

Topological phases in a two-dimensional lattice: Magnetic field versus spin-orbit coupling

W. Beugeling,¹ N. Goldman,² and C. Morais Smith¹

¹*Institute for Theoretical Physics, Utrecht University, Leuvenlaan 4, 3584 CE Utrecht, The Netherlands*

²*Center for Nonlinear Phenomena and Complex Systems,
Université Libre de Bruxelles (U.L.B.), B-1050 Brussels, Belgium*

(Dated: August 13, 2012)

In this work, we explore the rich variety of two-dimensional topological phases that arise when considering the competing effects of spin-orbit couplings, Zeeman splitting and uniform magnetic fields. We investigate minimal models, defined on a honeycomb lattice, which clarify the topological phases stemming from the intrinsic and Rashba spin-orbit couplings, and also from the Zeeman splitting. In this sense, our work provides an interesting path connecting the quantum Hall phases, generally produced by the uniform magnetic field, and the quantum spin Hall phases resulting from spin-dependent couplings. First, we analyze the properties of each coupling term individually and we point out their similarities and differences. Secondly, we investigate the subtle competitions that arise when these effects are combined. We finally explore the various possible experimental realizations of our model.

PACS numbers: 73.43.-f, 71.70.Ej, 37.10.Jk

I. INTRODUCTION

Today, unprecedented efforts are devoted to the study of novel topological insulating phases, which exhibit remarkable properties at their boundaries.^{1,2} In two dimensions, these quantum states are characterized by propagating edge states, which carry dissipationless currents along one-dimensional boundaries. From a theoretical point of view, these transport properties rely on the existence of large bulk gaps in the energy spectrum, which host robust and gapless edge excitations. Different edge state structures and transport properties can be realized according to the nature of the gaps: A magnetic field breaks time-reversal symmetry (TRS) and leads to chiral edge states (i.e., particles propagating in a given direction) and quantized Hall currents,³ whereas a large spin-orbit coupling preserves the TRS and produces helical edge structures.^{4,5} In the latter situation, referred to as the quantum spin Hall (QSH) effect, edge states with opposite spins counterpropagate and contribute to a vanishing Hall current, while producing a spin Hall current.^{6,7} The origins of the quantum Hall (QH) and QSH phases are fundamentally different, as the former is produced by an external field, whereas the latter relies on the intrinsic properties of the material. However, one can interpret the QSH phase as being two opposite QH phases, one for each spin component. This observation is easily shown through the Kane-Mele lattice model,^{4,5} which realizes the QSH effect by the inclusion of the intrinsic spin-orbit coupling (ISO), and corresponds to two copies of the (QH) Haldane model.⁸

In recent works, the ISO coupling has been combined with a magnetic field,^{9,10} or with a constant exchange term,¹¹ in order to study the effect of the TRS breaking on the QSH effect. It has been shown that the QSH effect persists in the presence of magnetic field, leading to the “TRS-broken”¹¹ or “weak” QSH effect.⁹ The term “weak” refers to the absence of robustness against spin-flip scattering due to magnetic disorder. In absence of this scattering process, the Hall and spin Hall conductivities are protected and quantized.

In this work, we present a different approach to study the

interplay between QH and QSH physics. Contrarily to the general trend in the field of topological insulators, which is to include many terms and study complicated Hamiltonians, here we investigate the minimal models that can produce topological phases. We consider a tight-binding model of the honeycomb lattice under a perpendicular magnetic field, which includes the ISO coupling, the Rashba spin-orbit (RSO) coupling, and the Zeeman effect. This spin-1/2 model was introduced in our recent paper,⁹ where we already studied the combined effects of some of these terms. Here, instead, we first investigate the effect of each of these terms individually on the QH phases generated by the magnetic field, in order to distill the problem. In particular, we demonstrate the equivalence of the ISO and Zeeman terms in generating the weak QSH phase at zero energy in the presence of a magnetic field. Then, we show that the ISO coupling can drive topological phase transitions between different topological insulating phases, an effect which totally relies on the combination of the ISO coupling and the magnetic field.

Secondly, we study the interplay between the terms incorporated in our model. For example, we discuss the competition between the ISO coupling and Zeeman effect, motivated by the similarity of these terms for generating the weak QSH effect. We show that the combination of these terms does not necessarily lead to a larger regime of parameters where the weak QSH effect is observed, which reveals the subtle competition between each of these terms and the magnetic field. We also study the exquisite effects produced by the Rashba coupling in the presence of a large exchange field, which is shown to generate several QH gaps, even in the absence of a magnetic field. From another perspective, we also discuss the fate of the weak QSH phase, as the RSO coupling is turned on, which introduces spin-flip terms into the Hamiltonian, so that the spins are no longer aligned perpendicularly to the plane, and the spin Hall conductivity is no longer protected.

An important motivation for the investigation of these topological phases and phase transitions is the rich variety of systems where they may be observed. Despite its weak ISO coupling, graphene remains an appealing candidate for the obser-

vation of these phases, e.g., by increasing the ISO coupling artificially by adatoms.¹² Secondly, several experimental groups have reported on synthetic honeycomb lattices in a condensed matter system, for instance arrays of quantum dots on a GaAs substrate,¹³ and “molecular graphene”, where the lattice is created by using the repulsive interactions of CO molecules deposited in a triangular array on a Cu(111) surface.¹⁴ These systems have larger lattice constants, which lead to a higher flux per plaquette at realistic magnetic field values, and allow for a more flexible control over the parameters than in real graphene. A third type of experiments that motivates our studies is the recent realization of a honeycomb optical lattice for ultracold atoms.^{15–17} In cold-atom systems, large magnetic fields¹⁸ and spin-orbit couplings¹⁹ are produced synthetically by adjustable lasers (see Ref. 20 for a review). In particular, such artificial gauge potentials could be produced in optical lattices,^{20–23} thus leading to the possibility to probe the high-flux regime. Furthermore, we note that optical *square* lattices subjected to well-designed gauge potentials can reproduce the properties of honeycomb lattice systems, exhibiting Dirac-type physics.^{24–27} Cold-atom systems carry the advantages of high tunability, necessary for studying phase transitions, and allow for the study of topological phases in absence of interparticle interactions²⁸ and disorder.²⁹ These possible experimental realizations are explored in detail in the final part of this work.

This article is structured as follows. We introduce the lattice model in Sec. II and present the four terms characterizing our tight-binding Hamiltonian: the usual hopping term, the two spin-orbit terms as well as the Zeeman splitting. Then, in Sec. III, we review the techniques used for our investigation and characterization of topological phases. Section IV discusses the combined effect of the magnetic field and spin-orbit coupling. Then, the combined effects of spin-orbit couplings, Zeeman splitting and uniform magnetic flux are explored in Sec. V. The possible experimental realizations of our model, as well as the possibility to detect the effects presented in Secs. IV and V, are reported in Sec. VI. Finally, the conclusions are drawn in Sec. VII. The Appendix provides further mathematical details of the calculations.

II. MODEL

We consider a tight-binding model of spinful electrons in a two-dimensional honeycomb lattice, subjected to a uniform perpendicular magnetic field $\mathbf{B} = B\mathbf{e}_z$. This model was introduced in Ref. 9 to investigate the effects of an external magnetic field on the Kane-Mele model (cf. Refs. 4 and 5). The magnetic field induces a Zeeman effect, and causes the hopping terms to acquire a phase. This phase is encoded by the so-called Peierls substitution, where one replaces the momentum \mathbf{p} by $\mathbf{p} - e\mathbf{A}$, where \mathbf{A} is the gauge potential associated with the magnetic field. Thus, any hopping term from site k to site j picks up the phase factor $e^{i\theta_{jk}}$, where $\theta_{jk} = (e/\hbar) \int_{\mathbf{r}_k}^{\mathbf{r}_j} \mathbf{A} \cdot d\mathbf{l}$. In the remainder of this text, we will express the magnetic field strength (flux density) in terms of the dimensionless quantity ϕ , defined as the flux per unit cell

of the lattice, expressed in units of the elementary flux quantum h/e .

In our model, we include four effects into the Hamiltonian, written as $H = H_{\text{NN}} + H_Z + H_I + H_R$, with

$$\begin{aligned} H_{\text{NN}} &= -t \sum_{\langle j,k \rangle} e^{i\theta_{jk}} c_j^\dagger c_k, \\ H_Z &= -2\pi\phi\lambda_Z \sum_j c_j^\dagger \sigma_z c_j, \\ H_I &= -it_I \sum_{\langle\langle j,k \rangle\rangle} e^{i\theta_{jk}} \nu_{jk} c_j^\dagger \sigma_z c_k, \\ H_R &= -it_R \sum_{\langle j,k \rangle} e^{i\theta_{jk}} c_j^\dagger (\boldsymbol{\sigma} \wedge \mathbf{d}_{jk}) c_k. \end{aligned} \quad (1)$$

Here, $c_k = (c_{k\uparrow}, c_{k\downarrow})$ is the annihilation operator for electrons at site k . The first term H_{NN} describes hopping between two nearest-neighbor (NN) sites j and k , where t denotes the hopping amplitude. The second term H_Z describes the Zeeman effect, which is an on-site term that assigns different potentials to the two spin components via the Pauli matrix σ_z . Its amplitude is proportional to the magnetic flux ϕ and the coefficient λ_Z , which is related to the material’s Landé g factor by $2\pi\phi\lambda_Z = g\mu_B B$, where μ_B is the Bohr magneton. Note that in a cold-atom emulation of this model, the parameters ϕ and λ_Z are tuned individually (cf. Sec. VI C). The Zeeman term lifts the degeneracy by shifting the two spin-degenerate copies of the spectrum up and down by an equal amount of energy. The third term H_I describes the ISO coupling, which corresponds to a next-nearest-neighbor (NNN) hopping with amplitude t_I .⁴ Here, the factor $\nu_{jk} = \pm 1$, where the sign depends on the value of $\mathbf{d}_{kl} \wedge \mathbf{d}_{lj}$, i.e., the outer product of the two bond vectors connecting site k to site j via their unique common neighbor at site l . This hopping also involves the matrix σ_z , so that the sign of the hopping amplitude is opposite for the two spin components \uparrow and \downarrow . Thus, the two different spin species are effectively subjected to opposite local (Haldane-type⁸) magnetic flux. In absence of a magnetic field, this term opens a topologically nontrivial gap, and causes the QSH effect, characterized by so-called helical edge states, which are protected by TRS.^{4,5} The final term H_R is the contribution to the NN hopping due to the RSO coupling. This hopping has an amplitude t_R , and involves the spin matrix $\boldsymbol{\sigma} \wedge \mathbf{d}_{jk} = \sigma_x d_{jk}^y - \sigma_y d_{jk}^x$, where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$, and \mathbf{d}_{jk} is the vector connecting sites k and j . The resulting spin matrix has only off-diagonal elements, so that this hopping involves a spin flip. This term couples the two spin components, and as a consequence, σ_z is no longer a good quantum number for $t_R \neq 0$. In practice, this means that spin states are no longer exclusively up or down, but may also point in different directions. Finally, we remark that the model considered here preserves inversion symmetry. In particular, we disregard the effect of an additional staggered potential, which was shown to induce topological phase transitions between trivial and non-trivial phases in Refs. 4 and 5. The possible physical realizations of our model (1) are discussed later in Sec. VI.

III. TOOLS FOR THE TOPOLOGICAL ANALYSIS

In the following sections, we will perform an analysis of the various topological phases produced by the terms in the model Hamiltonian (1). Before going into details, describing each term individually and their combined effects, we first present the framework and tools which we will use for this analysis. Readers who are already familiar with the notions of bulk and edge states, Chern numbers, and (spin) Hall conductivity may choose to skip this section.

A. Harper equation and the Hofstadter butterfly

We will illustrate our framework by studying the spin-degenerate model, i.e., we set $\lambda_Z = t_I = t_R = 0$ in Hamiltonian (1). We assume $t > 0$ throughout the whole text. The bulk band structure can be computed numerically by applying periodic boundary conditions, namely by considering a toroidal geometry. This requires to set $\phi = p/q$, where p and q are integers, in which case the system is periodic in both spatial directions. Under these conditions, and by choosing a proper gauge for the Peierls phases $e^{i\theta_{jk}}$, the system reduces to a $q \times 1$ magnetic unit cell. (See the Appendix for a description of the gauge structure and its spatial periodicity.) By applying a Fourier transform and invoking Bloch's theorem, the Schrödinger equation reduces to a $4q \times 4q$ eigenvalue problem, known as the discrete Harper equation,^{30,31}

$$(E/t)\Psi_n = \mathcal{D}_n\Psi_n + \mathcal{R}_n\Psi_{n+1} + \mathcal{R}_{n-1}^\dagger\Psi_{n-1}, \quad (2)$$

where $\Psi_n = (\psi_{nA\uparrow}, \psi_{nA\downarrow}, \psi_{nB\uparrow}, \psi_{nB\downarrow})$ denotes the single-particle wave function at site $n = 1, \dots, q$, and where the 4×4 matrices \mathcal{D}_n and \mathcal{R}_n are given in terms of the parameters defined in Hamiltonian (1) (see the Appendix for details). The four components of the wave function Ψ_n are due to the two sublattice (A, B) and the two spin (\uparrow, \downarrow) degrees of freedom. The $4q \times 4q$ Harper problem in Eq. (2) is analogous to the original (spinless, square-lattice) Hofstadter problem,³⁰ where the Harper equation involves a $q \times q$ matrix.

Solving the Harper equation (2) provides the bulk energy bands $E = E(k_x, k_y)$. It is instructive to evaluate the sizes of the bulk gaps as a function of the magnetic flux ϕ , which results in a diagram known as the Rammal-Hofstadter butterfly,^{30,32} shown in Fig. 1. This figure may be obtained in the limit of large q , where the bands become dispersionless. The fractal structure, which is periodic in ϕ with a period of 1, is a result of the competition between two length scales: The magnetic length $l_B = \sqrt{\hbar/eB}$ and the lattice spacing. Here, ϕ parametrizes this competition, as the ratio between the area of the unit cell and that of the cyclotron orbits ($\sim l_B^2$).

