Magnetic-Field-Induced Localization in 2D Topological Insulators

Pierre Delplace, Jian Li, and Markus Büttiker

Département de Physique Théorique, Université de Genève, CH-1211 Genève, Switzerland

(Received 16 July 2012; published 11 December 2012)

Localization of the helical edge states in quantum spin Hall insulators requires breaking time-reversal invariance. In experiments, this is naturally implemented by applying a weak magnetic field *B*. We propose a model based on scattering theory that describes the localization of helical edge states due to coupling to random magnetic fluxes. We find that the localization length is proportional to B^{-2} when *B* is small and saturates to a constant when *B* is sufficiently large. We estimate especially the localization length for the HgTe/CdTe quantum wells with known experimental parameters.

DOI: 10.1103/PhysRevLett.109.246803

PACS numbers: 73.20.Fz, 72.10.-d, 75.47.-m

The prediction and discovery of quantum spin Hall insulators (QSHIs) [1-4] has opened a door to an unexpected category of topological phases in condensed matter [5-8] and revealed a new route to investigations of edge- or boundary-state physics [9-11]. Although the prototypes of QSHIs [1,3] are mainly based on two copies of quantum Hall insulators, which have been investigated for more than three decades [12,13], it was soon realized that the fundamental importance of time-reversal invariance (TRI) distinguishes the two systems in a profound way [2]. Indeed, QSHIs, unlike the \mathbb{Z} -classified quantum Hall insulators [14], belong to a class of two-dimensional time-reversalinvariant \mathbb{Z}_2 topological insulators [2]. The defining feature of QSHIs, as their name suggests, is a pair of helical edge states that persist in the bulk insulating gap of the system [1-4,9].

The topological power of QSHIs lies precisely in the robustness of the helical edge states against generic perturbations due to unavoidable disorder in every experimental setup, unless TRI is broken. In the presence of both TRI breaking and disorder, the helical edge states will be localized, and the general framework of Anderson's localization theory applies [15]. Nevertheless, the localization of the helical edge states distinguishes itself from conventional one-dimensional localization when the focus is placed on the crucial role that TRI plays in the problem. This point becomes especially relevant as TRI can be broken continuously, for instance, by turning on a magnetic field gradually. Indeed, the sensibility of transport through helical edge states to a weak out-of-plane magnetic field has been demonstrated experimentally in the measurement of magnetoconductance in topologically nontrivial HgTe/CdTe quantum wells [4,16]. In contrast, various experiments have shown that the edge transport only weakly depends on the in-plane magnetic field [4,16,17]. Related theoretical analyses have been carried out that include the interplay between TRI breaking and disorder but mainly consider magnetic impurities [18,19] or a bulk short-range random potential combined with a magnetic field [20]. A transparent edge theory that focuses on the magnetic-field-dependent localization of the helical edge states, however, is still missing.

In this Letter, we propose a model that explicitly addresses the question of how the localization of helical edge states occurs as a weak magnetic field is gradually turned on. Our model is based on the scattering theory of edge states in the presence of generic long-range disorder. In particular, we consider the existence of alternative paths for the edge states due to, e.g., constrictions formed at the rough edges of realistic samples [see Fig. 1], which further allows for loops of the helical edge states. The magnetic field penetrating through these loops results in broken TRI that is experienced by the helical edge states in the form of finite random fluxes. We show that these random fluxes necessarily lead to localization of the helical edge states with universal behaviors in two regimes: Immediately after the magnetic field is turned on, the localization length becomes finite and decreases as B^{-2} ; when the standard deviation of the random fluxes



FIG. 1 (color online). (a) Helical edge states in a disordered QSHI in a uniform magnetic field. Occasional occurrences of constrictions along the edge lead to Fabry-Perot-type loops where Aharonov-Bohm phases due to magnetic fluxes can accumulate. (b) The scattering of the helical edges by one of these loops can be divided into two parts: the scattering between two pairs of helical edge states and the propagation of one of these pairs around the loop. (c) Relevant scattering probabilities in the first part.

