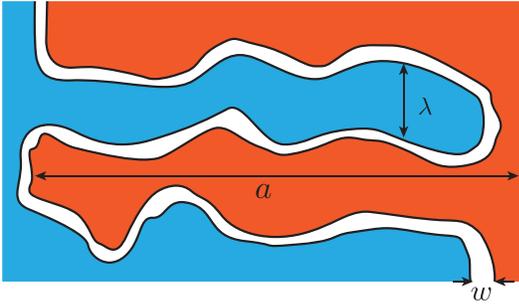


Figure 3. Schematic picture of interlocked “fingers” of neighboring puddles. Here the length of “fingers” a is of order of the puddle diameter. One can imagine that Fig. 1 shows a vertical cross section of Fig. 3.

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

$$\delta = (gA)^{-1} = \alpha^2 (\Delta/\Gamma)^{55/36} \Delta, \quad (26)$$

where g is the 2D DOS given by Eq. (14), and A is the area of a puddle given by Eq. (11). Therefore the ratio $\delta/E_C = (\Delta/\Gamma)^{7/36} \ll 1$ and our use of E_C is legitimate.

Let us now consider the ES conductivity which replaces AH at low temperatures. In the ground state, each puddle i of our system is charged by a random fractional charge $|q_i| \leq e/2$. This charging 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 a random potential that results in localized electron states and enables the formation of the Coulomb gap around the chemical potential [46, 47]. This Coulomb gap leads again to conductivity described by the ES law in the low temperature limit.

We can calculate the constant T_{ES} in the ES law starting from the standard expression $T_{ES} = e^2/\kappa\xi$ [2, 28]. Using Eq. (22) for ξ we arrive at Eq. (6). We see now that $T_{ES} \ll \Delta$. Equating $(T_{ES}/T)^{1/2}$ to E_C/T with the help of Eqs. (6), we arrive at the expression of T_2 as Eq. (8).

VI. THIN FILM OF 3D TOPOLOGICAL INSULATOR

In previous sections we dealt with a general model of a trivial 2D semiconductor with gapped Dirac spectrum Eq. (1). In this section we concentrate on the 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, 33–39] or may be created by a finite concentration of magnetic dopants like Cr [10–20, 40, 41]. Because of the promise of such films to achieve metrological precision of the quantum anomalous Hall effect [11, 13, 16, 19, 20, 31, 32] and the quantum spin Hall effect [9, 33–35], in this section

we are more specific with material parameters and numerical estimates.

We have in mind TI thin films based on $(\text{Bi}_x \text{Sb}_{1-x})_2 \text{Te}_3$ [8], or $(\text{Bi}_x \text{Sb}_{1-x})_2 (\text{Te}_y \text{Se}_{1-y})_3$ [9] which have very large dielectric constant $\kappa \sim 200$ [57–59]. Using $\kappa \sim 200$ and the Fermi velocity of TI $v \sim 4 \times 10^5$ m/s [60], one gets $\alpha \sim 0.027$. We assume that such a film of width $d \sim 7$ nm is deposited on a substrate with much smaller dielectric constant $\kappa_e \ll \kappa$, so that the electric field created by Coulomb impurities residing inside the film are trapped within the film [61–64]. This trapping slows down the decay of the Coulomb potential with distance and enhances the role of impurities. TI films typically have a large ($N \sim 10^{19}$ cm $^{-3}$) concentration of Coulomb impurities, which allows us to study only their effect and to assume that impurities inside the substrate play no role.

As in the previous section, we calculate the tunneling action S and the critical ratio $(\Gamma/\Delta)_c$, and we describe the hopping conductivity of the film. According to Ref. [64], due to the peculiar electrostatics of TI films, the expression of Γ is the same as Eq. (16), while the expression for λ becomes

$$\lambda = \alpha^{-2/3} (Nd^3)^{-1/6} d, \quad (27)$$

which is valid if $\lambda > d$. Using the above estimates of α , d and N , we get $\lambda \sim 50$ nm, so that this estimate is valid.