B. Topological invariants

Each bulk gap in the spectrum shown in Fig. 1 may be characterized by a topological invariant, which encodes the Hall conductivity of this system when the Fermi energy lies inside

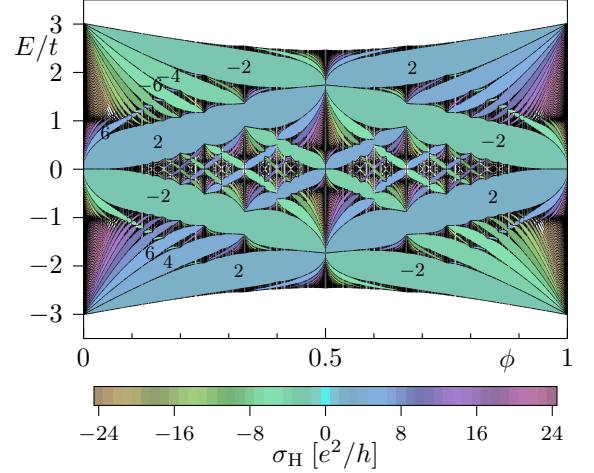


FIG. 1. (Color online) Hofstadter butterfly spectrum for the spin-degenerate model, Hamiltonian (1) with $\lambda_Z = t_I = t_R = 0$. The bulk bands, i.e., the regions where the density of bulk states is positive, are shown in black. In the bulk gaps, the colors and numbers indicate the quantized Hall conductivity in units of e^2/h .

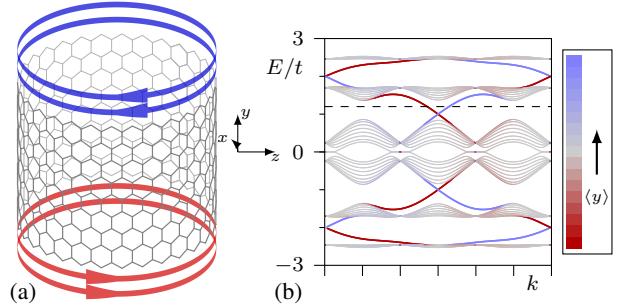


FIG. 2. (Color online) (a) The cylinder geometry representing a ribbon with periodic boundary conditions in the x direction. (b) Spectrum at $\phi = 1/3$ for the spin degenerate model ($\lambda_Z = t_I = t_R = 0$). The colors represent the expectation value $\langle y \rangle$: Red and blue indicate edge states at the bottom and top edge, respectively, and gray represents bulk states. The spin-degenerate edge states shown in (a) correspond to those in the QH gap at the Fermi energy indicated by the dashed line in (b).

this gap. The integer values of the Hall conductivity are robust against external perturbations: They remain constant as long as the bulk gaps remain open. In the following, we recall the relation between this transport coefficient and the concept of edge states and topological invariants.

1. Edge-state analysis

The edge-state analysis can be performed for a system featuring a boundary,³³ such as a cylinder, as shown in Fig. 2(a). When solving the Harper equation (2) on a cylinder, namely by applying periodic boundary conditions along one spatial

direction only, the spectrum consists of *bulk bands* and *topological edge states*.³³ Indeed, we typically find a few *edge states* within the bulk gaps, some of which *cross* the gap from one bulk band to the other; see Fig. 2(b). Importantly, each edge state contributes e^2/h (one quantum of conductivity) to the Hall conductivity of the system. The total Hall conductivity is therefore equal to an integer number (i.e., the number of edge-state branches) times e^2/h . The topological invariance of the Hall conductivity is due to the fact that the number of edge states inside the bulk gap cannot be altered unless the gap is closed. This result can be used to derive the topological invariant associated to any of the bulk gaps, leading to the colors in the spectrum in Fig. 1. For this spin-degenerate model, all topological invariants are multiples of 2 because the system is twofold spin degenerate.

More precisely, in order to evaluate the Hall conductivity of a bulk gap, we count the edge states whose dispersions intersect the Fermi energy, taking into account their location and direction of propagation. The location of each state is derived by computing its expectation value $\langle y \rangle$ from the eigenstate, which provides the coloring of Fig. 2(b). Secondly, the direction of propagation can be obtained from the sign of its momentum derivative dE/dk , where k is the momentum parallel to the edge. Edge states with opposite directions contribute with opposite signs to the Hall current and therefore to the Hall conductivity,

$$\sigma_H = (N_R - N_L) \frac{e^2}{h}, \quad (3)$$

where N_L and N_R denote the number of left- and right-moving states, respectively. The Hall conductivity does not depend on the magnitude of the velocity ($\propto dE/dk$), but instead is given by the number of edge-state channels.³⁴ In this spin-degenerate model, the states are all either left- or right-moving, i.e., they are *chiral* states. The chirality of the edge modes provides protection of the Hall conductivity against backscattering due to disorder in the presence of broken TRS.

2. Bulk topological invariants

Remarkably, the number of edge states is directly related to topological invariants, the *Chern numbers*, a quantity associated with each of the bulk bands. The Chern number is an integer topological index defined in the toroidal geometry,^{35,36}

$$C_n = \frac{1}{2\pi i} \int_{BZ} \left(\frac{\partial \mathbf{A}_{n,\mathbf{k}}^y}{\partial k_x} - \frac{\partial \mathbf{A}_{n,\mathbf{k}}^x}{\partial k_y} \right) d\mathbf{k}, \quad (4)$$

where $\mathbf{A}_{n,\mathbf{k}} = \langle u_{n,\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n,\mathbf{k}} \rangle$ is the Berry connection associated to the eigenstate $|u_{n,\mathbf{k}}\rangle$, and where BZ denotes the (toroidal) Brillouin zone. The number of edge states N_r inside the r th bulk gap, or equivalently the Hall conductivity assigned to this gap (cf. discussion above), is then equal to the sum of all the Chern numbers associated to the occupied bands,^{1,37}

$$N_r = \sum_{n \in \text{occupied bands}} C_n. \quad (5)$$

This identity is known as the *bulk-boundary correspondence*,^{33,36} as it relates the topological indices associated to the bulk bands to the number of edge excitations. This correspondence provides us with an important observation: Although the dispersions of the edge states depend on the shape of the edges (e.g., zigzag, bearded, or armchair) and on the system size, the number of edge-state branches does not: this number is a bulk property and therefore does not depend on the specific form of the edge. In the remainder of this paper, we will study systems with zigzag and bearded edges, without any loss of generality concerning the topological properties.

In the next sections, we will characterize the topological phases by their associated Hall conductivity and edge-state structures. As described in the following, the edge states produced by the magnetic field are affected in the presence of strong perturbing terms, such as the spin orbit couplings and Zeeman splitting. For the sake of simplicity, we will show these effects using the edge-state analysis obtained from a cylindrical geometry, rather than focusing on the analysis of the bulk topological invariants.

3. The Středa formula

An efficient method to compute the Hall conductivity can be formulated in terms of the integrated density of states. Inside a bulk gap, the integrated density of states $N_{\phi}(E)$ is defined as the fraction of all states that lie below the gap (which is an integer times $1/4q$). The Hall conductivity is then proportional to the ϕ -derivative of this quantity,

$$\sigma_H = 4 \frac{e^2}{h} \frac{\partial N}{\partial \phi}. \quad (6)$$

a result known as the Středa formula.^{38–40} We have used this powerful formula to compute the Hall conductivity, and generate the colorful butterfly spectra presented in the remainder of this paper.

4. Charge and spin Hall conductivities

In the presence of terms involving the matrix σ_z in the Hamiltonian, the spin SU(2) symmetry is broken. If spin-flip terms (i.e., linear combinations of σ_x and σ_y) are absent, the spin in the z direction is conserved, i.e., it is a good quantum number. In this case, which happens if the ISO coupling or the Zeeman splitting is present while the RSO is not, the Hamiltonian may be decomposed into the two spin components (\uparrow, \downarrow). Thus, one is able to define the component-wise Chern numbers $C_{\uparrow,n}$ and $C_{\downarrow,n}$ and Hall conductivities $\sigma_{H\uparrow}$ and $\sigma_{H\downarrow}$. The charge Hall conductivity of the whole system is the sum of the two component-wise conductivities, $\sigma_H = \sigma_{H\uparrow} + \sigma_{H\downarrow}$, since both of them contribute equally to the Hall current. Considering their contributions to the spin current, spin-up and spin-down edge modes have opposite “spin charges” $\pm \hbar/2$. Thus, the spin Hall conductivity is equal

to the difference of the two component-wise conductivities, $\sigma_H^{sp}/(e/4\pi) = (\sigma_{H\uparrow} - \sigma_{H\downarrow})/(e^2/h)$, where $e/4\pi$ is the elementary quantum of spin Hall conductivity.

In the spin-degenerate model discussed earlier, the two component-wise Hall conductivities are always equal, so that all spin Hall conductivities vanish. This is certainly different if the spin degeneracy is broken. For instance, in the presence of ISO, the QSH gap at zero magnetic field and at zero energy⁵ is characterized by $\sigma_{H\uparrow} = -\sigma_{H\downarrow} = -e^2/h$. Here, the contributions to the charge Hall conductivity cancel, but the contributions to the spin Hall conductivity add up to $\sigma_H^{sp} = -2e/4\pi$.

If the Hamiltonian contains spin-flip terms, e.g., the Rashba coupling, then the spin is not conserved, and the aforementioned definition of the spin Hall conductivity is no longer valid. However, the related spin Chern numbers, which equal $C_n^{sp} = C_{\uparrow,n} - C_{\downarrow,n}$ if spin is conserved, remain well-defined topological invariants, even in absence of spin conservation.^{41,42}

IV. EFFECTS OF THE MAGNETIC FIELD

A. The spin-degenerate model: Landau Levels and the anomalous quantum Hall effect

Zooming in on the butterfly of Fig. 1 near $\phi \approx 0$, we observe that the thickness of the bands decreases in the low-flux regime (not shown). Thus, in this limit we can treat the bands as being infinitely thin, i.e., as Landau levels. In the low-energy regime, between the Van Hove singularities at $E/t = \pm 1$, the band structure at low flux coincides with the Landau-level spectrum known from graphene,^{43,44} with energies $E \propto \pm\sqrt{2\pi l\phi}$ (where $l = 0, 1, 2, \dots$ is the Landau-level index).

Moreover, computing the Hall conductivity in this region, we find the “anomalous” Hall-conductivity sequence $\sigma_H = 4(n + 1/2)e^2/h$ ($n \in \mathbb{Z}$), as observed in graphene.⁴⁵ The breakdown of this structure around $E/t = \pm 1$ can be understood from the large density of states near the Van Hove singularities and from the Chern numbers associated to these many bands.⁴⁶ The edge-state structures, Hall conductivity plateaus and topological aspects of spinless electrons in a honeycomb lattice subjected to a magnetic field have been thoroughly described in Ref. 46.

B. Intrinsic spin-orbit coupling

Since the seminal work of Kane and Mele,⁴ it is known that the ISO coupling opens a topologically nontrivial gap of size $\Delta = 6\sqrt{3}t_1$ at zero magnetic field. This bulk gap hosts two counterpropagating edge modes per edge, with opposite spins (one Kramers pair): These *helical* edge states are related by TRS. This topological phase is known as the QSH state, and may be regarded as two opposite QH phases (i.e., each spin performs the QH effect, with opposite chirality). The system preserves TRS (due to absence of the magnetic field), which protects the QSH state against scattering processes caused by

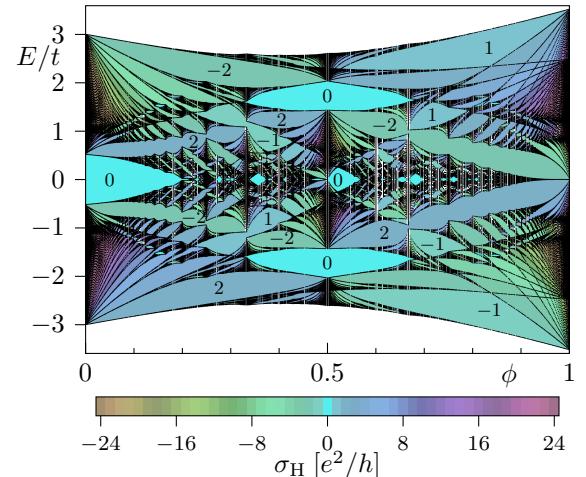


FIG. 3. (Color online) Hofstadter butterfly spectrum for the system with ISO coupling ($t_1/t = 0.1$). Although the periodicity in ϕ is 6 rather than 1 (see the Appendix), we present only the range $0 \leq \phi \leq 1$ for the sake of comparison with the other terms. The colors and numbers indicate the quantized Hall conductivity in units of e^2/h .

disorder. Although for any edge state a reversely propagating mode would be available to scatter to, the spin has to be flipped, and the two different ways to flip the spin interfere destructively with each other, due to a mutual phase factor of $e^{i\pi} = -1$.⁴⁷

The combination of the ISO coupling and a magnetic field leads to the breaking of the TRS satisfied by the QSH phase. However, due to the absence of spin flip terms in the Hamiltonian, the helical edge-state structure persists. In terms of charge and spin Hall conductivity, the resulting state is equivalent to the QSH state. As announced in the introduction, this state is called the *weak QSH* phase⁹ or *TRS broken QSH* phase.¹¹ From the Hofstadter butterfly in Fig. 3, we observe that the size of the weak QSH gap at zero energy tends to decrease if the flux is increased (but not monotonically), and closes eventually. In Fig. 3, in which $t_1 = 0.1t$, the QSH gap at $E = 0$ remains open in the range $0 \leq \phi \leq 0.2$. Here we observe the competition between the ISO coupling, which opens the (weak) QSH gap, and the magnetic field, which “tries” to destroy the weak QSH state by its TRS breaking property. Throughout the weak QSH gap, the edge states remain crossing at zero energy, which indicates that this phase is robust at least in the absence of magnetic disorder. We finally note that the spectrum remains particle-hole symmetric, as in the spinless case.