(proportional to *B*) is larger than one magnetic flux quantum, the localization length saturates to a constant. In between these two regimes, damped oscillations of the localization length may arise depending on the distribution of the random fluxes. Our results provide a clear illustration of the symmetry-breaking-induced localization in QSHIs.

We start to introduce our model by considering one of its possible realizations, depicted in Fig. 1. The edge roughness of a realistic QSHI sample may lead to occasional constrictions (at one edge) where the helical edge states can either tunnel across or pass around. As a consequence, loops can form and attach to the propagating path of the helical edge states. When a perpendicular magnetic field is applied to the sample [21], each of these loops acts as a magnetic flux impurity, to be distinguished from usual magnetic impurities. Individually, such an impurity works like a Fabry-Perot scatterer [see Fig. 1(b)], where the scattering probability amplitudes depend on the Aharonov-Bohm (AB) phase, owing to the magnetic flux, acquired by electrons circling around the loop. The collective action of a random distribution of these scatterers causes localization of the helical edge states with an explicit dependence on the magnetic field. The main part of this problem can be tackled consistently by scattering theory, as we now show.

To analyze the scattering of the helical edge states by a single magnetic flux impurity [see Fig. 1], we divide the full scattering process into two effective parts: the local scattering between two pairs of helical states and the free propagation of one pair of helical states that closes the loops. For simplicity, we assume that the first part does not depend on a magnetic field and, hence, respects *local* TRI, while the second part contains the entire information about the magnetic flux by means of AB phases that enter the propagators [22].

Owing to the local TRI, the scattering between two pairs of helical states, described by a 4×4 scattering matrix *S*, has the following constraint:

$$S = \Theta S^{\dagger} \Theta^{-1}, \tag{1}$$

where Θ is the time-reversal operator. We choose a specific basis ordered as (R_1, L_1, L_2, R_2) , where R_i (L_i) stands for the right (left) mover of the *i*th Kramers pair (i = 1, 2), such that $\Theta = \mathbb{1}_2 \otimes (-i\sigma_y \kappa)$ with σ_y the Pauli matrix and κ the complex-conjugate operator. Consequently, the scattering matrix *S* satisfying Eq. (1) necessarily has the following form:

$$S = \begin{pmatrix} t_1 & 0 & r' & s' \\ 0 & t_1 & -s & r \\ r & -s' & t_2 & 0 \\ s & r' & 0 & t_2 \end{pmatrix}.$$
 (2)

Here t_i stands for the direct transmission for Kramers pair i; r(r') stands for the reflection from a right (left) mover to

a left (right) mover; *s* and *s'* represent the transmission by switching to another Kramers pair. Importantly, zeros in *S* signify the absence of direct backscattering within one Kramers pair due to TRI. By taking into account the unitarity of the scattering matrix, the parametrization of *S* can be further simplified as (up to an unimportant global phase factor) $t_1 = -t_2^* = t$, $r' = r^*$, $s' = s^*$, and $T + R + T_s = 1$, where $T = |t|^2$, $R = |r|^2$, and $T_s = |s|^2$.

The free propagation of Kramers pair 2 leads to a scattering matrix S_{Φ} given in the basis (L_2, R_2) by

$$S_{\Phi} = \begin{pmatrix} e^{i(\varphi+\phi)} & 0\\ 0 & e^{i(\varphi-\phi)} \end{pmatrix}$$
(3)

with φ the dynamical phase and ϕ the AB phase (equal to the total flux enclosed by the loop in units of $h/2\pi e$). Therefore, the final scattering matrix S for Kramers pair 1, in the basis (R_1 , L_1), reads

$$S = \begin{pmatrix} |t| + RZ_+ + T_s Z_- & -(rs)^* \Delta Z \\ -rs \Delta Z & |t| + RZ_- + T_s Z_+ \end{pmatrix}, \quad (4)$$

where

$$Z_{\pm} = \frac{e^{i(\varphi \pm \phi)}}{1 + |t|e^{i(\varphi \pm \phi)}}, \qquad \Delta Z = Z_{+} - Z_{-}, \qquad (5)$$

and the phase of t^* has been absorbed into the dynamical phase φ . The backscattering probability can be obtained immediately:

$$\mathcal{R} = \frac{RT_s}{T} \frac{\sin^2 \phi}{|\cos \phi + \cosh(\ln|t| + i\varphi)|^2}.$$
 (6)

