Topologically protected extended states in disordered quantum spin-Hall systems without time-reversal symmetry

Zhong Xu, L. Sheng, D. Y. Xing, Emil Prodan, and D. N. Sheng

Phys. Rev. B **85**, 075115 — Published 13 February 2012

DOI: 10.1103/PhysRevB.85.075115

Topologically Protected Extended States in Disordered Quantum Spin-Hall Systems without Time-Reversal Symmetry

Zhong Xu[1], L. Sheng[1],* D. Y. Xing[1], Emil Prodan[2],† and D. N. Sheng[3]
[1] *National Laboratory of Solid State Microstructures and Department of Physics, Nanjing University, Nanjing 210093, China*
[2] *Department of Physics, Yeshiva University, New York, NY 10016, USA*
[3] *Department of Physics and Astronomy, California State University, Northridge, California 91330, USA*

We demonstrate the existence of robust bulk extended states in the disordered Kane-Mele model with vertical and horizontal Zeeman fields, in the presence of a large Rashba coupling. The phase diagrams are mapped out by using level statistics analysis and computations of the localization length and spin-Chern numbers $C_\pm$. $C_\pm$ are protected by the finite energy and spin mobility gaps. The latter stays open for arbitrarily large vertical Zeeman fields, or for horizontal Zeeman fields below a critical strength or at moderate disorder. In such cases, a change of $C_\pm$ is necessarily accompanied by the closing of the mobility gap at the Fermi level. The numerical simulations reveal sharp changes in the quantized values of $C_\pm$ when crossing the regions of bulk extended states, indicating that the topological nature of the extended states is indeed linked to the spin-Chern numbers. For large horizontal Zeeman fields, the spin mobility gap closes at strong disorder prompting a change in the quantized spin-Chern numbers without a closing of the energy mobility gap.

PACS numbers: 72.25.-b, 72.10.Fk, 73.20.Jc, 73.43.-f

I. INTRODUCTION

Topological insulators (TIs) are materials characterized by robust properties against smooth deformations and disorder. Representative examples are the Chern (CI),[1] quantum spin-Hall (QSH),[2–5] and the strong topological insulators.[6,7] Samples of these materials display robust conducting states at the edges while being insulators in the bulk. The edge modes are connected to robust metallic states residing in the bulk and away from the Fermi level.[8,9] The robustness of the edge and the bulk metallic states is a manifestation of a bulk topological invariant and, as such, there has been a concentrated effort on understanding the bulk states in disordered TIs.[10–19]

In two-dimensional bulk systems, the extended quantum states can survive localization only in extraordinary circumstances. In CIs, the bulk extended states carry a Chern number and, as such, they cannot be destroyed until these numbers are "annihilated" through collision with other extended states. In QSH insulators with time-reversal symmetry (TRS), it is believed that a similar scenario happens, but this time the bulk extended states carry the $\boldsymbol{Z}_2$ invariant,[3] which is protected by the insulating gap and TRS. Consequently, TRS is thought to be crucial for delocalization in the bulk QSH insulators. The lack of a definitive answer is due to the lack of a complete theory of the $\boldsymbol{Z}_2$ invariant for aperiodic systems (for recent progress see Refs. 20 and 21). For example, the behavior of the $\boldsymbol{Z}_2$ invariant is not understood when the insulating gap is filled with dense localized spectrum, and the numerical approaches[12,17] based on the twisted boundary conditions (BC) were not able to probe this regime. Nevertheless, transfer matrix calculations[11,15] and level statistics analyses[19] have firmly established the presence of bulk extended states in representative models of disordered QSH systems with TRS.

An alternative approach is provided by the spin-Chern invariants $C_\pm$.[10,22] While for QSH systems with TRS the $C_\pm$ and the $\boldsymbol{Z}_2$ invariants predict perfectly consistent phase diagrams,[19,23,24] they give conflicting predictions when TRS is broken. An early study[11] on the bulk extended states in disordered QSH systems concluded that the protection against localization comes from both topology and TRS: Lack of any of the two will result in the immediate destruction of the bulk extended states. This conclusion was aligned with the predictions based on the $\boldsymbol{Z}_2$ invariant and was quickly accepted by the community. But recent studies on QSH systems with broken TRS revealed a more complex picture. Ref. 25 showed that robust gapless edge states can happen in QSH systems with TRS replaced by other symmetries. Yang *et al.*[26] studied the Kane-Mele model[2,3] with a Rashba coupling and vertical Zeeman field and found a topological QSH-like phase characterized by $C_\pm$ that remain quantized until the bulk energy gap closes. Such quantization indicates that the topological order of these QSH systems is intact, and thus extended bulk states should exist.

In this paper, we establish the existence of such extended states in disordered QSH systems with strongly broken TRS. The phase diagrams are explored using the level statistics analysis,[27] the localization length calculations,[28,29] and the spin-Chern invariants.[10,22] We show that the Kane-Mele model[2,3] with a large vertical Zeeman field displays a nontrivial phase diagram in the (E_F, W) plane (E_F = Fermi level and W = disorder strength), with a topological phase characterized by quantized spin-Chern numbers, surrounded by CI phases. *Each phase is completely surrounded by lines of robust extended states.* The situation remains the same for horizontal Zeeman fields up to a critical strength when the spin gap closes at strong disorder and the extended states disappear without a closing of the energy mobility gap.

Our results can be consistently explained based on the non-commutative theory of the spin-Chern invariants.

II. THE NON-COMMUTATIVE SPIN-CHERN NUMBERS.

The non-commutative spin-Chern numbers and their quantization conditions have been discussed extensively in Refs. 22, 30 and 19 and they will be briefly summarized here. We elaborate, however, on the question of when do the quantization conditions hold and what are the physical effects triggered by a change of the spin-Chern numbers.