Notice that the electric field in the plane of the film created by charge fluctuations in a disk of radius r and thickness d is given by

$$E = \frac{e\sqrt{Nr^2d}}{\kappa rd} = \frac{e}{\kappa} \sqrt{\frac{N}{d}}, \quad (28)$$

which turns out to be independent of r . Therefore, there is no enhancement of the electric field at scales shorter than λ , as there was in Sec. IV. Substituting Eq. (28) (or, equivalently, $E = \Gamma/e\lambda$) into Eq. (18), we arrive at the action

$$S = \frac{w\Delta}{\hbar v} = \frac{\lambda\Delta^2}{\hbar v\Gamma} = \alpha^{-1/3} (Nd^3)^{1/6} (\Delta/\Gamma)^2 \quad (29)$$

for tunneling between neighboring puddles. 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. (19) by

$$(L/\lambda) \exp\left[-\frac{M^2}{N\lambda^2 d}\right] = 1. \quad (30)$$

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

$$M = \left\{ N\lambda^2 d \ln[(\Gamma/\Delta)^{7/3}] \right\}^{1/2}. \quad (31)$$

Now substituting the electric field $E = eM/\kappa\lambda d$ into the action Eq. (18) gives

$$S = \frac{\alpha^{-1/3} (Nd^3)^{1/6} (\Delta/\Gamma)^2}{\left\{ \ln[(\Gamma/\Delta)^{7/3}] \right\}^{1/2}} \simeq \alpha^{-1/3} (Nd^3)^{1/6} (\Delta/\Gamma)^{43/18}, \quad (32)$$

which is smaller than the action given by Eq. (29). In the last step, as in the Sec. IV we used the power-law approximation $\ln x = x^{1/3}$ valid for $x \in (3, 100)$ with accuracy better than 30%.

Substituting Eq. (32) into the expression of G , Eq. (23), and setting $G = 1$, we arrive at the critical point [65]

$$(\Gamma/\Delta)_c = \alpha^{-2/19}(Nd^3)^{1/19}. \quad (33)$$

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 the charging energy of a puddle. Similarly to Sec. III, capacitance of a puddle within the film is determined by long border of the puddle with neighboring puddles. Near the border there are two stripes of the length L and width λ with charges $-e$ and e . 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

$$E_C = (e^2/\kappa d)(\lambda/L) = (e^2/\kappa d)(\Delta/\Gamma)^{7/3}, \quad (34)$$

At lower temperatures $T < T_2$ conductivity obeys the ES law with the characteristic temperature $T_{\text{ES}} = e^2/\kappa_e \xi$, 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. (9) and (32) we get

$$T_{\text{ES}} = \alpha(\kappa/\kappa_e)\Delta(\Delta/\Gamma)^{49/18}. \quad (35)$$

Eqs. (35) and (33) show that in TI films, as in trivial semiconductors [c.f. Eqs. (6) and (24)], a reduction of T_{ES} and a crossover from strong localization to practically metallic conductivity happens dramatically quickly when Γ exceeds Δ .

We now estimate the characteristic energies Γ , Δ , E_C , T_{ES} , T_1 , and T_2 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 $\Gamma \simeq 17 \text{ meV}$. 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 $\Gamma \simeq 17 \text{ meV}$, $\Delta \simeq 15 \text{ meV}$, $S \simeq 3$, $E_C \simeq 0.8 \text{ meV}$, and $T_{\text{ES}} \simeq 130 \text{ K}$ (here assume that the film has BN on both sides and use $\kappa_e = 5$). Temperatures $T_1 = \Delta/S \simeq 60 \text{ K}$ and $T_2 = E_C^2/T_{\text{ES}} = 0.6 \text{ K}$ are obtained by equating Eqs. (2) to (3), and (3) to (5) respectively. In this case, apparently ES conductivity starts when $(T_{\text{ES}}/T_2)^{1/2} \sim 15$ so that ES law is hardly observable because of very large resistance. 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, one can show that in thinner films, $d < 7 \text{ nm}$, for which $\Delta > \Gamma$ activation energy $E_C = e^2/\kappa d$ and $T_{\text{ES}} = \alpha(\kappa/\kappa_e)\Delta$ 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 Refs. 8 and 9.

For the case of magnetically doped TI thin films, the exchange half-gap Δ induced by magnetic impurities is not directly related to d and is of order of 20 meV [19, 20, 45], so that we have practically the same numbers as in the previous example.