Evidently, for the helical edge states to be backscattered with finite probability, two conditions must be fulfilled. First, it is necessary that both *R* and *T_s* are finite. If one of these two tunneling probabilities is zero, the system essentially reduces to two decoupled copies of quantum Hall edge states, and backscattering is known to be prohibited for either copy [23]. Only when both tunneling processes (*R* and *T_s*) are allowed does the system belong truly to the \mathbb{Z}_2 -classified symmetry class where TRI plays a central role. It follows that the second necessary condition for backscattering to occur is to break the *global* TRI by having $\phi \neq 0 \mod \pi$. The cooperation of these two conditions clearly illustrates the underlying protection mechanism, from the scattering point of view, for the helical edge states of QSHIs.

At disordered sample edges, the magnetic flux impurities will occur randomly, and the helical edge states will eventually be localized. Here, we assume not only that the variables (including phases and scattering amplitudes) for each individual scatterer are random but also that different scatterers are independent. In this scenario, the total transmission probability \mathcal{T}_N through N scatterers obeys the following scaling law [24] (see also Supplemental Material [25])

$$\langle \ln \mathcal{T}_N \rangle = \langle \ln \mathcal{T}_{N-1} \rangle + \langle \ln |\tau_N|^2 \rangle, \tag{7}$$

where $|\tau_N|^2$ is the transmission probability for the *N*-th scatterer. The inverse localization length $\gamma = 1/\ell_{\text{loc}}$ is defined as [24,26,27]

$$\gamma = -\lim_{N \to \infty} \frac{n}{N} \langle \ln \mathcal{T}_N \rangle, \tag{8}$$

where *n* is the linear density of the scatterers. Thanks to the scaling law, γ can be expressed in terms of the scattering amplitudes for a *single* scatterer:

$$\gamma = -n\langle \ln(1 - \mathcal{R}) \rangle \simeq n\langle \mathcal{R} \rangle, \tag{9}$$

where we have assumed weak backscattering for each individual scatterer ($\mathcal{R} \ll 1$). The average over the arbitrary dynamical phase φ can be carried out exactly and yields

$$\gamma = n \left\langle \frac{RT_s}{R+T_s} \frac{1+T}{T} \frac{\sin^2 \phi}{\sin^2 \phi + \frac{(1-T)^2}{4T}} \right\rangle.$$
(10)

By further using the fact that $\phi = BA(2\pi e/h)$, where the magnetic field *B* is taken to be uniform and *A* is the area enclosed by the helical loops, we will need only to average over distributions of the scattering amplitudes and the area *A* in order to estimate the localization length $\ell_{\text{loc}} = 1/\gamma$. One immediate consequence of Eq. (10) is that the localization length is magnetic-field-symmetric, which is certainly expected.

An important regime that we are particularly interested in is the weak magnetic-field regime, where $B\bar{A} \ll h/e$ with \bar{A} the mean of A. In this regime Eq. (10) becomes (assuming T is not too close to 1)

$$\gamma = \alpha B^2 \tag{11}$$

with $\alpha = 4n(2\pi e/h)^2 \langle RT_s(1 + T)/(R + T_s)^3 \rangle \langle A^2 \rangle$ a constant factor for given distributions of scattering amplitudes and A. The B^2 dependence of the inverse localization length here is a *universal* result of our model in the sense that it does not depend on the specific distributions of variables for individual scatterers. It implies that the localization length of the helical edge states is finite at a weak magnetic field and diverges only as $1/B^2$ when the magnetic field is vanishing. Furthermore, the low-temperature magnetoconductance of a QSHI should also vary as B^2 in the weak magnetic-field limit, that is, $\delta G(B) = G(0) - G(B) \propto B^2$. This contrasts our result with the linear magnetoconductance behavior previously found on a lattice model with bulk impurity potentials [20].