A. Definition and The Quantization Condition

In a lattice model with many-orbitals per site $|\boldsymbol{n},\alpha,\sigma\rangle$, where $\boldsymbol{n}$ are the coordinates of a site, $\sigma=\pm1$ represents the spin degree of freedom and α some other possible quantum numbers. Let $\hat{\sigma}_z$ be the spin operator $\hat{\sigma}_z|\boldsymbol{n},\alpha,\sigma\rangle=\sigma|\boldsymbol{n},\alpha,\sigma\rangle$. Since we are dealing with a general Hamiltonian H, the space of occupied quantum states is not invariant under the action of the spin operator but, if P denotes the projector onto the occupied states, one can consider the projected spin operator: $P\hat{\sigma}_zP$, which does leave this space invariant. Therefore, if one considers the spectral projectors $P_\pm$ onto the positive/negative spectrum of $P\hat{\sigma}_zP$, one can effectively obtain an exact splitting $P=P_-\oplus P_+$ of the occupied space into two spin sectors. In the presence of disorder, one can associate non-commutative Chern numbers[31] $C_\pm$ to the projectors $P_\pm$,[22] which are referred to as the spin-Chern numbers. They are defined by the following explicit formula:[19]

$$C_\pm = 2\pi i\ \mathbb{E}\left\{\mathrm{tr}_0\left\{P_\pm\left[-i[\hat{x}_1,P_\pm],-i[\hat{x}_2,P_\pm]\right]\right\}\right\}, \quad (1)$$

where $\mathbb{E}\{\ \}$ denotes the disorder average, tr_0 means the trace over the states at $\boldsymbol{n}=\boldsymbol{0}$ and $\hat{\boldsymbol{x}}=(\hat{x}_1,\hat{x}_2)$ is the position operator.

The spin-Chern numbers $C_\pm$ remain quantized and invariant as long as $\lambda_\pm^2=\mathbb{E}\left\{\mathrm{tr}_0\left\{P_\pm\hat{\boldsymbol{x}}^2P_\pm\right\}\right\}<\infty$.[19,30] These are the quantization conditions. The quantities $\lambda_\pm$ can be viewed as localization lengths and $\lambda_\pm<\infty$ is enforced by a positive mobility gap at E_F in the Hamiltonian spectrum, and by a positive mobility gap at the origin in the spectrum of the projected spin operator $P\hat{\sigma}_zP$. We will refer to these two mobility gaps as the energy and the spin mobility gaps, respectively. The closing of either of these two gaps can prompt a sudden change of the spin-Chern numbers.

The non-commutative spin-Chern numbers defined in Eq. (1) can be numerically evaluated using only one (periodic-) boundary condition and one disorder configuration ($C_\pm$ are self-averaging).[19] This provides a tremendous numerical advantage over other existing methods for the evaluation of the Chern numbers, such as those based on twisted boundary conditions, and this together with the precise quantization conditions make $C_\pm$ very effective practical tools.[18]

B. Discussion of The Quantization Condition

Let us denote the localization length of the full projector P by λ_{E_F}: $\lambda_{E_F}^2=\mathbb{E}\left\{\mathrm{tr}_0\left\{P\hat{\boldsymbol{x}}^2P\right\}\right\}$. Note that in general $\frac{1}{2}\lambda_{E_F}^2\leq\lambda_-^2+\lambda_+^2$, so $\lambda_\pm<\infty$ implies $\lambda_{E_F}<\infty$, and $\lambda_{E_F}\to\infty$ implies λ_- or $\lambda_+\to\infty$. Therefore, whenever $C_\pm$ are seen to take quantized values, one can safely conclude that the quantum states are localized at E_F. In practice, this quantization was observed to hold with many digit precision, which is a remarkable fact given that it occur in the presence of strong disorder.[18] It also follows that $C_\pm$ will almost sure take non-quantized values when the states are delocalized at E_F. This allows one to map the regions of the phase diagrams that can harbor extended states, which are necessarily located in between the regions where $C_\pm$ assume different quantized values. The spin-Chern numbers can, in principle, take quantized values just by accident, even when the states are delocalized at E_F, but these are events with extremely low probability.

For clean QSH systems with TRS and certain models without TRS, it is known that the spin-gap stays open when the system crosses the QSH phase boundary where the energy gap closes.[19,23,26] This remains true in the weak disorder regime and, as such, a sudden change of $C_\pm$ necessarily implies delocalization at E_F. For QSH systems with TRS, the spin-gap is known to stay open beyond the weak disorder regime.[19]

In general, for systems with and without TRS, we have the following simple identity:

$$(P\hat{\sigma}_zP)^2=P\left(1-(i[\hat{\sigma}_z,P])^2\right)P, \quad (2)$$

from where one can see that the spectrum of $P\hat{\sigma}_zP$ near the origin is determined by the spectrum of $i[\hat{\sigma}_z,P]$ near ±1. We recall that we are interested in the nature of the spectrum of $P\hat{\sigma}_zP$ near the origin because we want to understand the localization properties of the projectors onto the positive and negative parts of its spectrum. So in some sense the origin plays here the same role as the Fermi level does for the energy spectrum. Now, the spectrum of $i[\hat{\sigma}_z,P]$ is contained inside the interval $[-1,1]$, so the spectrum of $P\hat{\sigma}_zP$ near the origin is determined by the edges of the spectrum of $i[\hat{\sigma}_z,P]$. Therefore, the main question becomes: what is the nature of the spectrum of $i[\hat{\sigma}_z,P]$ near its edges?

To address this question, we looked at the average and the fluctuations of the matrix elements of $i[\hat{\sigma}_z,P]$. Using $P=\oint_{\mathcal{C}}(\zeta-H)^{-1}\frac{d\zeta}{2\pi i}$, with $\mathcal{C}$ encircling the energy spectrum below E_F, one has

$$i[\hat{\sigma}_z,P]=\tfrac{1}{2\pi}\oint_{\mathcal{C}}(H-\zeta)^{-1}[\hat{\sigma}_z,H](H-\zeta)^{-1}d\zeta\ . \quad (3)$$