VII. SUMMARY AND CONCLUSION

In this paper we have considered the temperature-dependent conductivity of a two-dimensional insulator subjected to the random potential created by Coulomb impurities in the substrate. Our primary results can be summarized as follows. First, the random potential of charged impurities necessarily produces large band bending. We focus here on the case where the impurity concentration is large enough that the disorder potential $\Gamma \gg \Delta$, and the system can be described as a network of large and closely-spaced fractal puddles [Fig. 1] separated by narrow insulating barriers [Fig. 3]. This case is characterized by the ‘‘three-mechanism sequence’’ of temperature-dependent conductivity illustrated in Fig. 2. Only the highest-temperature regime has an activation energy E_a equal to half the band gap Δ . The middle regime, activated hopping between puddles (AH), exhibits a parametrically smaller activation energy whose value depends on the impurity concentration, while the lowest temperature regime corresponds to Efros-Shklovskii conductivity, which may appear as an even smaller activation energy when measured over a limited temperature range.

Second, when the impurity concentration N exceeds some critical value the tunnel barriers between puddles become thin enough to be nearly transparent, and electrons are delocalized across many puddles. In this limit the localization length grows exponentially with increased disorder, and the corresponding activation energy falls exponentially, so that in mesoscopic samples one effectively has a disorder-induced insulator-to-metal transition.

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 100 times smaller than the value expected theoretically or measured by probes like optical absorption or tunneling spectroscopy.

In principle, one can infer the band gap by measuring the activation energy at the highest temperature regime. However, the existence of this regime practically requires a low enough disorder that electron and hole puddles are small and well separated from each other. Even in this case, experiments using transport to estimate Δ should first demonstrate two distinct

regimes of constant activation energy, and then use only the value from the higher- T regime as an estimate of Δ .

The existence of an apparent disorder-induced IMT at $\Gamma/\Delta > (\Gamma/\Delta)_c$ 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 [66]. 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 [67, 68]. A full theory of this IMT in TI films is beyond the scope of our current analysis.

ACKNOWLEDGMENTS

We are grateful to Stevan Nadj-Perge, Koji Muraki, Shahal Ilani, Ilya Gruzberg, and David Goldhaber-Gordon for helpful conversations. Y.H. is supported by the William I. Fine Theoretical Physics Institute. B.S. was partly supported by NSF grant DMR-2045742.

Appendix A: Activation to the classical mobility edge

1. Percolation-based argument for $E_a = \Delta$

In this Appendix we consider the process of thermal activation of electrons from the chemical potential to the classical mobility edge. We show that, generically, this process leads to an activation energy

$$E_a = |E_{\text{be}} - \mu|, \quad (\text{A1})$$

where $E_{\text{be}} = \pm\Delta$ represents the mean energy of the nearest band edge (conduction band or valence band) and μ is the chemical potential, both defined relative to the mid-gap [69]. In the maximally-insulating state, where $\mu = 0$, this equation gives

$$E_a = \Delta, \quad (\text{A2})$$

as in the non-disordered system.

To show this result, we first consider the case $\mu = 0$. In this case the mean energy of the conduction band edge is $+\Delta$ and the mean energy of the valence band edge is $-\Delta$, although the disorder potential causes both band edges to wander in energy as a function of position. (See Fig. 1.) Let us now consider the process of drawing spatial contours of constant energy E for one of the two band edges (say, the conduction band). The contours corresponding to $E = 0$ are small closed curves which surround electron puddles. As the energy E is increased, these contours grow in diameter and an increasing fraction of the system's area is inscribed within such contours. This inscribed area corresponds to the regions of the system that is accessible to conduction band electrons with energy E .

At some point, as the contour energy E is increased, the area inscribed within contours becomes large enough that it comprises half of the total area of the system. At this energy

E_p there is a percolation transition, such that a conduction band electron with $E > E_p$ can move freely across a macroscopic distance and contribute to bulk conduction. The energy E_p is therefore equal to the position of the classical mobility edge for the conduction band, and the activation energy for this process $E_a = E_p$. That percolation occurs when half the area of the system is encompassed by contours can be argued simply on the basis of symmetry: in two dimensions it is not possible for a continuous percolation cluster and its complement to percolate simultaneously. Thus the critical area fraction for a continuous and symmetric random potential is equal to $1/2$ [2].