Another interesting regime is when the magnetic field is strong enough such that both $B\bar{A}$ and $B\sigma_A$ [with $\sigma_A \equiv \sqrt{\langle (A - \bar{A})^2 \rangle}$ being the standard deviation] are larger than a flux quantum h/e. In this regime we can approximate the average in terms of A as an average over a uniform distribution of ϕ , which yields

$$\gamma_{\rm sat} = 2n \left\langle \frac{RT_s}{R+T_s} \right\rangle. \tag{12}$$

This especially simple result again shows a universal behavior in our model—the inverse localization length saturates at a sufficiently strong magnetic field irrespective of the specific distributions of scattering variables. However, the actual value of γ_{sat} certainly depends on the distributions of *R* and *T_s*. Moreover, the above formula reemphasizes the importance of allowing both tunneling processes represented by *R* and *T_s* to evoke a true protection mechanism due to TRI.

In between the two regimes discussed above, we need to consider the specific distributions of the variables in Eq. (10). Let us first focus on the behavior of γ by assuming that A has a Gaussian distribution characterized by the mean \overline{A} and the standard deviation σ_A . It is instructive in this case to look at the function $\Gamma(B) \equiv [(1+T)/2T] \times$ $\langle \sin^2 \phi / [\sin^2 \phi + (1 - T)^2 / 4T] \rangle_A$ with the average taken only in terms of A. $\Gamma(B)$ has been defined such that it saturates to the value 1 at a sufficiently strong magnetic field. In Fig. 2, we plot the numerically obtained $\Gamma(B)$ for various T and fixed \overline{A} and σ_A . Right after the magnetic field is turned on, $\Gamma(B)$ shows a quadratic increase irrespective of the assumed T or the distribution of A. Before $\Gamma(B)$ saturates, it undergoes damped oscillations when σ_A/\bar{A} is small. These oscillations are due to the collective AB effect for the helical loops in our model: The loops enclosing a similar area lead to AB oscillations of a similar period; they contribute coherently to the backscattering of the helical edge states; the magnetic-field dependence of the total transmission is thus shaped by the AB effect at individual scatterers when $B\sigma_A$ is significantly smaller than $\phi_0 = h/e$. The period of the oscillations is roughly $\phi_0/2\bar{A}$, where the factor 1/2 is obviously a consequence of Γ (and hence γ) depending only on $\sin^2 \phi$. The overall amplitude of the oscillations is suppressed for large T and enhanced for small T. The reason is intuitively clear: The



FIG. 2 (color online). $\Gamma(B)$ for different values of *T*, where *A* is assumed to take a normal distribution with $\sigma_A/\bar{A} = 0.3$.

more the helical edge states are scattered into the loops, the more pronounced the resulting AB effect.

Now we address the issue of the scattering amplitudes which have only been assumed to be phenomenological parameters in the scattering matrix S so far. To this end, we investigate a constriction depicted in Fig. 3, which is described by the following effective Hamiltonian:

$$\mathcal{H} = \begin{pmatrix} \hbar \upsilon_F \hat{k}_x & 0 & m(x) & \delta(x) \\ 0 & -\hbar \upsilon_F \hat{k}_x & -\delta(x) & m(x) \\ m(x) & -\delta(x) & -\hbar \upsilon_F \hat{k}_x & 0 \\ \delta(x) & m(x) & 0 & \hbar \upsilon_F \hat{k}_x \end{pmatrix}, \quad (13)$$