Note that $[\hat{\sigma}_z, H]$ is independent of the (non-magnetic) disorder. Using the universal localization estimate:[32]

$$\mathbb{E}\{|\langle \boldsymbol{n}|(H-E)^{-1}|\boldsymbol{m}\rangle|^s\} \leq c_s e^{-s|\boldsymbol{n}-\boldsymbol{m}|/\lambda_E}, \qquad (4)$$

valid for any $s < 1$ (c_s = generic constant, λ_E = the localization length at E), one can establish (see Appendix):

$$\mathbb{E}\left\{|\langle \boldsymbol{n}|i[\hat{\sigma}_z, P]|\boldsymbol{m}\rangle|\right\} \leq c_s \overline{[\hat{\sigma}_z, H]}^{\frac{s}{2}} \lambda_{E_F}^2 e^{-s|\boldsymbol{n}-\boldsymbol{m}|/4\lambda_{E_F}} , \qquad (5)$$

where $\bar{\mathcal{O}}$ = the largest matrix element of $\mathcal{O}$. Thus the averaged matrix elements of $i[\hat{\sigma}_z, P]$ decay exponentially and are small when $[\hat{\sigma}_z, H]$ and λ_{E_F} are both small, and the same is true for the fluctuations. In such cases, the spectrum of $i[\hat{\sigma}_z, P]$ is expected to be absent or localized at the edges ± 1, and consequently the spectrum of the projected spin operator is expected to be localized near the origin. Thus quantization and invariance of the spin-Chern numbers should hold. Another important conclusion is that, whenever we see a sudden change in $C_\pm$ and $[\hat{\sigma}_z, H]$ is small, which is the case for the typical values of the Rashba interaction, E_F necessarily crosses an energy region where λ_{E_F} is large or infinite.

Heuristically, we can go further and make the following predictions: 1) Since a vertical Zeeman field enhances the spin-gap, in a QSH system with only vertical Zeeman field, a sudden change in $C_\pm$ signals the divergence of λ_{E_F}; 2) Since a horizontal Zeeman field has a direct effect on $i[\hat{\sigma}_z, P]$, the spin-gap is expected to decrease and then close as the field is increased. Nevertheless, the spin-gap is expected to remain open for fields below a threshold value or at moderate disorder, where a change in $C_\pm$ will signal again the divergence of λ_{E_F}.

III. THE MODEL

The above predictions will be confirmed by a numerical analysis. Our numerical simulations are based on the Kane-Mele model,[2,3] which is defined on a honeycomb lattice:

$$H = -t \sum_{\langle nm\rangle} c_n^\dagger c_m + \tfrac{2i}{\sqrt{3}} V_{SO} \sum_{\langle\langle nm\rangle\rangle} c_n^\dagger \hat{\boldsymbol{\sigma}} \cdot (\boldsymbol{d}_{km} \times \boldsymbol{d}_{nk}) c_m$$
$$+ iV_R \sum_{\langle nm\rangle} c_n^\dagger \mathbf{e}_z \cdot (\hat{\boldsymbol{\sigma}} \times \boldsymbol{d}_{nm}) c_m + \sum_n c_n^\dagger (w_n + \boldsymbol{h}\cdot\hat{\boldsymbol{\sigma}}) c_n \ .$$

Here, the first term is the usual nearest neighbor hopping term with $c_n^\dagger = (c_{n,\uparrow}^\dagger, c_{n,\downarrow}^\dagger)$ as the electron creation operator on site n. In the following calculations, the hopping integral t will be chosen as the units of energy, for simplicity. The second term is the intrinsic spin-orbit coupling with $\hat{\boldsymbol{\sigma}}$ the Pauli matrix, where n and m are two next nearest neighbor sites, k is their common nearest neighbor, and vector $\boldsymbol{d}_{nk}$ points from k to n. The third term stands for the Rashba spin-orbit coupling. The last term represents our addition of a spin independent on-site potential w_n with random amplitudes uniformly distributed in $[-W/2, W/2]$, and of a uniform Zeeman field.

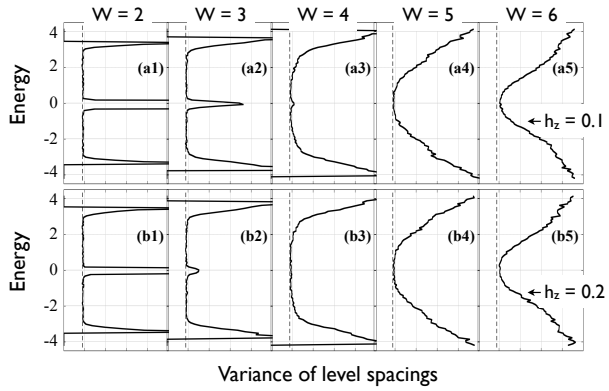

FIG. 1. Variance of the level spacings ensembles collected at various E's, W's and for $\boldsymbol{h} = h_z \mathbf{e}_z$ with $h_z = 0.1$ and 0.2. A lattice of 50×50 unit cells and 200 disorder configurations were used. The dotted line marks the 0.178 covariance of GUE. The windows span horizontally between 0 and 1.

Both cases will be studied, where the Zeeman field $\boldsymbol{h}$ is perpendicular or parallel to the plane of the lattice, i.e., $\boldsymbol{h} = h_z \hat{\mathbf{e}}_z$ or $\mathbf{h} = h_x \hat{\mathbf{e}}_x$. We will fix V_{SO} and V_R at 0.1, so that the model is inside the QSH phase when $\boldsymbol{h} = 0$ and $W = 0$. While we adopt this specific model and assume relatively high strengths of the spin-orbit couplings, which is mainly for the illustration purpose, our main conclusions are valid in general cases, as they depend only on the nontrivial topology of the energy bands and are independent of the details of the model.

IV. VERTICAL ZEEMAN FIELD

We consider first a vertical Zeeman field $\boldsymbol{h} = h_z \mathbf{e}_z$ and various disorder strengths. We investigate the localized/delocalized character of the energy spectrum by using level statistic analysis, the recursive Green's function method and scaling analysis.