Thus, the energy of activation to the conduction band is equal to the energy at which half of the system's area of accessible via conduction band states. For a symmetric potential this energy corresponds exactly to the mean energy of the conduction band, and therefore $E_a = \Delta$, as announced above.

This same argument can easily be extended to the case where the chemical potential is not at zero, but is instead at some finite energy μ relative to the mean position of the mid-gap. The percolation level E_p remains unchanged relative to the mid-gap, and thus the energy of activation to the conduction band is $\Delta - \mu$. The energy of activation to the valence band is $\Delta + \mu$.

Thus, the activation energy for activation to the nearest mobility edge is identical to what it would have been in the non-disordered case. Its origin, however, is nontrivial, and involves a symmetry of percolation in a two dimensional continuous potential. In three dimensions, for example, there is no such symmetry, since a given space and its complement may percolate simultaneously. Consequently the activation energy for percolation is significantly smaller than the difference in energy between the chemical potential and the mean position of the closest band edge, $E_a \approx 0.3|E_{\text{be}} - \mu|$ [3, 4].

2. Computer modeling

In order to confirm Eq. (A1), we consider a simple computer model of a 2D electron system of dimensions surrounded by a three-dimensional environment containing charged impurities. We simulate the electron system as a square grid of $L_0 \times L_0$ discrete points, embedded as the mid-plane of a cube containing L_0^3 charged impurities. We define our coordinates such that the 2D electron system comprises the $x - y$ plane, while impurities are uniformly and randomly distributed within the region $-L_0/2 < x, y, z < L_0/2$ (impurities are not constrained to reside at integer values of x, y, z). A given discrete point i in the $x-y$ plane may have charge $q_i = -1, 0, 1$, with the values corresponding to the location of the chemical potential relative to the conduction band and valence band edges at site i . Specifically, $q_i = -1$ corresponds to the chemical potential being above the conduction band edge, $q_i = +1$ indicates that the chemical potential is below the valence band edge, and $q_i = 0$ if the chemical potential is within the band gap. The corresponding semiclassical

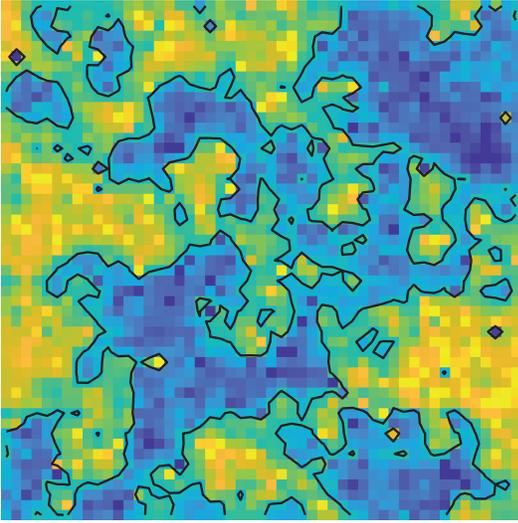


Figure 4. An example of the energy E_{CB} of the conduction band edge as a function of position, as given by our numerical simulation. Light (yellow) colours indicate high energy, and dark (blue) colors correspond to low energy. The black contour indicates the energy of the percolation level, $E_{CB} = \Delta$. This example corresponds to $\Delta = 5$, $L_0 = 50$.

Hamiltonian for this system is

$$H = \sum_i \phi_i^{\text{imp}} q_i - \Delta \sum_i q_i + \sum_{i,j \neq i} \frac{q_i q_j}{r_{ij}} \quad (\text{A3})$$

where r_{ij} is the distance between the sites i and j in the x - y plane. The quantity ϕ_i^{imp} represents the potential created by bulk impurities at site i , and is given by

$$\phi_i^{\text{imp}} = \sum_k \frac{q_k^{\text{imp}}}{r_{ik}}. \quad (\text{A4})$$