C. Zeeman effect

The Zeeman term models an on-site spin-splitting effect that is always present in real materials subjected to magnetic fields. Studying Hamiltonian (1) with $\lambda_Z \neq 0$ and $t_R = t_I = 0$, we observe that the spin degeneracy is lifted by merely shifting the two spin states up and down in energy

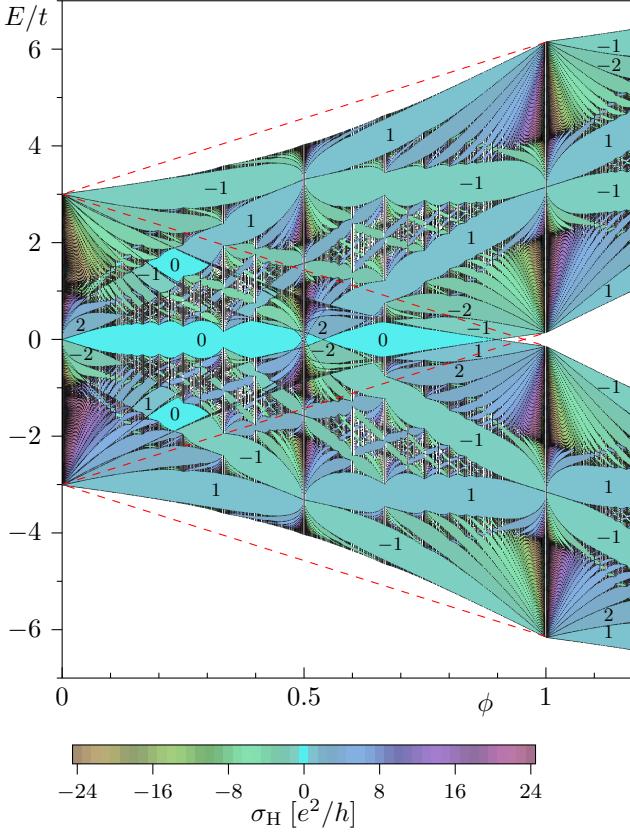


FIG. 4. (Color online) Hofstadter butterfly spectrum for the model described by Hamiltonian (1) with Zeeman effect, $\lambda_Z/t = 0.5$ and $t_I = t_R = 0$. The bulk bands are in black. The colors and numbers indicate the quantized Hall conductivity in units of e^2/h . The diagonal dashed lines indicate the outline of the butterfly (i.e., $E/t = \pm 3 \pm 2\pi\phi\lambda_Z/t$) for the two spin components.

by $2\pi\lambda_Z\phi$. The spectrum is otherwise left invariant. With respect to the Hofstadter butterfly represented in Fig. 1, this means that the spin-up and spin-down copies are “skewed” in opposite directions, as shown in Fig. 4. Because the lowest and highest bands of the spin-degenerate model are situated at approximately $E/t = \pm 3$, the two spin components are completely separated above a certain flux value, approximately equal to $\phi = 3/(2\pi\lambda_Z/t)$. Furthermore, the Zeeman shift breaks the periodicity of the butterfly spectrum along the ϕ axis.

1. The Zeeman-induced weak QSH phase

As we already mentioned, the spectrum of the spin-degenerate model shows Landau levels for low values of the flux (cf. Sec. IV A). The Hall conductivities in the gaps are $\sigma_H/(e^2/h) = \dots, -6, -2, 2, 6, \dots$. If we concentrate on the bulk gaps around zero energy, we observe one and then three doubly degenerate edge states (see Fig. 5(a)). When the Zeeman splitting is nonzero, the two copies of these states shift

up and down, as shown in Fig. 5(b) for $\phi = 1/61$. Assuming that the splitting is sufficiently large, a new gap opens at zero energy, where the original band was situated. The spin-up and spin-down edge states now connect to bulk bands that have been shifted up and down, respectively. As a consequence, in the newly opened gap at zero energy, we have $N_\uparrow = -1$ and $N_\downarrow = +1$. Thus, the charge Hall conductivity vanishes, and the spin Hall conductivity is equal to $\sigma_H^{\text{sp}} = -2e/4\pi$. Based on these values, we find that this state is a weak QSH phase.

Comparing this state to the one generated by ISO (see Fig. 5(c)), we observe the similarity between them. In both cases, the Landau levels for the two spin components shift in opposite directions. The difference between the two terms lies in the different values of the shift: The Zeeman shift $\pm 2\pi\phi\lambda_Z$ is linear in the flux, so that it vanishes for $\phi \rightarrow 0$, while the Landau-level shift induced by the ISO coupling equals $\pm 3\sqrt{3}t_I$.⁴ A recent study shows that a flux-independent Zeeman term (called an exchange term) may also generate the QSH state at zero magnetic field.¹¹

The similarity between the Zeeman and ISO terms may be understood at a formal level, in terms of the respective linearized (low-energy) Hamiltonians: The linearized Zeeman term acts as $\Psi^\dagger (\sigma_z^{\uparrow\downarrow} \otimes \mathbb{1}^{AB} \otimes \mathbb{1}^{KK'}) \Psi$, where the factors indicate the proper spin, sublattice pseudospin, and valley pseudospin, respectively. On the other hand, the ISO coupling acts as $\Psi^\dagger (\sigma_z^{\uparrow\downarrow} \otimes \sigma_z^{AB} \otimes \sigma_z^{KK'}) \Psi$. *A priori*, these Hamiltonians act differently on the eight-component field vector $\Psi = (\psi_{\sigma,\tau,\xi})_{\sigma=\uparrow\downarrow, \tau=AB, \xi=KK'}$. In the lowest Landau level, the four components $\psi_{\uparrow,A,K}, \psi_{\downarrow,A,K}, \psi_{\uparrow,B,K'}, \psi_{\downarrow,B,K'}$ vanish.⁴⁴ Substituting the remaining field components into the linearized Hamiltonians yields equality up to a sign,

$$\begin{aligned} & \Psi^\dagger \left(\sigma_z^{\uparrow\downarrow} \otimes \mathbb{1}^{AB} \otimes \mathbb{1}^{KK'} \right) \Psi \\ &= \psi_{\uparrow,A,K}^\dagger \psi_{\uparrow,A,K'} - \psi_{\downarrow,A,K}^\dagger \psi_{\downarrow,A,K'} \\ &+ \psi_{\uparrow,B,K}^\dagger \psi_{\uparrow,B,K} - \psi_{\downarrow,B,K}^\dagger \psi_{\downarrow,B,K} \\ &= -\Psi^\dagger \left(\sigma_z^{\uparrow\downarrow} \otimes \sigma_z^{AB} \otimes \sigma_z^{KK'} \right) \Psi, \end{aligned} \quad (7)$$

which shows that the Zeeman effect and the ISO coupling act equivalently on the lowest Landau level. We remark that this reasoning is only valid at this specific Landau level: For higher Landau levels, all the eight fields are present, and the two terms become inequivalent.

2. The spin-imbalanced quantum Hall phases

So far, most of the studies on the QSH effect have generally concentrated on the behavior of the system at zero energy. However, very interesting features also emerge at nonzero energies in the presence of external fields. For instance, in the low-flux regime, a Zeeman gap is formed at the $n = 1$ Landau level (at $E/t \approx 0.42$), as shown in Fig. 5(b). Similarly to the zero-energy gap, this gap also shows a difference between the number of edge states with spin up and spin down components, $N_\uparrow = 1$ and $N_\downarrow = 3$, which may be understood from the values above and below the corresponding Landau

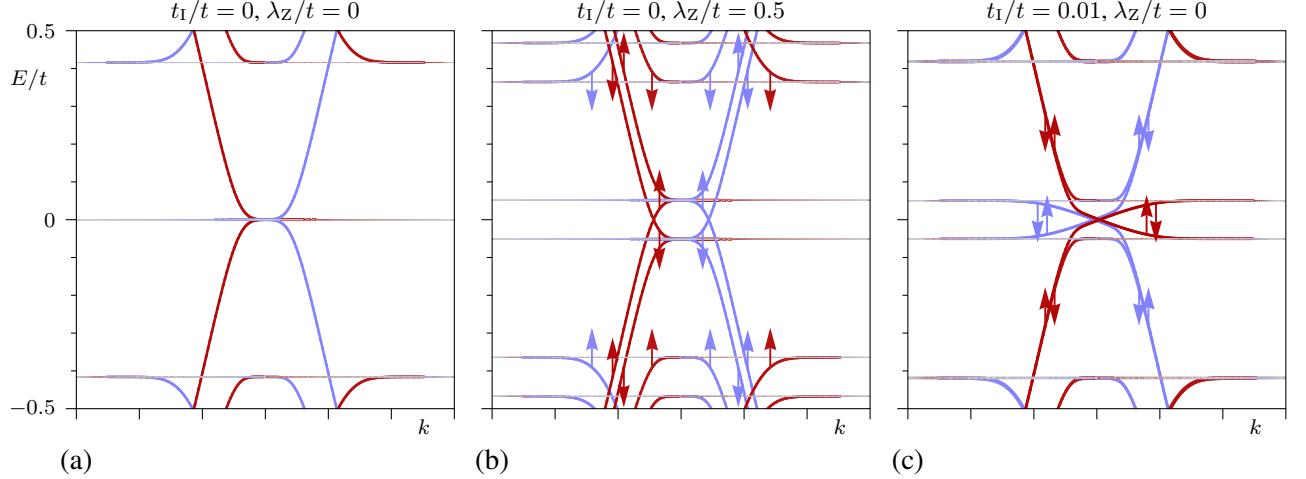


FIG. 5. (Color online) Comparison of dispersions at $\phi = 1/61$ for (a) the spin-degenerate case, (b) the system with Zeeman effect ($\lambda_Z/t = 0.5$) and (c) with ISO ($t_I/t = 0.01$). The coupling constants λ_Z and t_I have been chosen such that the (weak-QSH) gap sizes at zero energy become equal.

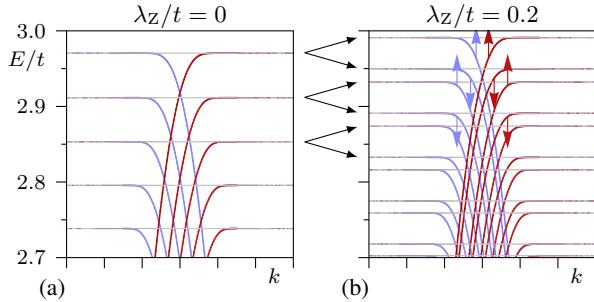


FIG. 6. (Color online) High-energy spectra for the model without spin-orbit couplings ($t_R = t_I = 0$) at $\phi = 1/61$. (a) spectrum for the spin-degenerate model ($\lambda_Z = 0$), with quantum Hall gaps corresponding to $\sigma_H = 2, 4, 6, \dots$ (in units of the conductivity quantum). (b) Spectrum with Zeeman effect ($\lambda_Z = 0.2t$), showing the spin-filtered quantum Hall state with $N_↑ = 1$ and $N_↓ = 0$. The other gaps show (alternatingly) quantum Hall phases ($N_↑ = N_↓$) and spin-imbalanced quantum Hall phases ($|N_↑| = |N_↓| + 1$).

level of the spinless model. Again, the spin Hall conductivity is nonzero, $\sigma_H^{sp} = -2e/4\pi$, but the difference with the weak quantum spin Hall gap is that the charge Hall conductivity is also nonzero, $\sigma_H = 4e^2/h$, as can be deduced from Fig. 5(b). Furthermore, the edge states in this gap all propagate in the same direction (i.e., they are chiral), thus providing robustness against disorder. Here, we refer to this phase as the *spin-imbalanced quantum Hall phase*.⁹

A special instance of the spin-imbalanced quantum Hall phase may occur if edge states of one of the spin-components are absent, i.e., either $N_↑ = 0$ or $N_↓ = 0$, in which case one speaks about the *spin-filtered quantum Hall phase*. In this state, the magnitudes of the charge Hall and spin Hall currents are therefore equal when expressed in units of their re-

spective conductance quanta. The spin-filtered quantum Hall phase appears at high energies and low flux in the presence of Zeeman coupling, as we now explain. In the absence of Zeeman coupling, we observe equally-spaced Landau levels, each of which has Chern number -2 , see Fig. 6(a). If the Zeeman term is present, the gap that forms between the two copies of the original highest-energy Landau level exhibits a spin-filtered quantum Hall phase, characterized by the presence of edge states of only one spin component; in this case $N_↑ = 1$ and $N_↓ = 0$, see Fig. 6(b). The other gaps visible in Fig. 6(b) are spin-imbalanced and ordinary spin-degenerate quantum Hall gaps, alternatingly. Let us mention that the spin-filtered and spin-imbalanced quantum Hall phases are ubiquitous in systems with Zeeman-split Landau levels, e.g., the quantum Hall plateaus corresponding to odd-integer filling factors in GaAs/AlGaAs heterostructures,⁴⁸ or in HgTe quantum wells, which have strong Zeeman effect and are thus ideal candidates for observation of these phases.⁴⁹

D. Rashba spin-orbit coupling

The RSO coupling differs in an essential way from the ISO coupling and the Zeeman effect, in the sense that the hopping involves a spin flip. In other words, the spin matrices involved in the Rashba hopping are off-diagonal (i.e., σ_x and σ_y), whereas the other terms in the Hamiltonian involve diagonal spin matrices. This has a profound effect on the spin structure of the system: If the RSO interaction is nonzero, mixing occurs between the spin components and we find that the spin orientation of the eigenstates is generally site-dependent. The spin direction is confined to the two-dimensional plane and is perpendicular to the momentum: Aside from the out-of-plane component (z -direction), the component perpendicular to the cylinder edge may also be nonzero, while the component par-

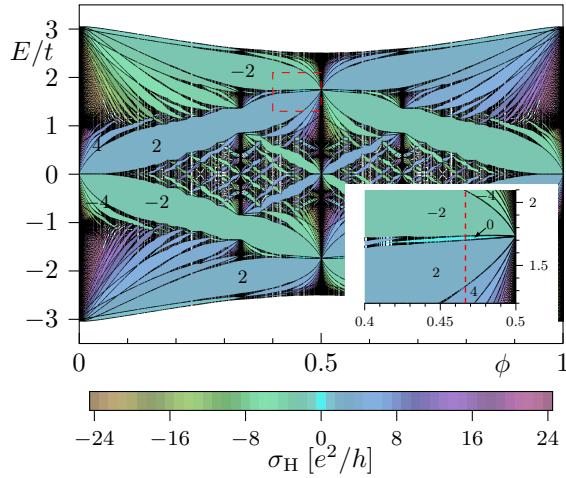


FIG. 7. (Color online) Hofstadter butterfly spectrum (with RSO coupling, $t_R/t = 0.2$). The inset shows a magnification of the region surrounded by the dashed lines. In the inset, the dashed line indicates the flux value $\phi = 7/15$ discussed in the text. The colors and numbers indicate the quantized Hall conductivity in units of e^2/h .

allel to the edge always vanishes. In other words, the helicity $(\sigma_x, \sigma_y, \sigma_z) \cdot \hbar \mathbf{k}$ of the (edge) states remains zero. This effect is caused by the nature of the Rashba coupling in the Hamiltonian (1), which involves the outer product of the spin vector σ with the bond vectors \mathbf{d}_{jk} , so that the spin is always perpendicular to the hopping direction. Because the edge-state spins are no longer in the up or down state, the Chern numbers for the respective spin components are no longer well-defined. Thus, the spin conductivity no longer takes quantized values. Nevertheless, the spin Chern number remains a well-defined integer-valued topological invariant,⁴¹ which can be used as a tool to distinguish between trivial and nontrivial gaps.