A. Level Statistics Analysis

The level statistic analysis involves an exact diagonalization of the disordered Hamiltonian on a large lattice with periodic boundary conditions and for a large number of disorder configurations (= 200 for the present analysis). The level spacings were collected in the following way. We picked an arbitrary energy E and, for each disorder configuration, we identified the unique energy levels E_i and E_{i+1} that satisfy: $E_i < E < E_{i+1}$. Then we recorded the level spacings: $\Delta E = E_{i+j+1} - E_{i+j}$, letting j take consecutive values between $-k$ and k. We have experimented with $k = 5$ to 10 and the outcome of the analysis was virtually the same. The results reported here are for $k = 10$, and as such the ensemble of level spacings contains 4200 spacings. Fig. 1 plots the variance $\langle(\Delta E)^2\rangle / \langle \Delta E\rangle^2 - 1$ of the ensembles of level spacings

recorded at various energies E and disorder strengths W for $h_z = 0.1$ and $h_z = 0.2$.

The level spacing analysis allows one to identify the regions of localized and delocalized spectrum. It is a firmly established fact that the level spacings follow a Poisson distribution when the spectrum is localized, for which the variance equal to 1. It is also a firmly established fact that in the region of the spectrum where the localization length exceeds the size of the system, the statistics of the level spacings coincides with that of a random Gaussian ensemble of same general symmetries.[33,34] In our case the ensemble is the random Gaussian unitary ensemble (GUE), which has the distribution[27] $P_{\rm GUE}(s)=\frac{32s^2}{\pi^2}e^{-\frac{4}{\pi}s^2}$ with s as the level spacing ΔE normalized by the mean level spacing, which has a variance $[\langle(\Delta E)^2\rangle/\langle\Delta E\rangle^2 - 1] = 0.178$.

Examining Fig. 1, one can see that for both $h_z = 0.1$ and $h_z = 0.2$, there are energy regions where the variance is large, indicating localization, but also two regions where the variance is practically equal to 0.178, the variance of the GUE. We can conclude that in these two regions the localization length is larger than the size of the simulated lattice. However, simulations with increasing lattice size returned practically the same level statistics and one should note that the disorder strengths in Fig. 1 are comparable or exceeding the width of the clean bands yet the variance remains pinned at 0.178. This leaves very little doubt that the states in those two energy regions are in fact delocalized. We mention that the level statistics has been extensively analyzed for Chern and quantum spin-Hall insulators,[16,19] and the conclusions there were similar with the ones we draw here.

One distinct feature that can be seen in Fig. 1 is the drifting of the extended states regions towards each other when the disorder strength W is increased. The regions eventually collide, annihilate and then disappear. This is a typical behavior of the extended states carrying topological numbers robust against disorder, which can be understood in the following way: Such extended states cannot suddenly localize because in that case the topological number will suddenly switch to a trivial value. But the spectrum of any short range disorder model will eventually localize at extreme disorder,[32] and the only mechanism to localize the extended states carrying a topological number is through the neutralization of the topological number via collisions with one or more extended states carrying the opposite topological number. Therefore, the drifting and annihilation will always occur for the states carrying a topological number robust against disorder. Such drifting and annihilation was never observed in a trivial insulator. Thus, the drifting and annihilation observed in Fig. 1 is a strong indication of the presence of extended states carrying topological numbers, which arguably are the spin-Chern numbers discussed above.

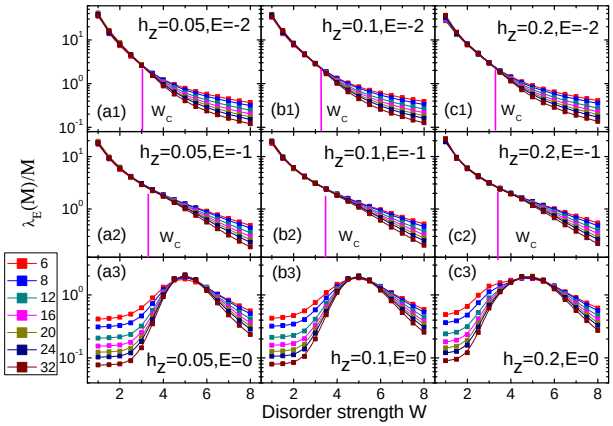

FIG. 2. Normalized localization length $\lambda_E(M)/M$ as a function of disorder strength W at various energies E, calculated for a long (10^6 unit cells in length) tube of M unit cells in circumference, in the presence of vertical Zeeman fields. Symbols with different colors correspond to different sizes.

B. The Recursive Green's Function Analysis

The delocalized character of the states can also be firmly established by using the recursive Green's function approach.[28,29] In Fig. 2 we plot the normalized localization length $\lambda_E(M)/M$ as a function of W for long tubes with circumferences ranging from $M = 6$ to 32 unit cells, and for $h_z = 0.05$, 0.1 and 0.2, and E=-2, -1 and 0. The scaling analysis for these data will be presented later. Now, for each E, we can see two regions separated by a critical disorder strength W_c, where $\lambda_E(M)/M$ behaves qualitatively differently (see later for how W_c is determined). For $W > W_c$, $\lambda_E(M)/M$ decays with M, corresponding to an insulator phase.[28,29] However, for $W < W_c$, it remains nearly independent of M, essentially forming a line of fixed points. This implies that $\lambda_E(M)$ grows linearly with M, therefore diverging in the thermodynamic limit. Consequently, the electron states are delocalized at these E's and W's. The independence of $\lambda_E(M)/M$ on M indicates that we are dealing with critical states characterized by a vanishing β scaling function,[28,29]. The extended states must have a topological origin, otherwise the system will belong to the trivial unitary class where all electron states are localized at positive W's. The consistency between the data in Figs. 1 and 2 is evident.

C. Phase diagram and computations of the spin-Chern numbers

To construct the phase diagram, we start from the limit $V_R = h_z = 0$, while holding V_{SO} at 0.1. In this limit, the spin up/down sectors decouple and the phase digram in the (E_F, W) plane, for each sector, consists of CI phases (with $C_\pm = \pm 1$) surrounded by a line of ex-

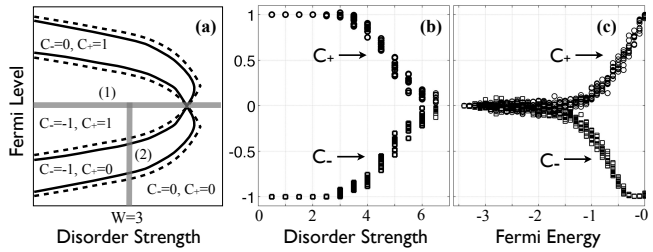

FIG. 3. (a) Phase diagram of a QSH system with a vertical Zeeman field. The phase boundaries (solid lines) harbor extended states. The dotted line shows the phase diagram when $V_R = 0$. (b-c) Computations of $C_\pm$ along the paths (1-2) of panel (a), respectively, completed for $h_z = 0.2$ on a 50×50 unit cells lattice, for 10 disorder configurations.