Here, r_{ik} represents the distance between the bulk impurity k and the site i ; the impurity charge is $q_k^{\text{imp}} = \pm 1$. The corresponding energies of the conduction band and valence band edges at site i are

$$E_i^{\text{CB}} = -\phi_i + \Delta \quad (\text{A5})$$

$$E_i^{\text{VB}} = -\phi_i - \Delta, \quad (\text{A6})$$

where

$$\phi_i = \phi_i^{\text{imp}} + \sum_{j \neq i} \frac{q_j}{r_{ij}} \quad (\text{A7})$$

is the total electric potential at site i . The intent of this model is to describe the spatial meandering of the band edge for a 2D system surrounded by bulk impurities. Once the corresponding energies $E_i^{\text{CB,VB}}$ are known, we can calculate the energy E_a that corresponds to the classical mobility edge for the nearest band. Specifically, E_a corresponds to the minimal energy such that there exists a percolating path across opposite

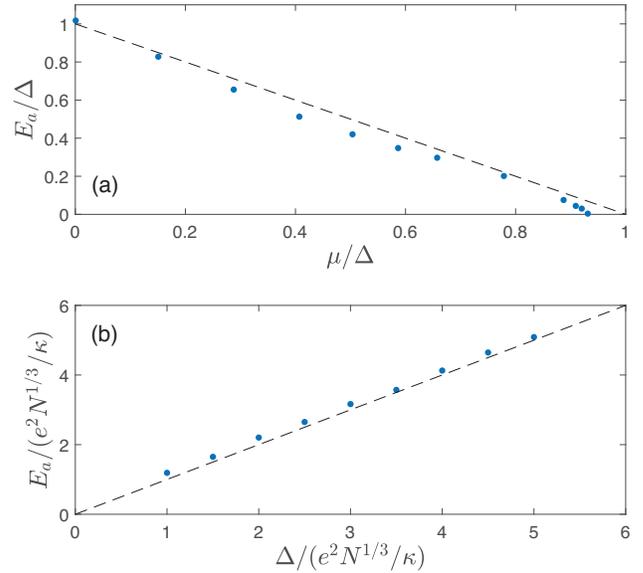


Figure 5. Numerical results for the activation energy as a function of (a) the chemical potential μ and (b) the band gap Δ . The dashed lines in (a) and (b) correspond to Eqs. (A1) and (A2), respectively. Both plots use a system size $L_0 = 50$.

faces of the system using only lattice sites with $E_i^{\text{CB}} < E_a$. The model discretizes the electron system on a length scale given by the mean distance between bulk impurities. The corresponding correlation length of the random potential is given by

$$\Lambda = \frac{\Delta^2 \kappa^2}{2\pi e^4 N}, \quad (\text{A8})$$

and so such discretization is unimportant when $\Delta/(e^2 N^{1/3}/\kappa)$ is large. We use nearest-neighbor percolation on the square lattice, but since the potential is correlated over length scales much larger than the lattice constant, this choice is unimportant in the limit of large Δ .

In order to find the energies $E_i^{\text{CB,VB}}$, we first need a solution for the charges q_i of each site i in the ground state. Finding such a solution is a difficult numerical problem. We use the numerical algorithm described in Refs. [2–4] to find a pseudo-ground state that is minimized with respect to changing any one or two values of q_i simultaneously. This algorithm is known to give a good approximation for the properties of the ground state. The resulting solution for the charge q_i of each site allows us to define the energies $E_i^{\text{CB,VB}}$, and thus to find the percolation level for both the CB and VB. We iterate this procedure over many random choices of the impurity positions, and average over all such iterations.

We also examine the dependence of the activation energy on the chemical potential μ . The chemical potential can be tuned by adjusting the net concentration of charged impurities, which mimics the effect of a gate voltage. $\mu = 0$ corresponds (on average) to equal numbers $L_0^{3/2}$ of + and – impurities. Since we are defining μ relative to the average energy of the mid-gap, which at site i has an energy $-\phi_i$. We can thus de-

fine μ for the system as

$$\mu = -\frac{1}{L_0^2} \sum_i \phi_i. \quad (\text{A9})$$

This chemical potential is calculated after the pseudo-ground state is determined. We then report E_a relative to this chemi-

cal potential.

Figure 4 shows an example plot of the energy of the CB bottom, E^{CB} , as a function of position. Figure 5 gives the calculated activation energy as a function of chemical potential and band gap. It closely matches Eq. A1.