The second difference with the ISO coupling is that the zero-flux spectrum remains gapless and the spin degeneracy is lifted by the RSO coupling (considering $t_R \neq 0$ and $t_I = \lambda_Z = 0$). Indeed, upon setting $t_R \neq 0$, the two Dirac cones from the spinless model (at the special points K and K') are broken into four cones each: A central isotropic cone at K or K' , and three anisotropic “satellite” cones around it. This effect is known as trigonal warping.^{50,51} Between the main cones and the satellite ones, there is a Van Hove singularity at low energy, which scales as $\pm(t_R/t)^3$ for small t_R/t . The Berry phases associated to the main and satellite cones are $-\pi$ and π , respectively, which add up to the Berry phase 2π of the Dirac cones in the spin-degenerate model (π for each spin component).⁵² For energies $|E/t| < (t_R/t)^3$, the low-flux Landau level spectrum is characterised by two different sets of Landau levels with Chern numbers 2 and 6. Outside this regime, there are only twofold-degenerate Landau levels, because the satellite cones can no longer be resolved. Around $|E/t| \approx (t_R/t)^3$ there is a crossover regime where the sixfold-degenerate bands split into three, each one being twofold degenerate.

Adding a nonzero Rashba term to the spin-degenerate

model discussed in Sec. III A will cause a spin-splitting of some bulk bands in the Hofstadter butterfly, see Fig. 7. In particular, at low energies, the Rashba term splits the fourfold-degenerate Landau levels into twofold-degenerate Landau levels. In the limit $\phi \rightarrow 0$, we also note that the spectrum indeed remains gapless and particle-hole symmetric. However, this is not true for other Dirac regimes in the spectrum. For instance, at $\phi = 7/15$ and $E/t \approx 1.7$ (see the inset of Fig. 7), the RSO coupling splits the central Landau level, and the resulting gap exhibits two counterpropagating edge modes on each edge. However, the spin Hall conductivity is undefined because the spin direction varies as a function of the Fermi energy. Nevertheless, at each energy inside the gap the spins of the two edge states are approximately equal, so that the spin current nearly vanishes. Besides, the charge Hall conductivity is exactly zero when the Fermi energy is located in this gap.

The RSO coupling has drastic effects on the edge states, as their spins no longer align perpendicularly to the sample plane, but get an additional in-plane component perpendicular to the edge. Besides, the direction of the edge-state spins depends on the Fermi energy. This observation distinguishes the Rashba effect from the Zeeman effect in a tilted field, which puts all edge-state spins in the same direction. In addition, the in-plane component is opposite for edge states at the opposite edges, whereas the perpendicular component is the same. Importantly, due to the dependence of the spin direction on the Fermi energy, the latter can be conveniently used for spin manipulations (in addition to the amplitude of the Rashba coupling itself).

We illustrate this phenomenon around the Rashba-split bulk band at $\phi = 7/15$ and $E/t \approx 1.7$. In Fig. 8(a), we show that in the limit of $t_R \rightarrow 0$, the eigenstates are in-plane (indicated as “left” and “right” in the figure, meaning that the spin direction is $\pm \hat{y}$). This result shows that for infinitesimal RSO coupling, it is more natural to decompose the states in terms of the eigenstates of the Pauli matrix σ_y than those of σ_z . If the Rashba coupling is then increased, the gap opens, and in the vicinity of the bulk bands, the edge states tend to rotate to the vertical direction, see Fig. 8(b). In addition, we observe that not all spins have the same length. This phenomenon occurs because the spins displayed in this plot represent the expectation values of the spin components (see the Appendix). Since the spin direction depends on the lattice position (y coordinate), the length of this expectation value may be less than unity.

We expect that the canting of the spins is most easily observed in gaps with a single edge state on each edge. For instance, in Ref. 9, it has been demonstrated that the RSO coupling cant the spin in a spin-filtered gap generated by the Zeeman effect. In this setting, the spin textures may be controlled at will by tuning the coupling parameters and the Fermi energy.

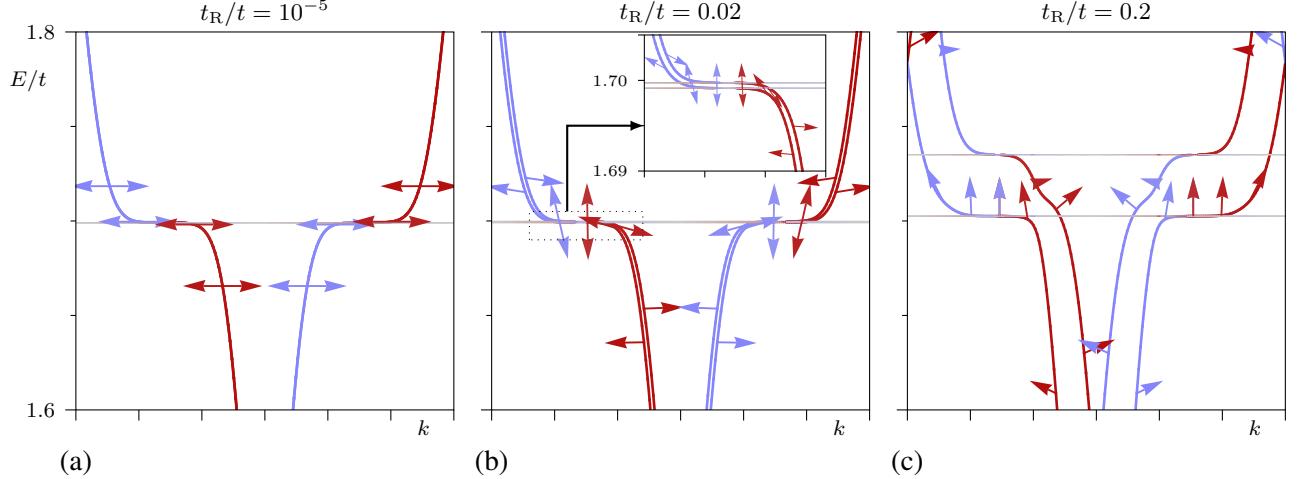


FIG. 8. (Color online) Dispersions at $\phi = 7/15$, showing the rotation of spin directions as the Rashba coupling t_R is increased.

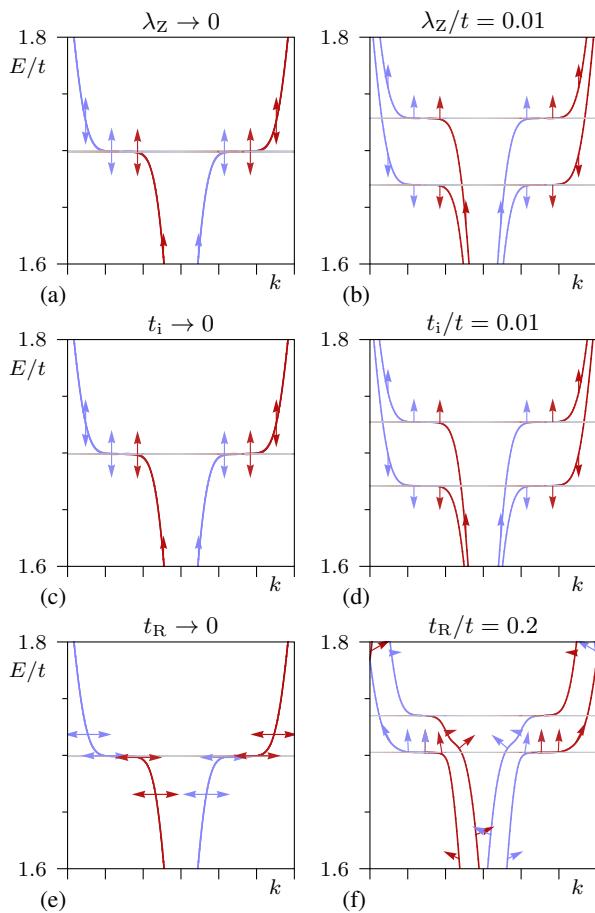


FIG. 9. (Color online) Spin splitting in an example at flux $\phi = 7/15$ and near $E/t = 1.7t$. Zeeman effect: (a) $\lambda_Z \rightarrow 0$ and (b) $\lambda_Z = 0.01t$; ISO coupling: (c) $t_I \rightarrow 0$ and (d) $t_I = 0.01t$; RSO coupling: (e) $t_R \rightarrow 0$ and (f) $t_R = 0.2t$.

E. Comparison between the low-flux limit $\phi \approx 0$ and the “half-flux” regime $\phi \approx 1/2$

As shown before in the low-flux limit $\phi \rightarrow 0$, the ISO coupling and the Zeeman effect affect the zero-energy modes equivalently in the sense that they both open a weak QSH gap. Indeed, it has been shown that for this mode, the ISO coupling and Zeeman effect are formally equivalent.⁴⁴ On the other hand, the RSO coupling does not open a gap at zero energy. This fact hampers the comparison of the RSO coupling to the aforementioned two terms in this energy regime.

From the Hofstadter butterfly shown in Fig. 1, one observes that the Landau-level structure in the low-flux, low-energy limit is not unique. In fact, similar structures appear wherever the shape of the dispersion is characterized by Dirac cones. For example, at $\phi = 1/2$ and for energies close to $E/t = \pm\sqrt{3}$, the dispersion can be approximated by a Dirac dispersion of the form $|E/t| = \sqrt{3} \pm \sqrt{\frac{3}{2}|\mathbf{k}| - \frac{1}{4}\sqrt{3}|\mathbf{k}|^2} + \mathcal{O}(|\mathbf{k}|^3)$, where \mathbf{k} is the momentum relative to the position of one of the two Dirac cones. The linear term is responsible for the square-root behavior of the (fourfold-degenerate) Landau levels for flux values close to $\phi = 1/2$. The quadratic term causes the asymmetry of the Landau level spectrum with respect to the energy $|E/t| = \sqrt{3}$. Furthermore, we observe that the energy $|E_{LL,0}/t| \approx \sqrt{3} - (\pi/3)|\phi - 1/2|$ of the central Landau level is linear in the flux rather than constant.

The approximate linear dependence of the spectrum $E(\mathbf{k})$ at $\phi = 1/2$ and $|E/t| \approx \sqrt{3}$ suggests that the Zeeman and ISO terms should behave similarly in this regime, in direct analogy to the case $\phi, E \approx 0$ discussed in Sec. IV C 1. Although the energy of the lowest Landau level is no longer constant in ϕ , the approximation holds, thus leading to a similar behavior with regard to the opening of the gaps, as shown in Figs. 9(a-d) for the flux value $\phi = 7/15$ which lies close to $\phi = 1/2$. The Zeeman effect and ISO coupling split the central Landau level at $E/t \approx 1.7$ and open up a weak QSH gap, similar to the

zero mode in the low-flux regime. As discussed in Sec. IV D, the RSO coupling splits the central Landau level, unlike the case for $\phi \rightarrow 0$, see Figs. 9(e,f). Although there are edge states inside the gap, it is a trivial state, because the spin Hall conductivity approximately vanishes. This comparison shows an important difference between Zeeman coupling and ISO coupling on one hand, and the Rashba coupling on the other hand: The former two create the weak QSH phase, while the latter does not. In the following section, we show that in the presence of Zeeman or ISO coupling, the Rashba coupling tends to destroy the weak QSH phases generated by the former effects.

V. COMPETITION AND PHASE TRANSITIONS

After analyzing the effects of the three terms H_I , H_R , and H_Z in Hamiltonian (1) independently, a natural question arises: what effects emerge when these terms are combined? Here, we study the competition featuring two terms, by tuning the ratio between their corresponding coupling constants. Tuning these coupling constants generally leads to opening and closing of gaps, so that the topological nature of the gaps may change. In that case, we deal with *topological phase transitions*. We recall that Kane and Mele⁴ have studied the phase transition from a (time-reversal symmetric) QSH state to a trivial state, by increasing the ratio t_R/t_I . Here, we extend this study and we explore the more exotic phase transitions that are realized in the presence of a magnetic field and the Zeeman effect.

A. Phase transitions driven by the ISO coupling

Let us analyze the interplay between the ISO coupling and the magnetic field by investigating a gap at a fixed flux value, in absence of the Zeeman effect and of the Rashba coupling. By tuning the value of the ISO coupling amplitude t_I , we can close the gaps and obtain different topological phases at both sides of the transition. In Fig. 10, we have shown the dispersions for $\phi = 1/3$, for values of t_I below, at, and above the values where the gap around $E/t = 2.3$ closes. Around this energy, there are three bulk bands: Two (overlapping) ones above and one below the gap. For values of $t_I < t_{I,0} \approx 0.4285t$, below the critical value, we observe the weak QSH phase, i.e., two counterpropagating edge modes with opposite spins on each edge, $N_\uparrow = 1$ and $N_\downarrow = -1$. The spin-down pair of edge modes connects the top and bottom bulk bands visible in this plot, and the spin-up pair connects the middle band to a lower band. For $t_I = t_{I,0}$, the gap disappears: The bulk bands touch each other in three points at $E/t \approx 2.3$, and two of these points are connected by the spin-down edge states in the middle of the figure. If the value of t_I is increased further, a gap opens again, and a new pair of edge states (with spin down) appears. Moreover, the existing spin-down edge states have inverted their direction of propagation. Thus, for $t_I > t_{I,0}$, we have a spin-imbalanced quantum Hall state with $N_\uparrow = 1$ and $N_\downarrow = 2$. Note that only the spin-down

states are modified in this phase transition, while the spin-up states remain the same.

In Ref. 9, this phase transition was investigated in the presence of a nonzero Zeeman coupling in addition to the ISO coupling. Here, we emphasize that the Zeeman effect is actually superfluous and that this transition is driven exclusively by the ISO coupling. Variation of λ_Z has no qualitative effect on the transition, but only shifts the energies of the bands and gaps. Within a large range of values for λ_Z , we obtain an identical phase transition as in Fig. 10. The critical value $t_{I,0}$ of the ISO strength is independent of λ_Z .

The phase transition presented here differs in a fundamental way from a crossing of Landau levels. In the latter case, the Chern numbers associated to the Landau levels do not change at the transition, and as a consequence the number of edge states inside the gap between them remains unmodified as well. For the phase transition presented here, the number of edge states and consequently the Chern numbers do change at the transition. Focusing on the spin-down edge states, we observe that $N_\downarrow = -1$ for $t_I < t_{I,0}$ and $N_\downarrow = 2$ for $t_I > t_{I,0}$, a difference of $\Delta N_\downarrow = 3$. The Chern numbers of the bands below and above the gap are 1 and 1 in the former case and -2 and 4 in the latter. The sum of these Chern numbers is unchanged, as required by the bulk-boundary correspondence, Eq. (5).