where v_F is the Fermi velocity for the helical edge states, m(x) and $\delta(x)$ represent x-dependent coupling between the edge states, and the basis is ordered as (R_1, L_1, L_2, R_2) . The above Hamiltonian manifestly respects TRI: $\mathcal{H} = \Theta \mathcal{H} \Theta^{-1}$ [22]. This Hamiltonian can be derived from microscopic models such as the Bernevig-Hughes-Zhang model for HgTe/CgTe quantum wells [3,4,16,28,29]. For simplicity, we take $m(x) = m_W \theta(x) \theta(L - x)$ and $\delta(x) = \delta_W \theta(x) \theta(L - x)$ with $\theta(x)$ the Heaviside step function and m_W and δ_W two constants determined by the constriction separation W. In the case of HgTe/CgTe quantum wells, a nonvanishing δ term owes its existence to the presence of bulk-inversion asymmetry [4,16].

The scattering amplitudes for this constriction, corresponding to S in Eq. (2), can be easily derived (see Supplemental Material [25]):

$$t = i\cos(\delta_W L/\hbar v_F)/\zeta, \qquad (14)$$

$$s = \sin(\delta_W L/\hbar v_F)/\zeta, \qquad (15)$$

$$r = m_W \sin(qL)/\hbar v_F q \zeta, \tag{16}$$

where $q = \sqrt{E^2 - m_W^2}/\hbar v_F$ can be either real or imaginary, depending on the energy *E*, and $\zeta = |(E/\hbar v_F q) \times \sin(qL) + i \cos(qL)|$ is a normalization factor. Clearly,



FIG. 3 (color online). A schematic view of a constriction where two pairs of helical edge states (indicated by the linear bands) are coupled (indicated by the mixed bands) in a region of length L and separation W.

r vanishes when $m_W = 0$, which shows the fact that *m* couples R_1 (L_1) to L_2 (R_2); *s* vanishes when $\delta_W = 0$, which shows the fact that δ couples R_1 (L_1) to R_2 (L_2). For low energy ($|E| < m_W$) scattering states, $R/(T + T_s) \approx \sinh^2(m_W L/\hbar v_F)$ if $|E| \ll m_W$, and $R/(T + T_s) \approx (m_W L/\hbar v_F)^2$ if $|E| \approx m_W$, meaning that the reflection dominates as long as *L* is large compared with $\hbar v_F/m_W$. In this regime, Eq. (10) reduces to $\gamma \approx 4n\langle T_s \rangle \langle \sin^2 \phi \rangle$.

More generally the average with respect to the scattering amplitudes has to be performed numerically by taking certain distributions of W and L (at a certain energy E). The advantage of this change of variables is that W and L, unlike the scattering amplitudes, are in principle independent to each other. In total, this leads to three independent geometric variables, W, L, and A, that remain to be averaged on in our final evaluation of γ as a function of magnetic field B (we leave the density n of scatterers as a parameter). After carrying out these averages numerically (see Supplemental Material [25]), we show the typical results in Fig. 4. The qualitative behavior of $\gamma(B)$ in Fig. 4 is essentially the same as that of $\Gamma(B)$ in Fig. 2, except that $\gamma(B)$ is shown for various ratios σ_A/\bar{A} whereas the scattering amplitudes have been averaged out. As universal features, γ increases as B^2 at a weak magnetic field and saturates at a sufficiently strong magnetic field. Despite the fact that the exact value of γ_{sat} depends on the energy E and the distributions of W and L, the order of magnitude of γ_{sat} turns out to be consistently 0.01*n* for all cases with realistic considerations (see Supplemental Material [25]). We also observe in the intermediate regime damped oscillations of γ which are pronounced if σ_A/\bar{A} is small but suppressed as long as σ_A/\bar{A} is close to or larger than 1. We point out that the localization length has a local maximum (minimum) whenever B is roughly an integer



FIG. 4 (color online). Inverse localization length γ as a function of magnetic field *B*, for different distributions of loop area *A*. It shows a B^2 increase in the weak *B* limit (inset) and a saturation at a relatively high field. In between, damped oscillations may occur if $\sigma_A/\bar{A} < 1$. The saturation value γ_{sat} is sample-dependent but has an order of magnitude 0.01n (with *n* the linear density of the scatterers) in our realistic estimations.