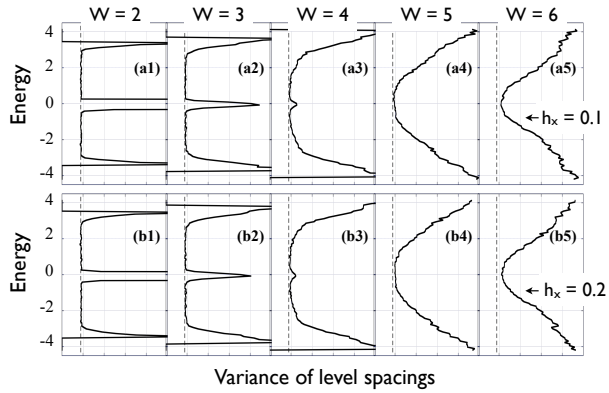

FIG. 4. Same as Fig. 1 but for horizontal Zeeman fields.

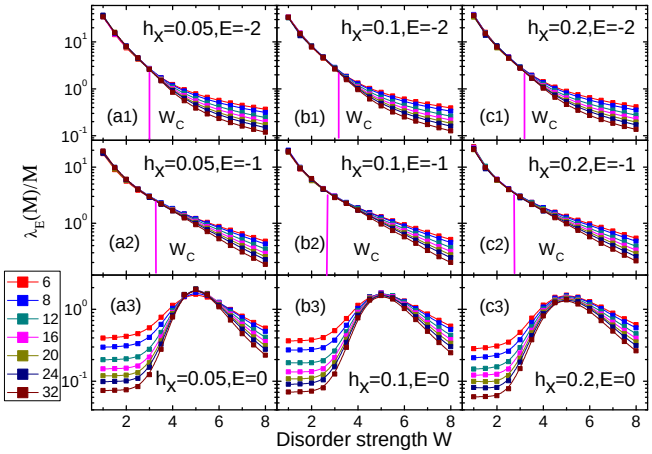

FIG. 5. Same as Fig. 2 but for horizontal Zeeman fields.

tended states and then by a trivial insulating phase[16,19]. TRS assures that the phase diagrams of the two sectors overlap perfectly. As in Fig. 3(a), when h_z is turned on, the phase diagrams shift revealing two regions where the total Chern number $C = C_- + C_+$ takes the values ± 1, surrounding a region characterized by $C_\pm = \pm 1$. The stability of the Chern numbers assures that the CI phases do not disappear when V_R is adiabatically turned on, but they could part away and open the phase diagram at large W's. This cannot happen here because the spin-gap stays open so the $C_\pm = \pm 1$ and $C_\pm = 0$ phases must be separated by extended states, a fact that can be seen explicitly in Fig. 2(a3-c3). We must conclude that the $C_\pm = \pm 1$ phase remains completely surrounded by the CI phases as illustrated in Fig. 3(a).

The calculated values of $C_\pm$ in Fig. 3(b-c) for $h_z = 0.2$, along the paths 1 and 2 in Fig. 3(a) further confirm our conclusions. Along path 1, both $C_\pm$ remain strictly quantized to ± 1 until they simultaneously start an abrupt descent to 0. $C_\pm$ are seen to cross the values ± 0.5 at about $W = 5$, exactly where the annihilation is observed in Figs. 1 and 2. Along path 2, C_- remains quantized at -1 well after C_+ started its descent to 0, indicating indeed the crossing of a region with total Chern number -1 (also confirmed by a direct calculation of C). This also provides direct evidence that the spin-gap remains open when crossing the boundary between $C_\pm = \pm 1$ and the CI phases, because its closing would have effected the quantization of both $C_\pm$ but in Fig. 3(c) C_- clearly remains quantized. A last remark here is that, since the CI phases are surrounded by $C = 0$ phases, they are necessarily surrounded by a line of extended states as illustrated in Fig. 3(a).

V. HORIZONTAL ZEEMAN FIELD

We now consider the horizontal Zeeman field. In Fig. 4 we have repeated the level statistics analysis discussed in the previous section, this time for $h_x = 0.1$ and 0.2. Examining Fig. 4, at smaller $W's$ we can again see regions where the variance stays very close to 0.178, and we still see these regions drifting towards each other. However, the variance is clearly seen to move away from 0.178 before the annihilation is complete (see panels a4-5 and b4-5). This is a first indication that the line of extended states is broken, which is also confirmed by the recursive Green's function analysis shown in Fig. 5. At $E = 0$ for $h_x = 0.1$ and 0.2, in panels (b3) and (c3) where the annihilation more or less takes place, we see $\lambda_E(M)/M$ decreasing with M, indicating a finite localization length. Nevertheless, both Figs. 4 and 5 indicate that extended states are still present at lower W's. This is also evident in Fig. 2 of Ref. 11. For the lower $h_x = 0.05$, the data clearly demonstrate that the extended states survive here all the way till the annihilation point.

The predicted phase diagram for large h_x is illustrated in Fig. 6(a), together with computations of $C_\pm$. To understand this diagram, it is useful to start from the one shown in Fig. 3(a) and consider a rotation of $\boldsymbol{h}$ from the z to the x-direction. During this rotation, we expect the phase diagram to open at large W's and the CI phases to separate and to continuously reduce size. Since there is an abrupt simultaneous change in the quantized

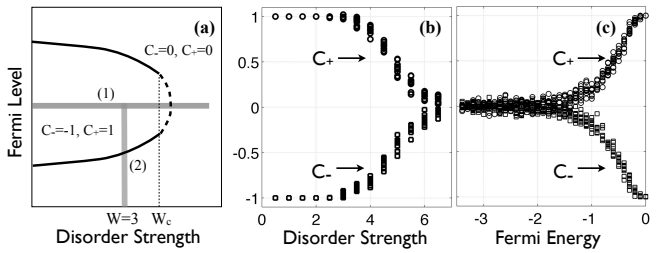

FIG. 6. Same as Fig. 3 but for a horizontal Zeeman field. The solid lines in panel (a) harbor extended states. The dotted line is where the spin-gap closes.