The question arises as to whether the difference in the number of edge states at both sides of the transition may be predicted. *A priori*, this difference cannot be predicted since it is the result of a complicated interplay of the ISO coupling and the magnetic flux. However, the difference is always a multiple of the denominator q of the flux $\phi = p/q$. This result may be understood from the fact that at a given flux value $\phi = p/q$, the Chern numbers of all the Hofstadter bands obey $C_n \equiv c \pmod{q}$. Here, c is the *modular multiplicative inverse* of p modulo q , defined as the unique integer c ($0 \leq c < q$) such that $cp \equiv 1 \pmod{q}$.⁵³ Consequently, the difference between two possible values of the Chern numbers is always a multiple of q , which proves the aforementioned claim. This property may also be understood from the superlattice structure: The magnetic field produces q copies of the original unit cell, and each of them reacts in the same way to the perturbation leading to the phase transition.

In Fig. 11, we show the phase diagram as a function of the ISO coupling t_I and the Fermi energy E_F with the flux value $\phi = 1/3$ held fixed. From this figure, we can easily read off the values of t_I and the energies E_F where the gaps close and the topological phase transitions occur. A few points must be noticed: First of all, we clearly identify the weak-QSH phase at half-filling for $t_I/t = 0.4$. The (spin-degenerate) QH phases are progressively destroyed for $t_I/t < 0.3$. For $t_I/t > 0.3$, we get spin-filtered QH [e.g., $(N_\uparrow, N_\downarrow) = (-1, 0)$], spin-imbalanced QH [e.g., $(N_\uparrow, N_\downarrow) = (+1, +2)$] and the weak QSH phase. In addition, there are also more exotic phases where the edge states are neither chiral nor helical, e.g., $(N_\uparrow, N_\downarrow) = (+1, -2)$. The spin-degenerate phases do not arise for $t_I/t > 0.3$. Secondly, the gaps are large (of the order of t) in wide regions of the parameters, which is favorable for the detection of the transitions in cold-atoms experi-

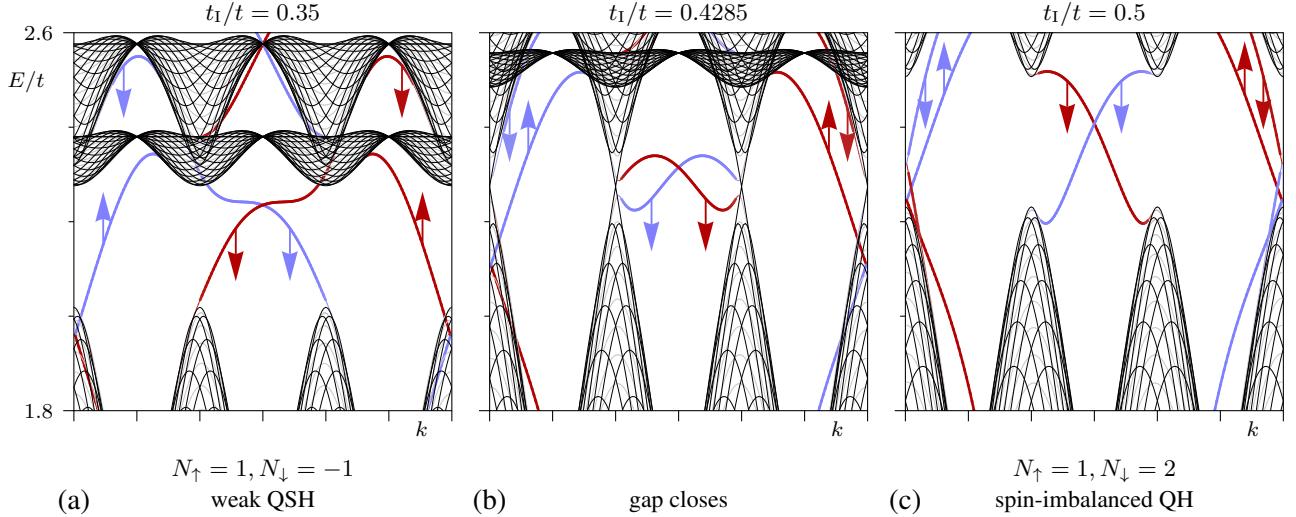


FIG. 10. (Color online) Spectra for $\phi = 1/3$, with (a) $t_1/t = 0.35$, (b) $t_1/t = 0.4285$, (c) $t_1/t = 0.5$. In (a), we observe a (weak) QSH gap, where $N_\uparrow = -N_\downarrow = 1$. (b) By increasing the ISO coupling constant, the bulk gap closes for $t_1/t \approx 0.4285$. (c) If t_1 is increased further, the gap reopens, but with a different number of edge states, namely $N_\uparrow = 1$ and $N_\downarrow = 2$, so that it enters a spin-imbalanced phase. Thus, the system undergoes a topological phase transition from a helical phase (weak QSH) to a chiral phase (spin-imbalanced QH).

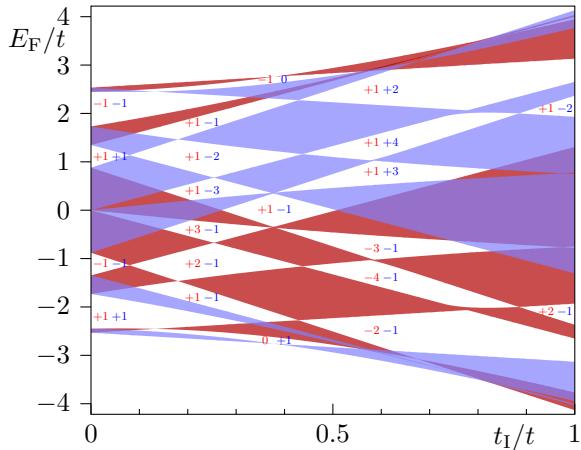


FIG. 11. (Color online) Phase diagram as a function of the ISO coupling t_1 and the Fermi energy E_F for the fixed flux value $\phi = 1/3$. In the shaded regions, the system is metallic. The colors red (dark gray) and blue (light gray) distinguish the spin-up and -down components, respectively. In the white regions, the bulk is insulating, and the edges conduct. There, the pairs of numbers indicate the number of edge states N_\uparrow and N_\downarrow for spin up and down, respectively.

ments. In addition, the property that the number of edge states changes by a multiple of 3 at each phase transition is clearly observed in Fig. 11. Finally, the energies where the phase transitions occur may be shifted by tuning the strength of the Zeeman effect. Since the spin-up and spin-down components are uncoupled, the Zeeman effect will shift all red (blue) areas in Fig. 11 up (down) by a fixed amount of energy. We note

that the Zeeman effect alone does not induce the same type of phase transitions as the ISO coupling. The Zeeman effect can only close and open gaps between bands of different spin components, and does therefore not modify the Chern numbers of the bands, whereas the ISO coupling can also close gaps between two bands of the same spin component. Nevertheless, tuning the Zeeman coupling strength λ_Z allows one to modify the nature of the phase transitions driven by ISO coupling.

B. Intrinsic spin-orbit coupling and Zeeman effect for $\phi \neq 0$

If the Zeeman effect and ISO coupling are present simultaneously, the interplay between the two terms is governed by the value of the flux ϕ , because the amplitude of the former is linear in ϕ while that of the latter is constant. Thus, in the butterfly spectrum illustrated in Fig. 12, we observe the ISO-dominated regime at low flux (cf. Fig. 3), and the Zeeman-dominated regime at high flux (cf. Fig. 4). The crossover between these two regimes takes place where the two terms are comparable in strength, approximately where $4\pi\phi\lambda_Z \sim 6\sqrt{3}t_1$, indicated by the vertical dashed line in Fig. 12 at $\phi \approx 0.165$. We remark that this crossover effect is present more generally in the situation where the Zeeman term competes with any other term with a constant amplitude: Due to its linear ϕ -dependence, the Zeeman term will always dominate any other one at sufficiently high values of ϕ . Naturally, this phenomenon only occurs in systems where the Zeeman splitting depends linearly on the magnetic flux ϕ . In the presence of magnetic impurities, the Zeeman term may exhibit a nonlinear dependence on the applied magnetic flux, as discussed in Ref. 49 for Mn-doped HgTe quantum wells, but

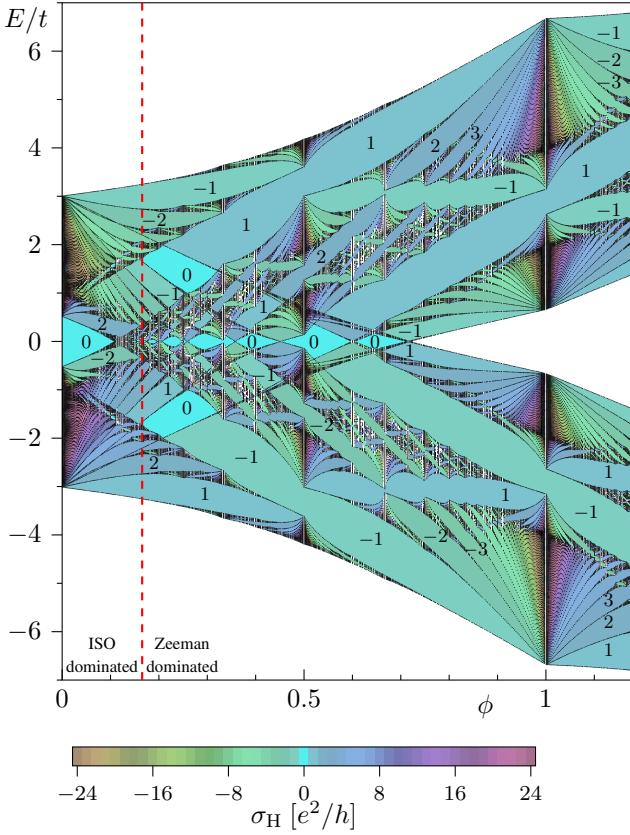


FIG. 12. (Color online) Hofstadter butterfly spectrum with ISO coupling ($t_1/t = 0.1$) and Zeeman effect ($\lambda_Z/t = 0.5$). The colors and numbers inside the bulk gaps indicate the quantized Hall conductivity in units of e^2/h . The vertical dashed line indicates the flux value where the two effects are comparable in strength. At this value, the crossover from the ISO-dominated (low ϕ) to the Zeeman-dominated (high ϕ) regime takes place.

this subject is beyond the scope of this paper.

Although the Zeeman effect dominates over the ISO coupling for $\phi \gg 0.165$, the influence of the ISO coupling is still visible in the high-flux spectrum: At $\phi = 1/2$ and $\phi = 1$, there are gaps at $E/t \approx 3$ (see Fig. 12), which are absent in the butterfly without the ISO coupling (see Fig. 4). We note that the dispersions around these points are of Dirac type for $t_1 = 0$, consequently, the ISO coupling opens these gaps in a way analogous to the gaps at zero flux and zero energy. However, the gaps at $\phi = 1/2$ and $\phi = 1$ are spin-filtered QH phases and not QSH, due to the subtle competition between the ISO coupling and the TRS-breaking effects.

C. Rashba spin-orbit coupling and exchange term at $\phi = 0$

In two recent works,⁵⁴ it has been shown that the combination of RSO and an exchange field, similar in structure to the Zeeman term, leads to chiral edge states at zero energy. In addition, the problem has been investigated for a sodium/lithium

iridate model which involves NN and NNN hopping terms.⁵⁵ In the latter and more complex setup, nontrivial QH gaps appear also away from half filling. Here, we consider the combined effects of the Zeeman and Rashba couplings, away from half-filling and in the absence of ISO coupling. We demonstrate that *several* QH gaps appear in the absence of NNN hopping terms, without including any gauge field (i.e., the Peierls phases): Namely, non-trivial QH phases are produced in our model by setting $\phi = t_1 = 0$, $t_R \neq 0$ and by applying a flux-independent Zeeman, or exchange,^{11,54} term

$$H_{\text{ex}} = g_{\text{ex}} \sum_j c_j^\dagger \sigma_z c_j, \quad (8)$$

which differs from H_z by the fact that the strength g_{ex} no longer depends on the flux ϕ (which we now set to zero). We stress that it is the simple association of a constant exchange term with a NN Rashba hopping term that leads to non-trivial topological phases, at half-filling but also at $E_F \approx g_{\text{ex}}$, as discussed in detail below.

When the exchange term is added to the tight-binding model with strength $0 < g_{\text{ex}} < 3t$, the energy bands associated to the spin-up and spin-down components are shifted in opposite directions but overlap at half-filling. In this situation, which is illustrated in Fig. 13(a), the system is a semi-metal for $E_F = \pm g_{\text{ex}}$ and a metal otherwise. When the Rashba coupling is turned on, bulk gaps open not only at half-filling $E = 0$ but also at $E = \pm g_{\text{ex}}$ (cf. Fig. 13(b)). We have computed the edge-state structures as well as the topological indices associated to the gaps: we obtain that all these gaps are related to non-trivial QH phases. More precisely, we find that the gaps at $E = \pm g_{\text{ex}}$ host a single edge state per edge, the spin orientation of which is given by $\langle \sigma_z \rangle \approx \pm 1$: these gaps correspond to spin-filtered QH phases, with equal (resp. opposite) charge (resp. spin) Hall conductivities. We observe that the spin orientation of these isolated edge states is nearly vertical, which is due to the fact that the exchange term is important in this high-energy regime. At half-filling, the bulk gap hosts two edge states per edge, with same velocity and quasi-horizontal spin orientation. This bulk gap corresponds to a topological QH phase with $\sigma_H = 2$ and a non-trivial spin orientation. Therefore the spin orientation of the edge states highly depends on the Fermi energy. We note that the gap openings occur for large exchange coupling, but for arbitrarily low Rashba coupling.