-
- [1] B. I. Shklovskii and A. L. Efros, *Zh. Eksp. Theor. Fiz.* **62**, 1156 (1972), [*Sov. Phys. - JETP* 35, 610 (1972)].
- [2] B. I. Shklovskii and A. L. Efros, *Electronic properties of doped semiconductors*, Vol. 45 (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, *Science Advances* **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, *Nature Portfolio* (2021), 10.21203/rs.3.rs-519444/v1.
- [10] J. G. Checkelsky, J. Ye, Y. Onose, Y. Iwasa, and Y. Tokura, *Nature Physics* **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* **340**, 167 (2013).
- [12] K. He, X.-C. Ma, X. Chen, L. Lü, Y.-Y. Wang, and Q.-K. Xue, *Chinese Physics B* **22**, 067305 (2013).
- [13] M. Mogi, R. Yoshimi, A. Tsukazaki, K. Yasuda, Y. Kozuka, K. S. Takahashi, M. Kawasaki, and Y. Tokura, *Applied Physics Letters* **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 Materials* **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, *Nature Physics* **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 Letters* **19**, 3409 (2019), pMID: 31038971, <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(R) (2019).
- [19] L. K. Rodenbach, I. T. Rosen, E. J. Fox, P. Zhang, L. Pan, K. L. Wang, M. A. Kastner, and D. Goldhaber-Gordon, *APL Materials* **9**, 081116 (2021), <https://doi.org/10.1063/5.0056796>.
- [20] K. M. Fijalkowski, N. Liu, P. Mandal, S. Schreyeck, K. Brunner, C. Gould, and L. W. Molenkamp, *Nature Communications* **12**, 1 (2021).
- [21] K. Zou and J. Zhu, *Phys. Rev. B* **82**, 081407(R) (2010).
- [22] T. Taychatanapat and P. Jarillo-Herrero, *Phys. Rev. Lett.* **105**, 166601 (2010).
- [23] M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, *Science* **367**, 900 (2020).
- [24] 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* **583**, 375 (2020).
- [25] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, *Nature* **592**, 43 (2021).
- [26] 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* **556**, 80–84 (2018).
- [27] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Nature* **556**, 43–50 (2018).
- [28] A. L. Efros and B. I. Shklovskii, *J. Phys. C: Solid State Phys.* **8**, L49 (1975).
- [29] Y. Xie, B. Lian, B. Jäck, X. Liu, C.-L. Chiu, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, *Nature* **572**, 101 (2019).
- [30] Y. Choi, H. Kim, C. Lewandowski, Y. Peng, A. Thomson, R. Polski, Y. Zhang, K. Watanabe, T. Taniguchi, J. Alicea, and S. Nadj-Perge, “Interaction-driven band flattening and correlated phases in twisted bilayer graphene,” (2021), [arXiv:2102.02209](https://arxiv.org/abs/2102.02209).
- [31] R. Yu, W. Zhang, H.-J. Zhang, S.-C. Zhang, X. Dai, and Z. Fang, *Science* **329**, 61 (2010).
- [32] 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* **339**, 1582 (2013), <https://www.science.org/doi/pdf/10.1126/science.1230905>.
- [33] C.-X. Liu, H. J. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, *Phys. Rev. B* **81**, 041307(R) (2010).
- [34] H.-Z. Lu, W.-Y. Shan, W. Yao, Q. Niu, and S.-Q. Shen, *Phys. Rev. B* **81**, 115407 (2010).
- [35] J. Linder, T. Yokoyama, and A. Sudbø, *Phys. Rev. B* **80**, 205401 (2009).
- [36] 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, *Nature Physics* **6**, 584 (2010).
- [37] Y. Sakamoto, T. Hirahara, H. Miyazaki, S.-i. Kimura, and S. Hasegawa, *Phys. Rev. B* **81**, 165432 (2010).
- [38] T. Zhang, J. Ha, N. Levy, Y. Kuk, and J. Stroschio, *Phys. Rev. Lett.* **111**, 056803 (2013).