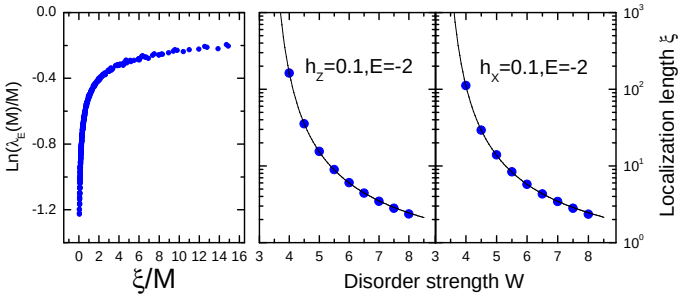

FIG. 7. Scaling analysis with data in Figs. 2 and 5 at $E = -1$ and $E = -2$ for different parameter sets. (a) Logarithmic plot of renormalized localization length on the insulating side as a function of ξ/M, where ξ is the localization length in the thermodynamic limit, as a scaling parameter. (b,c) Fitting of the localization length ξ (solid symbols) with the exponential function $\xi \propto e^{\alpha/\sqrt{W-W_C}}$ (solid lines) for $h_z = 0.1$ and $h_z = 0.1$, respectively, at $E = -2$. The fitting parameters are found to be (b) $W_c = 3.19$ and $\alpha = 6.38$, and (c) $W_c = 3.12$ and $\alpha = 6.25$.

$C_\pm$ along path 2, we conclude that all that is left when the rotation is completed is a doubly-degenerate line of extended states, which terminates abruptly at a critical W_c (which is possible if the line is doubly degenerate). For $W < W_c$, the mobility spin-gap stays open and the change in the quantized $C_\pm$ when moving along path 2 must be due to the closing of the mobility energy-gap. From Figs. 4 and 5, we know that the mobility energy-gap stays open when moving along path 1, so the change in the quantized $C_\pm$ seen in Fig. 6(b) must be due to the closing of the mobility spin-gap. Therefore, as illustrated in Fig. 6(a), there must exist a line in the phase diagram where the mobility spin-gap closes, and this line together with the line of extended states completely encircle the region where $C_\pm = \pm 1$. Our final note here is that the change in the quantized $C_\pm$ brings measurable physical effects even beyond W_c, where we can still see energy regions with very large localization lengths, levitating and annihilating each other, in stark contrast with what happens in trivial insulators,[16,19] where the extended spectrum simply fades away when the disorder is increased.

VI. SCALING ANALYSIS

According to the well-established one-parameter scaling theory,[28,29] the normalized localization length can be fitted with a universal scaling function $\lambda_E(M)/M = f(M/\xi)$, where ξ is the localization length in the insulator region or the correlation length in the metallic region, in the thermodynamic limit. Indeed, we find that all the data in the insulator phase shown in Figs. 2 and 5 at energy $E = -1$ and $E = -2$ can be well fitted with a universal function, as shown in Fig. 7(a). For example, the resulting localization length ξ for $h_z = 0.1$ and $h_x = 0.1$ at $E = -2$ is plotted as a function of disorder strength W in Figs. 7(b,c), respectively. It is found that the W dependence of ξ can be better fitted with an exponential function[35,36] $\xi \propto e^{\alpha/\sqrt{W-W_C}}$ than a power law function, with α and W_c as two fitting parameters. The critical disorder strength W_c determined in this way for different parameter sets is indicated in Figs. 2 and 5.

VII. SUMMARY

In summary, we used the spin-Chern numbers to predict the existence of robust bulk extended states in TIs with broken TRS and to map out the phase diagrams of the models. The predictions were well confirmed by the numerical computations of level statistics and localization lengths. Our study settles a long debated issue, namely, what protects the bulk extended states in TIs in the strong disorder regime. The answer is: topology alone. In contradistinction with the present common belief that the breaking of TRS in QSH systems will lead to the sudden and complete localization of the bulk states, we found that QSH systems with large Zeeman fields continue to display robust bulk extended states.

VIII. ACKNOWLEDGMENTS

This work is supported by the State Key Program for Basic Researches of China under Grant Nos. 2009CB929504 (LS), 2011CB922103 and 2010CB923400 (DYX), by the National Natural Science Foundation of China under Grant No. 10874066, 11074110 (LS), and 11174125 (DYX), and a project funded by the PAPD of Jiangsu Higher Education Institutions. This research was also supported by the U.S. NSF grants DMS-1066045 and DMR-1056168 (EP), and by the U.S. DOE Office of Basic Energy Sciences under Grant No. DE-FG02-06ER46305 (DNS).

Appendix: Derivation of Inequality (5)

The following is a detailed presentation of how the inequality (5):

$$|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle| \leq c_s \overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}} \lambda_{E_F}^s e^{-s|\boldsymbol{n}-\boldsymbol{m}|/4\lambda_{E_F}}. \tag{A.1}$$

has been obtained. The derivation presented here follows the lines used in Ref. 37 for the analysis of the Chern invariant. For us, the starting point is:

$$\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle = \tfrac{1}{2\pi}\oint_{\mathcal{C}}\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle d\zeta. \tag{A.2}$$

where we will later use the identity:

$$[\hat{\sigma}_z,(H-\zeta)^{-1}] = (H-\zeta)^{-1}[\hat{\sigma}_z,H](H-\zeta)^{-1}. \tag{A.3}$$

We are interested in the case when $[\hat{\sigma}_z,H]$ is small. Although the above formulas contain this factor explicitly, before we can say that $[\hat{\sigma}_z,P]$ is small we must resolve the following issue: the Green's function $(H-\zeta)^{-1}$ diverges when ζ crosses the dense localized spectrum at E_F. This can be resolved if we look at the averaged $[\hat{\sigma}_z,H]$.