When the exchange strength $g_{\text{ex}} > 3t$, the energy bands associated to the spin-up and spin-down components are well separated: a trivial gap opens at half-filling, while the Dirac points are shifted to the energies $E = \pm g_{\text{ex}}$, cf. Fig. 13(c). In this situation, the Rashba coupling still opens two non-trivial QH phases at $E_F = \pm g_{\text{ex}}$, cf. Fig. 13(d). Again, these phases are characterized by spin-filtered edge states with nearly vertical spin orientation. Therefore, for sufficiently large exchange coupling, it is necessary to set the Fermi energy away from half-filling in order to detect nontrivial topological phases. This behavior is summarized in the phase diagram of Fig. 13(e): For nonzero Rashba coupling t_R , the Hall conductivity at zero energy is $\sigma_H = \pm 2$ for $|g_{\text{ex}}/t| < 3$ and zero otherwise. Thus, topological phase transitions of the

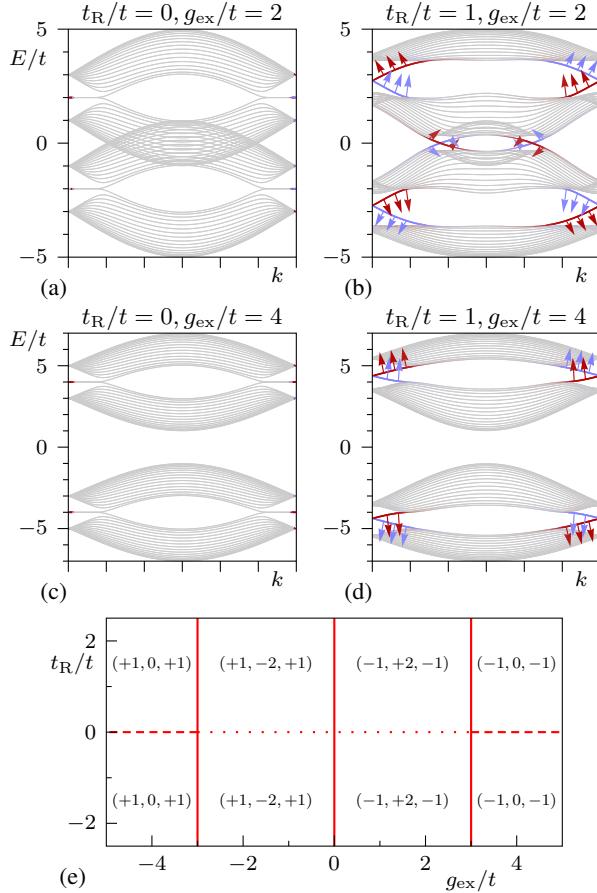


FIG. 13. (Color online) (a-d) Energy spectra $E = E(k)$ for $\phi = t_1 = 0$. The Zeeman (exchange) and Rashba coupling strength are indicated. (e) Phase diagram in the parameter space defined by the coupling strengths g_{ex} and t_{R} . The number triplets indicate the charge Hall conductivity in the lower, middle, and upper gaps, respectively. The solid lines indicate the phase transitions for the gap at zero energy. At the dashed lines, the lower and upper gaps close and the trivial gap at zero energy remains open [as in (c)], and at the dotted lines, all three gaps close [as in (a)]. At the dashed and dotted lines, the charge Hall conductivities in the gaps do not change.

zero-energy gap take place at the gap closings at $g_{\text{ex}}/t = \pm 3$. The lower and upper gaps (at $E_F \approx \pm g_{\text{ex}}$) always exhibit a Hall conductivity of ± 1 , where the sign depends on that of g_{ex}/t but not on that of t_{R}/t . Inverting the sign of g_{ex} inverts the directions of the edge currents, as well as the z component of the spins. Inverting the sign of t_{R} inverts the in-plane component of the spins. Continuous variation of t_{R} through zero changes the spin-direction continuously, where the spins tend to the z direction in the limit $t_{\text{R}}/t \rightarrow 0$. Thus, the gap closing at $t_{\text{R}} = 0$ does not constitute a topological phase transition.

Let us emphasize the fact that the chiral QH phases presented here are not the consequence of the external magnetic flux (i.e., Peierls phases), since they are produced when $\phi = 0$, and solely in the presence of the exchange and Rashba terms. When $\phi \neq 0$, new QH gaps open in the system, altering the exchange-Rashba-induced phases presented in Figs. 13(b)

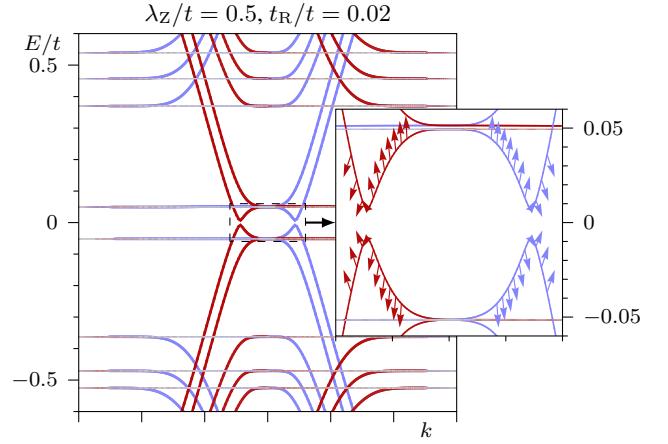


FIG. 14. (Color online) Spectrum for the system with nonzero Zeeman ($\lambda_Z/t = 0.5$) and RSO coupling ($t_{\text{R}}/t = 0.02$), for $\phi = 1/61$. The inset shows a zoom of the region where a gap opens between the edge states at zero energy. For comparison with the spectrum in absence of Rashba coupling, refer to Fig. 5(b).

and 13(d). This competition between QH gaps of different origins leads to complex and rich quantum phase transitions.

D. Zeeman effect and Rashba spin orbit coupling for $\phi \neq 0$

In the low-energy regime, where the (weak) QSH phase is generated by either ISO coupling or the Zeeman effect, the additional inclusion of the Rashba coupling has several profound effects on the physics. First of all, the RSO coupling tends to decrease the size of the gap (cf. Ref. 4), eventually completely closing it. Secondly, the combination of RSO coupling with ISO coupling and/or Zeeman effect destroys the particle-hole symmetry of the spectrum. Finally, at finite magnetic field, the RSO coupling opens up a gap (i.e., an avoided crossing) between the edge states at their crossing around zero energy. Around this crossing, the spin direction rotates from down to up (or vice versa), as illustrated in Fig. 14. The size of this gap is approximately linear in ϕ and in t_{R} (in the limit $\phi \rightarrow 0$ and $t_{\text{R}} \rightarrow 0$). Thus, the bulk gap is trivial in the sense that no edge states cross it from one bulk band to the other. However, in the limit where the edge-state gap is small, one would still observe a weak QSH state. At finite temperatures, the thermal energy can “bridge” the edge-state gap, so that the charge carriers may tunnel from one edge state to another. As the edge-state gap size is increased, the backscattering is enhanced, due to the decrease in the amount of tunneling across the edge-state gap. Here, we expect that the value of the spin Hall conductivity will deviate from its quantized value $\pm 2e/4\pi$. The amount of deviation increases when the gap size increases.

VI. EXPERIMENTAL REALIZATIONS

In this section, we will discuss the several experimental approaches that allow for the observation of the topological phases and phase transitions discussed in Secs. IV and V. For each of them, we present the opportunities and challenges, and we discuss the range of parameters that can be probed.

A. Condensed-matter systems

Condensed-matter systems typically have a lattice spacing in the order of 1 Å, so that one flux quantum per unit cell ($\phi = 1$) corresponds to a magnetic field in the order of 10^4 – 10^5 T. Thus, only the low-flux limit is relevant with respect to realistic condensed-matter realizations of our model. The prime example of a two-dimensional system with a honeycomb lattice is graphene, which has received much attention both at the theoretical and experimental levels.^{43,44} In the presence of a magnetic field, graphene shows an anomalous quantum Hall effect, where the Landau-level energies⁴⁵ depend on the magnetic field as $\sqrt{\phi}$, showing a radically different behavior than in semiconductor heterostructures (e.g., GaAs/AlGaAs).

Besides, the effects of spin-orbit coupling have been extensively studied for graphene, leading to the concept of Z_2 topological insulators.⁴ However, recent analysis has shown that graphene exhibits only a very weak ISO coupling ($t_1/t \sim 10^{-6}$ – 10^{-5} , where $t \sim 2.8$ meV),^{51,56} which unfortunately prevents the observation of the QSH state in this material. Nevertheless, the ISO interactions may be enhanced by putting heavy adatoms on top of the graphene surface,¹² which would open up a route to build a topological heterojunction.¹⁰ The Zeeman effect is of appreciable order of magnitude, $g \sim 2$, which corresponds to gaps of 100 K at a magnetic field of 30 T.⁵⁷ Values of the Rashba coupling as high as $t_R \sim 0.1t$ have been reported for graphene on a Ni(111) substrate.⁵⁸ Adsorption of metal atoms on the graphene surface increases this coupling as well.⁵⁹ Alternatively, high values of the Rashba coupling (accompanied by the spin re-orientation at the edges) may be induced by curving graphene nanoribbons.⁶⁰ Besides, it has been suggested that large Rashba coupling leads to topological phases in bilayer graphene.⁶¹ In addition, large effective pseudomagnetic fields, greater than 300 T, have been recently realized in strain-induced graphene nanobubbles.⁶² In summary, the observation of the topological phases presented here remains elusive in graphene, but should not be definitively ruled out.

The QSH effect has nevertheless been observed in other two-dimensional compounds that have considerably larger values of t_1 , such as Hg(Cd)Te quantum wells,^{6,7} where values of the order $t_1 \sim 1$ – 10 meV have been reported.^{63,64} Additionally, this material exhibits a strong Zeeman effect as well ($g \sim 20$).⁶⁵ The spin-filtered QH states have already been observed in these systems at low fields.⁶⁶ Hg(Cd)Te quantum wells doped with magnetic ions (e.g., Mn) exhibit an exchange interaction leading to the quantum anomalous Hall effect at zero magnetic field,⁶⁷ and to a nonlinear Zeeman effect

that causes reentrant behavior of the topological phases.⁴⁹

B. Artificial honeycomb lattices

The experimental limitation of the low flux values can be overcome by studying systems with a larger lattice spacing. Recently, several approaches have led to the engineering of artificial honeycomb lattices. First of all, arrays of quantum dots on a GaAs, arranged in a honeycomb pattern, have been engineered to simulate actual graphene.^{13,68} The lattice spacing in these systems is of the order of 100 nm, so that high flux values ($\phi \sim 1$) are obtained already with small magnetic fields. However, this large lattice constant leads to a small hopping amplitude, which has hampered the observation of the Hofstadter butterfly until now. Patterned superlattices able to probe the Hofstadter regime with usual magnetic field intensities have been realized in the past, but for square geometries.^{69,70} To probe the physics described here, one needs a superlattice with a smaller lattice constant, grown on a substrate that could lead to the generation of a strong Rashba interaction, for example. In this case, chiral topological states could be probed also away from zero energy, as we discuss in Sec. IV E.

Secondly, artificial graphene lattices have been created by manipulating CO molecules with the tip of a scanning tunneling microscope and arranging them in a triangular lattice on a Cu(111) substrate.¹⁴ These molecules repel the electrons on the Cu surface, thus creating a honeycomb array of sites between which the electrons hop. Measurements of the density of states show that the electrons indeed exhibit a Dirac dispersion. The typical lattice constants of these systems are approximately 7 times larger than in real graphene, thus allowing for flux values ~ 50 times larger than in graphene, with reasonable magnetic fields. In addition, effective magnetic fields up to 60 T can be induced by triaxial strain of the artificial lattice. It can be speculated that by using a heavier metal as substrate, it may be possible to engineer such a lattice with strong spin-orbit coupling, leading to the QSH phase.

C. Ultracold atoms in optical lattices

Although the parameters of the Hamiltonian (1) can be controlled to some extent in condensed-matter systems, the possibility to engineer the physics discussed here with ultracold atoms in optical lattices is very appealing. The advantage of the ultracold-atom realization of this model is the flexibility to control the parameters separately across a large range, compared to condensed-matter systems where these parameters are generally fixed by the material properties and the sample geometry.^{29,71} A second major advantage of ultracold atoms over condensed-matter systems is that the magnetic fields are produced synthetically, i.e., the Peierls phases are not due to a real magnetic field, but induced by external fields acting on atomic internal states.²⁰ In this way, large effective magnetic flux can be produced (i.e. $\phi \sim 1$), which allows the system to be studied beyond the low-flux limit.²¹ The spin-orbit couplings can also be synthesized in optical lattices using simi-

lar methods.²³ Moreover, in these configurations, the Zeeman coupling can be controlled independently, as it is not related to the synthetic magnetic fields⁷² (see Sec. V C, where we study the extreme situation where the effect of the Zeeman coupling is investigated in the absence of synthetic magnetic field, in which case this term is called an “exchange term”). Another important feature offered by optical-lattice setups is the possibility to control the interactions between the particles: by exploiting Feshbach resonances,²⁹ it is possible to reach the non-interacting regime and from there on, to include interactions in a controlled way to explore the robustness of topological order against interactions. Let us mention that these systems are usually free of disorder, although disorder can also be generated and tailored at will in optical lattices.⁷³ These outstanding features allow us to study the TRS-broken QSH phase, namely the survival of the QSH phase in the presence of a TRS-breaking perturbation.⁹ Indeed, contrarily to condensed-matter experiments, the absence of (magnetic) disorder can stabilize the helical (counterpropagating) edge states even when TRS is broken. Let us also comment on the fact that all the topological phases and edge-state excitations discussed here rely on the existence of robust bulk energy gaps. Therefore, when the temperature of the system approaches the size of the topological gaps of interest, one expects the effects to be strongly affected, and eventually to disappear. For the ultracold temperatures achieved in cold-atom laboratories, the gaps of interest should typically be of the order $\Delta \sim t$, where t is the tunneling amplitude. Therefore, we anticipate that only the largest topological gaps presented in our figures should be considered from the cold-atom point of view.

The quantum phase transitions presented in this work could be equally characterized either by a change in the topological invariants associated to the bulk gaps of interest, or by the modification of the edge-state structures. Although cold atoms do not offer the possibility to directly measure transport coefficients, several methods have been proposed to detect the topological invariants related to the quantized conductivities.^{39,74} These proposals are based on the fact that non-trivial band structures modify the atomic densities, leading to several signatures that can be directly seen in time-of-flight or density measurements. On the other hand, it is possible to directly probe edge state structures using Bragg spectroscopy techniques.^{75–77} In a recent work,⁷⁶ it was shown that a state-dependent light probe, focused on the edge of the cloud and transferring angular momentum to the atoms, clearly identifies the presence of chiral edge states in a QH optical lattice. This method allows us to project the topological edge states on a dark background, making them visible in *in situ* density measurements. In principle, this efficient scheme could be generalized to spin-1/2 systems in order to detect and distinguish between chiral and helical edge structures.