(half-integer) multiple of $\phi_0/2\bar{A}$ —this is where the TRI is maximally preserved (broken).

In summary, we have investigated a simple yet illuminating model that demonstrates a magnetic-field-dependent localization of the helical edge states in quantum spin Hall insulators. We have identified universal, sample-independent features, as well as an interesting but sample-specific feature in this model. With known parameters for the HgTe/CgTe quantum wells, we have also estimated quantitatively the localization length. Both the qualitative and the quantitative results can be examined by experiments.

P.D. was supported by the European Marie Curie ITN NanoCTM, and J.L. was supported by the Swiss National Center of Competence in Research on Quantum Science and Technology. In addition, we acknowledge the support of the Swiss National Science Foundation. The authors thank Alberto Morpurgo and Mathias Albert for inspiring discussions.

- C. L. Kane and E. J. Mele, Phys. Rev. Lett. 95, 226801 (2005).
- [2] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
- [3] B.A. Bernevig, T.A. Hughes, and S.-C. Zhang, Science 314, 1757 (2006).
- [4] M. Konig, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science 318, 766 (2007).
- [5] A. Kitaev, AIP Conf. Proc. 1134, 22 (2009).
- [6] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, AIP Conf. Proc. **1134**, 10 (2009).
- [7] M.Z. Hasan and C.L. Kane, Rev. Mod. Phys. 82, 3045 (2010).
- [8] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. 83, 1057 (2011).
- [9] A. Roth, C. Brune, H. Buhmann, L. W. Molenkamp, J. Maciejko, X. Qi, and S. Zhang, Science 325, 294 (2009).
- [10] M. Büttiker, Science 325, 278 (2009).

- [11] C. Brüne, A. Roth, H. Buhmann, E. M. Hankiewicz, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Nat. Phys. 8, 486 (2012).
- [12] K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. 45, 494 (1980).
- [13] F.D.M. Haldane, Phys. Rev. Lett. 61, 2015 (1988).
- [14] D.J. Thouless, M. Kohmoto, M.P. Nightingale, and M. den Nijs, Phys. Rev. Lett. 49, 405 (1982).
- [15] P.W. Anderson, Phys. Rev. 109, 1492 (1958).
- [16] M. Konig, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, J. Phys. Soc. Jpn. 77, 031007 (2008).
- [17] I. Knez, R.-R. Du, and G. Sullivan, Phys. Rev. Lett. 107, 136603 (2011).
- [18] Y. Tanaka, A. Furusaki, and K. A. Matveev, Phys. Rev. Lett. 106, 236402 (2011).
- [19] K. Hattori, J. Phys. Soc. Jpn. 80, 124712 (2011).
- [20] J. Maciejko, X.-L. Qi, and S.-C. Zhang, Phys. Rev. B 82, 155310 (2010).
- [21] We focus on the weak magnetic field regime, such that both the Zeeman effect and the separation between originally time-reversed trajectories can be neglected.
- [22] The assumption of local TRI is valid approximately in the weak magnetic field regime and when the length scale of the local tunneling is small compared with that of the loops.
- [23] M. Büttiker, Phys. Rev. B 38, 9375 (1988).
- [24] P.W. Anderson, D.J. Thouless, E. Abrahams, and D.S. Fisher, Phys. Rev. B 22, 3519 (1980).
- [25] See Supplemental Material http://link.aps.org/ supplemental/10.1103/PhysRevLett.109.246803 for the derivation of the scaling formula Eq. (7), the derivation of the scattering amplitudes Eqs. (14)–(16), and the details of the numerical simulations.
- [26] J. Pendry, Adv. Phys. 43, 461 (1994).
- [27] C.A. Müller and D. Delande, in *Ultracold Gases and Quantum Information*, Proceedings of the Les Houches Summer School, Session XCI (World Scientific, Singapore, 2011), Chap. 9.
- [28] B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).
- [29] V. Krueckl and K. Richter, Phys. Rev. Lett. 107, 086803 (2011).