We will use the following localization estimate[32]

$$\mathbb{E}\{|\langle \boldsymbol{n}|(H-E)^{-1}|\boldsymbol{m}\rangle|^s\} \leq c_s e^{-s|\boldsymbol{n}-\boldsymbol{m}|/\lambda_E} \tag{A.4}$$

where s is any number strictly between 0 and 1 and c_s is a generic constant. The constant c_s diverges when $s\to 1$, which reflects the known fact that the averaged Green's function diverges when E is inside the dense localized spectrum (this is not the case when the Green's function is taken to be a power s with $s<1$). We have successively:

$$\begin{aligned}&\tfrac{1}{2\pi}\oint_{\mathcal{C}}\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle d\zeta\\ =\ &\tfrac{1}{2\pi}\oint_{\mathcal{C}}\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle^{1-s/2}\\ &\times\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle^{s/2}d\zeta.\end{aligned} \tag{A.5}$$

For the first term in the integrand we will use the bound:

$$|\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle|^{1-s/2} < \frac{1}{|\mathrm{Im}\ \zeta|^{1-s/2}}. \tag{A.6}$$

The key here is that, although the right hand side diverges when ζ crosses the real axis, its integral is finite because of the power $1-s/2$. For the second term we use Eq. A.3:

$$\begin{aligned}&\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle^{s/2} =\\ &\langle \boldsymbol{n}|(H-\zeta)^{-1}[\hat{\sigma}_z,H](H-\zeta)^{-1}|\boldsymbol{m}\rangle^{s/2}\\ =\ &\Big[\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}\langle \boldsymbol{n}|(H-\zeta)^{-1}|\boldsymbol{k}\rangle\langle \boldsymbol{k}|[\hat{\sigma}_z,H]|\boldsymbol{k}+\boldsymbol{\delta}\rangle\\ &\times\langle \boldsymbol{k}+\boldsymbol{\delta}|(H-\zeta)^{-1}|\boldsymbol{m}\rangle\Big]^{s/2},\end{aligned} \tag{A.7}$$

where $\boldsymbol{k}+\boldsymbol{\delta}$ goes over the first nearest neighbors of $\boldsymbol{k}$ (thus $\boldsymbol{\delta}$ takes a finite number of values around the origin). Using Eq. A.4 and the fact that $|x_1+x_2+\ldots|^{s/2} \leq |x_1|^{s/2}+|x_2|^{s/2}+\ldots$, we can bound this term as:

$$\begin{aligned}&\big|\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle\big|^{s/2}\\ &\leq \overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}|\langle \boldsymbol{n}|(H-\zeta)^{-1}|\boldsymbol{k}\rangle|^{s/2}\\ &\times|\langle \boldsymbol{k}+\boldsymbol{\delta}|(H-\zeta)^{-1}|\boldsymbol{m}\rangle|^{s/2}\end{aligned} \tag{A.8}$$

At this point we arrived at:

$$\begin{aligned}&|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|\\ &\leq \tfrac{1}{2\pi}\oint_{\mathcal{C}}|\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle|^{1-s/2}\\ &\times|\langle \boldsymbol{n}|i[\hat{\sigma}_z,(H-\zeta)^{-1}]|\boldsymbol{m}\rangle|^{s}|d\zeta|\\ &\leq \overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\tfrac{1}{2\pi}\oint_{\mathcal{C}}\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}|\langle \boldsymbol{n}(H-\zeta)^{-1}|\boldsymbol{k}\rangle|^{s/2}\\ &\times|\langle \boldsymbol{k}+\boldsymbol{\delta}|(H-\zeta)^{-1}|\boldsymbol{m}\rangle|^{s/2}\tfrac{|d\zeta|}{|\mathrm{Im}\ \zeta|^{1-s/2}}.\end{aligned} \tag{A.9}$$

We now take the expected value on both sides and use $\mathbb{E}\{XY\} \leq [\mathbb{E}\{X^2\}\mathbb{E}\{Y^2\}]^{1/2}$:

$$\begin{aligned}&|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle| \leq \overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\tfrac{1}{2\pi}\oint_{\mathcal{C}}\tfrac{|d\zeta|}{|\mathrm{Im}\ \zeta|^{1-s/2}}\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}\\ &\big[\mathbb{E}\{|\langle \boldsymbol{n}(H-\zeta)^{-1}|\boldsymbol{k}\rangle|^s\}\mathbb{E}\{|\langle \boldsymbol{k}+\boldsymbol{\delta}|(H-\zeta)^{-1}|\boldsymbol{m}\rangle|^s\}\big]^{1/2}\end{aligned} \tag{A.10}$$

and use the localization estimate Eq. A.4:

$$\begin{aligned}&\ldots \leq c_s\overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\tfrac{1}{2\pi}\oint_{\mathcal{C}}\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}\\ &e^{-s(|\boldsymbol{n}-\boldsymbol{k}|+|\boldsymbol{k}+\boldsymbol{\delta}-\boldsymbol{m}|)/2\lambda_\zeta}\tfrac{|d\zeta|}{|\mathrm{Im}\ \zeta|^{1-s/2}}.\end{aligned} \tag{A.11}$$