VII. CONCLUSIONS

We have shown an in-depth analysis of the edge-state structures generated by the combination of the ISO coupling, the RSO coupling, and the Zeeman effect, together with a uniform

and perpendicular magnetic field. In Sec. IV, we have studied each term individually in order to pinpoint their effect on the quantum Hall states generated by the magnetic field. In the low-field limit, the ISO coupling and Zeeman effect are formally equivalent, and give rise to the weak QSH phase. Away from zero energy, we observe spin-filtered, spin-imbalanced, and ordinary quantum Hall phases. The RSO coupling is different in the sense that it involves spin-flip terms, which causes the spin states to acquire an in-plane component. In addition, the spin direction inside the gap depends on the Fermi energy, and consequently the spin Hall conductivity of the system is ill-defined. In the case of high magnetic fields, we observe Dirac-type regimes away from zero energy. Here, the Zeeman and ISO coupling open weak QSH gaps in a way analogous to the low-flux limit, whereas the RSO coupling opens a trivial gap (unlike the case at low flux, where the spectrum remains gapless). In addition, the topological phase transitions reported in an earlier work⁹ have been analyzed in more detail here.

A combination of the spin-orbit terms and the Zeeman effect leads to subtle competitions, several examples of which have been discussed in Sec. V. We have shown that, although the Zeeman effect and ISO coupling both give rise to the QSH phase, they do not always reinforce each other, which has been illustrated by the size of the parameter regime of the weak QSH phase in the low-flux limit. At high flux, the Zeeman effect dominates over the ISO coupling, but the effects of the latter are still well visible in the spectrum. Interestingly, a combination of the Rashba term with an exchange coupling leads to nontrivial topological phases at zero magnetic field, away from half-filling. Finally, inclusion of the Rashba coupling in the case where we have a weak QSH gap generated by the Zeeman effect opens a gap at zero energy between the edge states, which provides an illustration of the breakdown of the QSH effect in the absence of TRS and spin conservation.

This work emphasizes the various topological phase transitions and edge-state structures that could be observed in a wide range of materials and quantum emulators. Although we focused our analysis on the honeycomb lattice, where the spin-orbit couplings and the magnetic field configurations are expressed by specific tunneling terms and Peierls phases, we note that our results do not rely on this specific geometry. The topological phases and the phase transitions discussed here are universal, in the sense that they only rely on the inclusion of general ingredients, which can be found or engineered in many different contexts and geometries. Indeed, this universal property can be understood in the low-energy limit, in which various (*a priori*, very different) tight-binding models actually reduce to the same (Dirac-type) equations. In general, we observe that the main ingredients needed to generate such physics are nontrivial gauge fields, which are coupled to the particles of interest (e.g., electrons in materials and atoms in optical lattices). In this work, we have aimed to emphasize the effects produced by each of these ingredients independently: We have investigated the effects of spin splitting, spin mixing, and gauge potentials leading to magnetic or Haldane-type local flux. We hope that our detailed analysis will deepen the understanding of topological insulating phases and that it will

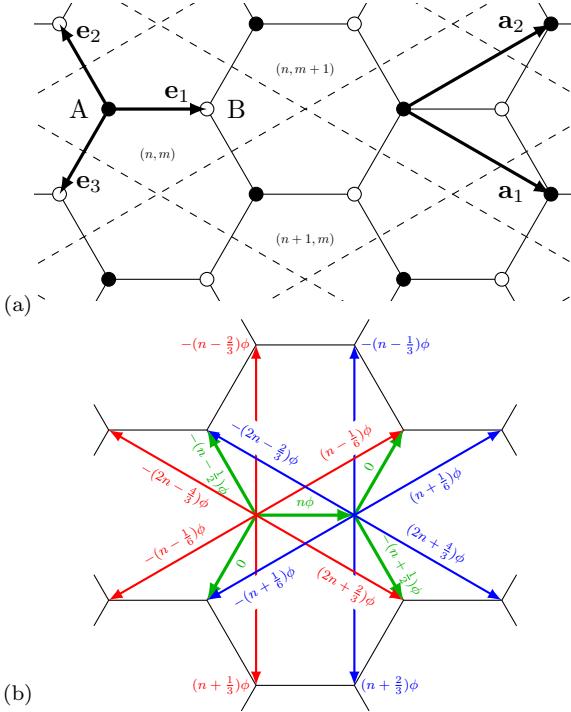


FIG. 15. (Color online) (a) The honeycomb lattice. We indicated the NN vectors e_j ($j = 1, 2, 3$) and the lattice vectors a_1 and a_2 . The sites of the sublattices A and B are indicated with solid and open circles, respectively. The dashed lines define boundaries of the unit cells, labeled as (n, m) . (b) Hopping phases $\theta_{jk}/2\pi$ for the unit cell (n, m) . The phases for the NN hopping are in green. The phases for the NNN hopping are in red and blue for the sublattices A and B, respectively.

motivate further developments in this exciting and rapidly developing field.

ACKNOWLEDGEMENTS

We are grateful to M. O. Goerbig and P. W. Brouwer for useful discussions. WB and CMS acknowledge financial support by the Netherlands Organisation for Scientific Research (NWO). NG thanks the F.R.S-F.N.R.S for financial support.

Appendix: Geometry and gauge

Geometrically, the honeycomb lattice is equivalent to a distorted square lattice where each unit cell has the shape of a rhombus. Since there are two inequivalent sets of sites, there are two sites per unit cell, and hence two sublattices, labeled A and B. The sites, labeled as j and k in the Hamiltonian (1), may be uniquely identified by two integer labels n and m for the unit cell and A or B for the sublattice: In the following, we shall write $j = (n, m, \tau)$ (where $n, m \in \mathbb{Z}$ and $\tau \in \{A, B\}$) to uniquely identify the sites.

The coordinates of the sites are given as follows. We choose the bond vectors e_1 , e_2 and e_3 as displayed in Fig. 15(a), where their length $a \equiv 1$ is chosen as the unit of length throughout this text. Subsequently, the lattice vectors are given by $a_1 = e_1 - e_2$ and $a_2 = e_1 - e_3$, so that the coordinates of the sites are

$$\mathbf{r}_{n,m,A} = n\mathbf{a}_1 + m\mathbf{a}_2, \quad \mathbf{r}_{n,m,B} = n\mathbf{a}_1 + m\mathbf{a}_2 + \mathbf{e}_1.$$

Furthermore, the reciprocal space is spanned by the two vectors b_1 and b_2 , defined by $b_i \cdot a_j = 2\pi\delta_{ij}$. We note that the results derived in this paper do not depend on the values of the vectors e_i ; the only effect of a change of the e_i is that the lattice vectors and the reciprocal lattice vectors will be different. A convenient choice for the bond vectors is $e_1 = (0, 0)$, $e_2 = (-1, 0)$ and $e_3 = (0, -1)$, effectively putting both sites of the unit cell (n, m) on the coordinate (n, m) .

The phases due to the magnetic field, associated with the hopping, are given in terms of the gauge potential \mathbf{A} as $\theta_{jk} = (e/\hbar) \int_{\mathbf{r}_k}^{\mathbf{r}_j} \mathbf{A} \cdot d\mathbf{l}$, where the integral is over the line between sites k and j . We may employ the gauge freedom for the gauge potential to choose it in such a way that the phase θ_{jk} for two sites $j = (n, m, \tau)$ and $k = (n', m', \tau')$ (NNs or NNNs) does not depend on the unit cell index m . The hopping phases resulting from such a choice are given in Fig. 15(b) for hopping from the sites of the unit cell (n, m) . The reader may check that the phase picked up from hopping around a loop is proportional to the area of the loop. In particular, when one considers one hexagon as the loop, the phase picked up is $2\pi\phi$, which corresponds to an enclosed flux of ϕ (in units of the flux quantum). Here, we note that the area of one hexagon is $3\sqrt{3}/2$, so that ϕ is related to the magnetic field strength B as $\phi = (3\sqrt{3}/2)eB/h$.

In addition to the gauge choice, let us furthermore assume that the flux value is a rational number, $\phi = p/q$, where p and q are coprime integers. Then, the phase assigned to each hopping is periodic in the index n with periodicity q . In other words, the unit cell of the superlattice (i.e., the lattice together with the hopping phases) is $q \times 1$ unit cells of the original lattice. With this observation, we may invoke Bloch's theorem, which states that we could write the electron wave functions $\varphi_{n,m,\tau}$ as

$$\begin{aligned} \varphi_{n,m,A} &= \psi_{nA} e^{i\mathbf{k} \cdot (n\mathbf{a}_1 + m\mathbf{a}_2)}, \\ \varphi_{n,m,B} &= \psi_{nB} e^{i\mathbf{k} \cdot (n\mathbf{a}_1 + m\mathbf{a}_2 + \mathbf{e}_1)}. \end{aligned} \quad (A.1)$$

With this ansatz, the Schrödinger equation which the fields have to satisfy reduces to a $4q$ -component matrix equation: The degrees of freedom are all fields $\Psi_n = (\psi_{nA\uparrow}, \psi_{nA\downarrow}, \psi_{nB\uparrow}, \psi_{nB\downarrow})$ in the unit cell of the superlattice, so that $n = 1, \dots, q$, and there are two sublattice and two spin components per unit cell of the original lattice. This matrix equation is then merely an eigenvalue equation of a $4q \times 4q$ matrix, known as the Harper equation,^{30,78}

$$\frac{E}{t} \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \vdots \\ \Psi_{q-1} \\ \Psi_q \end{pmatrix} = \begin{pmatrix} \mathcal{D}_1 & \mathcal{R}_1 & 0 & \cdots & 0 & \mathcal{R}_q^\dagger \\ \mathcal{R}_1^\dagger & \mathcal{D}_2 & \mathcal{R}_2 & \cdots & 0 & 0 \\ 0 & \mathcal{R}_2^\dagger & \mathcal{D}_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathcal{D}_{q-1} & \mathcal{R}_{q-1} \\ \mathcal{R}_q & 0 & 0 & \cdots & \mathcal{R}_{q-1}^\dagger & \mathcal{D}_q \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \vdots \\ \Psi_{q-1} \\ \Psi_q \end{pmatrix}. \quad (\text{A.2})$$

Here, the 4×4 matrices $\mathcal{D}_n = \mathcal{D}_n^{(\text{NN})} + \mathcal{D}_n^{(\text{R})} + \mathcal{D}_n^{(\text{I})} + \mathcal{D}_n^{(\text{Z})}$ and $\mathcal{R}_n = \mathcal{R}_n^{(\text{NN})} + \mathcal{R}_n^{(\text{R})} + \mathcal{R}_n^{(\text{I})}$ encode hopping between unit cells with the same index n and from unit cells with index $n+1$ to n . (Note that the Zeeman term is diagonal.) Thus, Eq. (A.2) can be written equivalently as Eq. (2).

The 4×4 matrices \mathcal{D}_n and \mathcal{R}_n may be written in terms of the sum of four contributions, that correspond to each of the terms in Hamiltonian (1). The contributions for the ordinary NN hopping are given by

$$\mathcal{D}_n^{(\text{NN})} = \begin{pmatrix} 0 & e^{i\mathbf{k}\cdot\mathbf{e}_1} e^{2\pi i n \phi} + e^{i\mathbf{k}\cdot\mathbf{e}_3} \\ e^{-i\mathbf{k}\cdot\mathbf{e}_1} e^{-2\pi i n \phi} + e^{-i\mathbf{k}\cdot\mathbf{e}_3} & 0 \end{pmatrix} \otimes \mathbb{1}_{\text{spin}}, \quad \mathcal{R}_n^{(\text{NN})} = \begin{pmatrix} 0 & 0 \\ e^{-i\mathbf{k}\cdot\mathbf{e}_2} e^{-2\pi i (n+\frac{1}{2}) \phi} & 0 \end{pmatrix} \otimes \mathbb{1}_{\text{spin}}. \quad (\text{A.3})$$

The Zeeman effect is encoded by $\mathcal{D}_n^{(\text{Z})} = (2\pi\lambda_Z\phi/t)\mathbb{1}_{\text{AB}} \otimes \sigma_z$ and $\mathcal{R}_n^{(\text{Z})} = 0$. The matrices for the ISO coupling are

$$\mathcal{D}_n^{(\text{I})} = -i \frac{t_{\text{I}}}{t} \begin{pmatrix} -e^{-2\pi i (n-\frac{1}{6}) \phi} e^{i\mathbf{k}\cdot\mathbf{f}_2} + e^{2\pi i (n-\frac{1}{6}) \phi} e^{-i\mathbf{k}\cdot\mathbf{f}_2} & 0 \\ 0 & e^{-2\pi i (n+\frac{1}{6}) \phi} e^{i\mathbf{k}\cdot\mathbf{f}_2} - e^{2\pi i (n+\frac{1}{6}) \phi} e^{-i\mathbf{k}\cdot\mathbf{f}_2} \end{pmatrix} \otimes \sigma_z, \quad (\text{A.4})$$

$$\mathcal{R}_n^{(\text{I})} = -i \frac{t_{\text{I}}}{t} \begin{pmatrix} -e^{2\pi i (2n+\frac{2}{3}) \phi} e^{i\mathbf{k}\cdot\mathbf{f}_3} + e^{2\pi i (n+\frac{1}{3}) \phi} e^{-i\mathbf{k}\cdot\mathbf{f}_1} & 0 \\ 0 & e^{2\pi i (2n+\frac{4}{3}) \phi} e^{i\mathbf{k}\cdot\mathbf{f}_3} - e^{2\pi i (n+\frac{2}{3}) \phi} e^{-i\mathbf{k}\cdot\mathbf{f}_1} \end{pmatrix} \otimes \sigma_z, \quad (\text{A.5})$$

where $\mathbf{f}_1 = \mathbf{e}_2 - \mathbf{e}_3$, $\mathbf{f}_2 = \mathbf{e}_3 - \mathbf{e}_1$, and $\mathbf{f}_3 = \mathbf{e}_1 - \mathbf{e}_2$ denote the NNN vectors. Here, the phase factors due to the flux correspond to those in Fig. 15(b). Due to the fact that the phases involve fractions of ϕ which are all multiples of $\frac{1}{6}$, the ϕ -periodicity of the butterfly spectra is 6, rather than the periodicity of 1 which is observed in absence of the ISO coupling. Finally, for the RSO coupling, the hopping matrices are given by

$$\mathcal{D}_n^{(\text{R})} = \frac{t_{\text{R}}}{t} \begin{pmatrix} 0 & 0 & 0 & -a_-^{(\text{R})} \\ 0 & 0 & a_+^{(\text{R})} & 0 \\ 0 & a_+^{(\text{R})*} & 0 & 0 \\ -a_-^{(\text{R})*} & 0 & 0 & 0 \end{pmatrix}, \quad (\text{A.6})$$

$$\mathcal{R}_n^{(\text{R})} = \frac{t_{\text{R}}}{t} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & b_+^{(\text{R})*} & 0 & 0 \\ b_-^{(\text{R})*} & 0 & 0 & 0 \end{pmatrix}, \quad (\text{A.7})$$

where $a_{\pm}^{(\text{R})} = e^{2\pi i n \phi} e^{i\mathbf{k}\cdot\mathbf{e}_1} + e^{\pm 2\pi i/3} e^{i\mathbf{k}\cdot\mathbf{e}_3}$ and $b_{\pm}^{(\text{R})} = e^{-2\pi i (n+\frac{1}{2}) \phi} e^{\mp 2\pi i/3} e^{i\mathbf{k}\cdot\mathbf{e}_2}$.