- [39] D. Kim, P. Syers, N. P. Butch, J. Paglione, and M. S. Fuhrer, *Nature Communications* **4**, 2040 (2013).
- [40] 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. Husain, and Z. X. Shen, *Science* **329**, 659 (2010).
- [41] 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, *Nature Physics* **8**, 616 (2012).
- [42] 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, *Nature Communications* **6**, 8913 (2015).
- [43] Y. Tokura, K. Yasuda, and A. Tsukazaki, *Nature Reviews Physics* **1**, 126 (2019).
- [44] Y. Deng, Y. Yu, M. Z. Shi, Z. Guo, Z. Xu, J. Wang, X. H. Chen, and Y. Zhang, *Science* **367**, 895 (2020).
- [45] 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, C. Liu, K. Deng, W. Huang, B. Shen, K. Shimada, E. F. Schwier, C. Liu, Q. Liu, and C. Chen, *Phys. Rev. X* **11**, 011039 (2021).
- [46] T. Chen, B. Skinner, and B. I. Shklovskii, *Phys. Rev. Lett.* **109**, 126805 (2012).
- [47] J. Zhang and B. I. Shklovskii, *Phys. Rev. B* **70**, 115317 (2004).
- [48] F. W. Van Keuls, X. L. Hu, H. W. Jiang, and A. J. Dahm, *Phys. Rev. B* **56**, 1161 (1997).
- [49] K. Bennaceur, P. Jacques, F. Portier, P. Roche, and D. C. Glattli, *Phys. Rev. B* **86**, 085433 (2012).
- [50] M. B. Isichenko, *Rev. Mod. Phys.* **64**, 961 (1992).
- [51] B. Skinner and B. I. Shklovskii, *Phys. Rev. B* **87**, 075454 (2013).
- [52] B. I. Shklovskii and A. L. Efros, *Zh. Eksp. Teor. Fiz.* **59**, 1343 (1970), [*Sov. Phys. - JETP* **32**, 733 (1971)].
- [53] M. E. Raikh and I. M. Ruzin, *Sov. Phys. Semicond.* **19**, 745 (1985).
- [54] In the limit of $\alpha \rightarrow 0$, the asymptotic expression to first order reads $(\Gamma/\Delta)_c = \alpha^{-1/3} [\ln(\alpha^{-1})]^{-2/3}$.
- [55] Note that this mesoscopic optimization method based on Eq. (19) is self-consistent if $\Gamma/\Delta < (\Gamma/\Delta)_c$ (or $G < 1$), so that $e^S > L/\lambda$.
- [56] H. Samavati, A. Hajimiri, A. Shahani, G. Nasserbakht, and T. Lee, *IEEE Journal of Solid-State Circuits* **33**, 2035 (1998).
- [57] W. Richter and C. R. Becker, *physica status solidi (b)* **84**, 619 (1977).
- [58] 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).
- [59] T. Bömerich, J. Lux, Q. T. Feng, and A. Rosch, *Phys. Rev. B* **96**, 075204 (2017).
- [60] 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, *Nature communications* **2**, 574 (2011).
- [61] N. S. Rytova, *Moscow University Physics Bulletin* **3**, 18 (1967).
- [62] A. Chaplik and M. Entin, *Journal of Experimental and Theoretical Physics* **34**, 1335 (1971), [*Zh. Eksp. Teor. Fiz.* **61**, 2496 (1971)].
- [63] L. V. Keldysh, *Soviet Journal of Experimental and Theoretical Physics Letters* **29**, 658 (1979), [*Pis'ma Zh. Eksp. Teor. Fiz.* **29**, 716 (1979)].
- [64] Y. Huang and B. I. Shklovskii, *Phys. Rev. B* **103**, 165409 (2021).
- [65] In the limit of $\alpha(Nd^3)^{-1/2} \rightarrow 0$, the asymptotic expression to first order reads $(\Gamma/\Delta)_c = \alpha^{-1/6} (Nd^3)^{1/12} \left\{ \ln[\alpha^{-1} (Nd^3)^{1/2}] \right\}^{-3/4}$.
- [66] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, *Phys. Rev. Lett.* **42**, 673 (1979).
- [67] S. Hikami, A. I. Larkin, and Y. Nagaoka, *Progress of Theoretical Physics* **63**, 707–710 (1980).
- [68] R. S. K. Mong, J. H. Bardarson, and J. E. Moore, *Phys. Rev. Lett.* **108**, 076804 (2012).
- [69] M. M. Mahmoodian and M. V. Entin, *Phys. Rev. B* **101**, 125415 (2020).