The path $\mathcal{C}$ can be always chosen so that λ_ζ is maximum when ζ crosses the real axis at E_F, thus $\lambda_\zeta \leq \lambda_{E_F}$ so we can make the sum independent of ζ:

$$\begin{aligned}&\ldots \leq c_s\left(\tfrac{1}{2\pi}\oint_{\mathcal{C}}\tfrac{|d\zeta|}{|\mathrm{Im}\ \zeta|^{1-s/2}}\right)\overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\\ &\times\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}e^{-s(|\boldsymbol{n}-\boldsymbol{k}|+|\boldsymbol{k}+\boldsymbol{\delta}-\boldsymbol{m}|)/2\lambda_{E_F}}\\ &\leq c_s\left(\tfrac{1}{2\pi}\oint_{\mathcal{C}}\tfrac{|d\zeta|}{|\mathrm{Im}\ \zeta|^{1-s/2}}\right)\overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\\ &\times\textstyle\sum_{\boldsymbol{k},\boldsymbol{\delta}}e^{-s(|\boldsymbol{n}-\boldsymbol{k}|+|\boldsymbol{k}-\boldsymbol{m}|-|\boldsymbol{\delta}|)/2\lambda_{E_F}}\end{aligned} \tag{A.12}$$

and at this point we use

$$|\boldsymbol{n}-\boldsymbol{k}|+|\boldsymbol{k}-\boldsymbol{m}| \leq \tfrac{1}{2}(|\boldsymbol{n}-\boldsymbol{m}|+|2\boldsymbol{k}-\boldsymbol{n}-\boldsymbol{m}|) \tag{A.13}$$

to conclude

$$\begin{aligned}&\ldots \leq c_s\left(\tfrac{1}{2\pi}\oint_{\mathcal{C}}\tfrac{|d\zeta|}{|\mathrm{Im}\ \zeta|^{1-s/2}}\right)\left(\textstyle\sum_{\boldsymbol{\delta}}1\right)\overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\\ &\times e^{-s(|\boldsymbol{n}-\boldsymbol{m}|+1)/4\lambda_{E_F}}\textstyle\sum_{\boldsymbol{k}}e^{-s|2\boldsymbol{k}-\boldsymbol{n}-\boldsymbol{m}|/4\lambda_{E_F}}\end{aligned} \tag{A.14}$$

The last sum is independent of $\boldsymbol{n}$ and $\boldsymbol{m}$ and is proportional to $\lambda_{E_F}^2$. The round parentheses contain generic constants independent of λ_{E_F} so we can absorb them into c_s to finally obtain the desired bound:

$$\mathbb{E}\{|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|\} \leq c_s\overline{[\hat{\sigma}_z,H]}^{\frac{s}{2}}\lambda_{E_F}^s e^{-s(|\boldsymbol{n}-\boldsymbol{m}|+1)/4\lambda_{E_F}}, \tag{A.15}$$

where the extra 1 at the exponent becomes irrelevant in the asymptotic regime $|\boldsymbol{n}-\boldsymbol{m}| >> 1$.

Same bound can be used for fluctuations since:

$$\begin{aligned}&\mathbb{E}\{|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|^2\}-\mathbb{E}\{|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|\}^2\\ &\leq \mathbb{E}\{|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|^2\}\leq \mathbb{E}\{|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|\},\end{aligned} \tag{A.16}$$

where we used the fact that $|\langle \boldsymbol{n}|i[\hat{\sigma}_z,P]|\boldsymbol{m}\rangle|\leq 1$.

* shengli@nju.edu.cn
† prodan@yu.edu

[1] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[2] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[3] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[4] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).
[5] M. Koenig, S. Wiedmann, C. Bruene, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
[6] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).
[7] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature **452**, 970 (2008).
[8] E. Prodan, J. Phys. A: Math. Theor. **42**, 065207 (2009).
[9] E. Prodan, J. Math. Phys. **50**, 083517 (2009).
[10] D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, Phys. Rev. Lett. **97**, 036808 (2006).
[11] M. Onoda, Y. Avishai, and N. Nagaosa, Phys. Rev. Lett. **98**, 076802 (2007).
[12] A. M. Essin and J. E. Moore, Phys. Rev. B **76**, 165307 (2007).
[13] H. Obuse, A. Furusaki, S. Ryu, and C. Mudry, Phys. Rev. B **76**, 075301 (2007).
[14] H. Obuse, A. Furusaki, S. Ryu, and C. Mudry, Phys. Rev. B **78**, 115301 (2008).
[15] A. Yamakage, K. Nomura, K. I. Imura, and Y. Kuramoto, J. Phys. Soc. Jpn. **80**, 053703 (2011).
[16] E. Prodan, T. L. Hughes, and B. A. Bernevig, Phys. Rev. Lett. **105**, 115501 (2010).
[17] H. M. Guo, Phys. Rev. B **82**, 115122 (2010).
[18] E. Prodan, Phys. Rev. B **83**, 195119 (2011).
[19] E. Prodan, J. Phys. A: Math. Theor. **44**, 113001 (2011).
[20] T. A. Loring and M. B. Hastings, Europhys. Lett. **92**, 67004 (2010).
[21] M. B. Hastings and T. A. Loring, Ann. Phys. **326**, 1699 (2011).
[22] E. Prodan, Phys. Rev. B **80**, 125327 (2009).
[23] H. Li, L. Sheng, D. N. Sheng, and D. Y. Xing, Phys. Rev. B **82**, 165104 (2010).
[24] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New J. Phys. **12**, 043048 (2010).
[25] B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).
[26] Y. Yang, Z.Xu, L. Sheng, B. Wang, D. Y. Xing, and D. N. Sheng, Phys. Rev. Lett. **107**, 066602 (2011).
[27] M. L. Mehta, *Random Matrices*, 3rd ed. (Elsevier, London, 2004).
[28] A. MacKinnon and B. Kramer, Phys. Rev. Lett. **47**, 1546 (1981).
[29] A. MacKinnon and B. Kramer, Z. Phys. B **53**, 1 (1983).
[30] E. Prodan, New J. Phys. **12**, 065003 (2010).
[31] J. Bellissard, A. van Elst, and H. Schulz-Baldes, J. Math. Phys. **35**, 5373 (1994).
[32] M. Aizenman and S. Molchanov, Comm. Math. Phys. **157**, 245 (1993).
[33] K. B. Efetov, Zh. Eksp. Teor. Fiz. **83**, 833 (1982).
[34] K. B. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge University Press, Cambridge, UK, 1997).
[35] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).
[36] X. C. Xie, X. R. Wang, and D. Z. Liu, Phys. Rev. Lett. **80**, 3563 (1998).
[37] M. Aizenman and G. M. Graf, J. Phys. A: Math. Gen. **31**, 6783 (1998).
[38] T. Richter and H. Schulz-Baldes, J. Math. Phys. **42**, 3439 (2001).