The solution of the Harper equation yields the dispersions of the bulk states, which are periodic over the so-called magnetic Brillouin zone, q times smaller than the original Brillouin zone. Effectively, the geometry of the system is a torus, in this case. However, to study edge states, a cylindrical geometry (as displayed in Fig. 2) is more convenient. The edge-state spectrum is obtained from the Harper equation similar

to Eq. (A.2), but with a matrix of dimension $4w \times 4w$, where w is the width of the ribbon in unit cells, and with the blocks in the top-right and bottom-left corner replaced by zeros (i.e. $\mathcal{R}_w \equiv 0$). In this case, the dispersion depends on just one component of \mathbf{k} , and its periodicity is the reciprocal lattice vector.

To determine states as bulk states or edge states, we derive the density profile $|\Psi_n|^2$ as a function of the unit cell index n (or equivalently, the coordinate component y). For convenience, we choose $n=0$ to be the center of the ribbon, so that $n = -\tilde{w}, \dots, \tilde{w}$, where $w = 2\tilde{w}+1$ is the width of the ribbon. The expectation value of this position, $\langle n \rangle = \sum_n n |\Psi_n|^2$ is used to distinguish edge states and bulk states. Edge states are characterized by a density profile that is sharply peaked at one edge, so that $\langle n \rangle \approx \pm \tilde{w}$. On the other hand, if the expectation value $\langle n \rangle$ is close to zero, then it means that there is a significant contribution of the density away from the edge; such a state is identified as a bulk state. In the edge state spectra of this paper, we have used this expectation value to color the states: Red and blue colors indicate the two opposite edges, while gray is used for the bulk states. In the same way, we use the spin expectation values $\langle \sigma_i \rangle = \sum_n \Psi_n^\dagger \sigma_i \Psi_n$ ($i = x, y, z$) to gain information about the spin states. We note that this expectation is only reliable in the case that the spin is constant (or almost constant) where the density of the state is concentrated. This may not be the case in the presence of Rashba coupling. In the case when the spin strongly depends on position, the length of the vector $\langle \sigma \rangle$ is less than unity.

¹ M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

² X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

- ³ K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).
- ⁴ C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
- ⁵ C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
- ⁶ M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
- ⁷ B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).
- ⁸ F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
- ⁹ N. Goldman, W. Beugeling, and C. Morais Smith, EPL **97**, 23003 (2012).
- ¹⁰ O. Shevtsov, P. Carmier, C. Petitjean, C. Groth, D. Carpentier, and X. Waintal, Phys. Rev. X **2**, 031004 (2012).
- ¹¹ Y. Yang, Z. Xu, L. Sheng, B. Wang, D. Y. Xing, and D. N. Sheng, Phys. Rev. Lett. **107**, 066602 (2011).
- ¹² C. Weeks, J. Hu, J. Alicea, M. Franz, and R. Wu, Phys. Rev. X **1**, 021001 (2011).
- ¹³ A. Singha, M. Gibertini, B. Karmakar, S. Yuan, M. Polini, G. Vignale, M. I. Katsnelson, A. Pinczuk, L. N. Pfeiffer, K. W. West, and V. Pellegrini, Science **332**, 1176 (2011).
- ¹⁴ K. K. Gomes, W. Mar, W. Ko, F. Guinea, and H. C. Manoharan, Nature **483**, 306 (2012).
- ¹⁵ P. Soltan-Panahi, J. Struck, P. Hauke, A. Bick, W. Plenkers, G. Meineke, C. Becker, P. Windpassinger, M. Lewenstein, and K. Sengstock, Nature Phys. **7**, 434 (2011).
- ¹⁶ L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, and T. Esslinger, Nature **483**, 302 (2012).
- ¹⁷ K. L. Lee, B. Grémaud, R. Han, B.-G. Englert, and C. Miniatura, Phys. Rev. A **80**, 043411 (2009).
- ¹⁸ Y.-J. Lin, R. L. Compton, K. Jiménez-García, J. V. Porto, and I. B. Spielman, Nature **462**, 628 (2009).
- ¹⁹ Y.-J. Lin, K. Jiménez-García, and I. B. Spielman, Nature **471**, 83 (2011).
- ²⁰ J. Dalibard, F. Gerbier, G. Juzeliūnas, and P. Öhberg, Rev. Mod. Phys. **83**, 1523 (2011).
- ²¹ M. Aidelsburger, M. Atala, S. Nascimbène, S. Trotzky, Y.-A. Chen, and I. Bloch, Phys. Rev. Lett. **107**, 255301 (2011).
- ²² N. R. Cooper, Phys. Rev. Lett. **106**, 175301 (2011).
- ²³ N. Goldman, I. Satija, P. Nikolic, A. Bermudez, M. A. Martin-Delgado, M. Lewenstein, and I. B. Spielman, Phys. Rev. Lett. **105**, 255302 (2010); L. Mazza, A. Bermudez, N. Goldman, M. Rizzi, M. A. Martin-Delgado, and M. Lewenstein, New J. Phys. **14**, 015007 (2012); K. Osterloh, M. Baig, L. Santos, P. Zoller, and M. Lewenstein, Phys. Rev. Lett. **95**, 010403 (2005); B. Béri and N. R. Cooper, Phys. Rev. Lett. **107**, 145301 (2011).
- ²⁴ L.-K. Lim, C. Morais Smith, and A. Hemmerich, Phys. Rev. Lett. **100**, 130402 (2008); L.-K. Lim, A. Hemmerich, and C. Morais Smith, Phys. Rev. A **81**, 023404 (2010).
- ²⁵ N. Goldman, A. Kubasiak, A. Bermudez, P. Gaspard, M. Lewenstein, and M. A. Martin-Delgado, Phys. Rev. Lett. **103**, 035301 (2009).
- ²⁶ A. Bermudez, L. Mazza, M. Rizzi, N. Goldman, M. Lewenstein, and M. A. Martin-Delgado, Phys. Rev. Lett. **105**, 190404 (2010).
- ²⁷ Z. Lan, N. Goldman, A. Bermudez, W. Lu, and P. Öhberg, Phys. Rev. B **84**, 165115 (2011).
- ²⁸ C. Chin, R. Grimm, P. Julienne, and E. Tiesinga, Rev. Mod. Phys. **82**, 1225 (2010).
- ²⁹ I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
- ³⁰ D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).
- ³¹ In a more general form, the Harper equation is known as an *almost Mathieu equation* in mathematical literature; see e.g., J. Bellissard and B. Simon, J. Funct. Anal. **48**, 408 (1982).
- ³² R. Rammal, J. Phys. (Paris) **46**, 1345 (1985).
- ³³ Y. Hatsugai, Phys. Rev. B **48**, 11851 (1993).
- ³⁴ M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).
- ³⁵ M. Kohmoto, Ann. Phys. **160**, 343 (1985).
- ³⁶ Y. Hatsugai, Phys. Rev. Lett. **71**, 3697 (1993).
- ³⁷ D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
- ³⁸ P. Středa, J. Phys. C **15**, L717 (1982); J. Phys. C **15**, L1299 (1982).
- ³⁹ R. O. Umucalilar, H. Zhai, and M. Ö. Oktel, Phys. Rev. Lett. **100**, 070402 (2008).
- ⁴⁰ A. Bermudez, N. Goldman, A. Kubasiak, M. Lewenstein, and M. A. Martin-Delgado, New J. Phys. **12**, 033041 (2010).
- ⁴¹ D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, Phys. Rev. Lett. **97**, 036808 (2006).
- ⁴² E. Prodan, Phys. Rev. B **80**, 125327 (2009).
- ⁴³ A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
- ⁴⁴ M. O. Goerbig, Rev. Mod. Phys. **83**, 1193 (2011).
- ⁴⁵ K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature **438**, 197 (2005); Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, Nature **438**, 201 (2005).
- ⁴⁶ Y. Hatsugai, T. Fukui, and H. Aoki, Phys. Rev. B **74**, 205414 (2006).
- ⁴⁷ X.-L. Qi and S.-C. Zhang, Physics Today **63**, 33 (2010).
- ⁴⁸ S. Das Sarma and A. Pinczuk, *Perspectives in Quantum Hall Effects: Novel Quantum Liquids in Low-Dimensional Semiconductor Structures* (John Wiley & Sons, 2008).
- ⁴⁹ W. Beugeling, C. X. Liu, E. G. Novik, L. W. Molenkamp, and C. Morais Smith, Phys. Rev. B **85**, 195304 (2012).
- ⁵⁰ M. Zarea and N. Sandler, Phys. Rev. B **79**, 165442 (2009).
- ⁵¹ R. van Gelderen and C. Morais Smith, Phys. Rev. B **81**, 125435 (2010).
- ⁵² For bilayer graphene, which is characterized by a Hamiltonian that is formally equivalent to $H_{NN} + H_R$, these Berry phases have been derived in R. de Gail, M. O. Goerbig, and G. Montambaux, Phys. Rev. B **86**, 045407 (2012).
- ⁵³ This result has been verified empirically for the spinless model. The observation that $C_n \equiv c \pmod{q}$ for all n implies automatically that $cp \equiv 1 \pmod{q}$, assuming that the Středa formula and the bulk-boundary correspondence hold. These results are valid more generally for the model described by Hamiltonian (1) as long as the two spin components are decoupled.
- ⁵⁴ Z. Qiao, S. A. Yang, W. Feng, W.-K. Tse, J. Ding, Y. Yao, J. Wang, and Q. Niu, Phys. Rev. B **82**, 161414 (2010); Z. Qiao, H. Jiang, X. Li, Y. Yao, and Q. Niu, Phys. Rev. B **85**, 115439 (2012).
- ⁵⁵ G. Murthy, E. Shimshoni, R. Shankar, and H. A. Fertig, Phys. Rev. B **85**, 073103 (2012).
- ⁵⁶ J. C. Boettger and S. B. Trickey, Phys. Rev. B **75**, 121402 (2007).
- ⁵⁷ D. A. Abanin, P. A. Lee, and L. S. Levitov, Phys. Rev. Lett. **96**, 176803 (2006).
- ⁵⁸ Y. S. Dedkov, M. Fonin, U. Rüdiger, and C. Laubschat, Phys. Rev. Lett. **100**, 107602 (2008).
- ⁵⁹ J. Ding, Z. Qiao, W. Feng, Y. Yao, and Q. Niu, Phys. Rev. B **84**, 195444 (2011).
- ⁶⁰ D. Gosálbez-Martínez, J. J. Palacios, and J. Fernández-Rossier, Phys. Rev. B **83**, 115436 (2011).
- ⁶¹ Z. Qiao, W.-K. Tse, H. Jiang, Y. Yao, and Q. Niu, Phys. Rev. Lett. **107**, 256801 (2011).
- ⁶² N. Levy, S. A. Burke, K. L. Meaker, M. Panlasigui, A. Zettl, F. Guinea, A. H. Castro Neto, and M. F. Crommie, Science **329**, 544 (2010).
- ⁶³ C. Brüne, C. X. Liu, E. G. Novik, E. M. Hankiewicz, H. Buhmann, Y. L. Chen, X. L. Qi, Z. X. Shen, S. C. Zhang, and L. W.

- Molenkamp, Phys. Rev. Lett. **106**, 126803 (2011).
- ⁶⁴ E. G. Novik, A. Pfeuffer-Jeschke, T. Jungwirth, V. Latussek, C. R. Becker, G. Landwehr, H. Buhmann, and L. W. Molenkamp, Phys. Rev. B **72**, 035321 (2005).
- ⁶⁵ M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, J. Phys. Soc. Jpn **77**, 031007 (2008).
- ⁶⁶ B. Büttner, C. X. Liu, G. Tkachov, E. G. Novik, C. Brüne, H. Buhmann, E. M. Hankiewicz, P. Recher, B. Trauzettel, S. C. Zhang, and L. W. Molenkamp, Nature Phys. **7**, 418 (2011).
- ⁶⁷ C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. Lett. **101**, 146802 (2008).
- ⁶⁸ G. De Simoni, A. Singha, M. Gibertini, B. Karmakar, M. Polini, V. Piazza, L. N. Pfeiffer, K. W. West, F. Beltram, and V. Pellegrini, Appl. Phys. Lett. **97**, 132113 (2010).
- ⁶⁹ S. Melinte, M. Berciu, C. Zhou, E. Tutuc, S. J. Papadakis, C. Harrison, E. P. De Poortere, M. Wu, P. M. Chaikin, M. Shayegan, R. N. Bhatt, and R. A. Register, Phys. Rev. Lett. **92**, 036802 (2004).
- ⁷⁰ Y. Zhang, Z. Jiang, J. P. Small, M. S. Purewal, Y.-W. Tan, M. Fazlollahi, J. D. Chudow, J. A. Jaszczak, H. L. Stormer, and P. Kim, Phys. Rev. Lett. **96**, 136806 (2006).
- ⁷¹ I. Bloch, J. Dalibard, and S. Nascimbène, Nature Phys. **8**, 267 (2012).
- ⁷² D. Makogon, I. B. Spielman, and C. Morais Smith, EPL **97**, 33002 (2012).
- ⁷³ L. Sanchez-Palencia and M. Lewenstein, Nature Phys. **6**, 87 (2010).
- ⁷⁴ E. Alba, X. Fernandez-Gonzalvo, J. Mur-Petit, J. K. Pachos, and J. J. Garcia-Ripoll, Phys. Rev. Lett. **107**, 235301 (2011); E. Zhao, N. Bray-Ali, C. J. Williams, I. B. Spielman, and I. I. Satija, Phys. Rev. A **84**, 063629 (2011).
- ⁷⁵ T. D. Stanescu, V. Galitski, J. Y. Vaishnav, C. W. Clark, and S. Das Sarma, Phys. Rev. A **79**, 053639 (2009); X.-J. Liu, X. Liu, C. Wu, and J. Sinova, Phys. Rev. A **81**, 033622 (2010).
- ⁷⁶ N. Goldman, J. Beugnon, and F. Gerbier, Phys. Rev. Lett. **108**, 255303 (2012).
- ⁷⁷ M. Buchhold, D. Cocks, and W. Hofstetter, Phys. Rev. A **85**, 063614 (2012).
- ⁷⁸ M.-C. Chang and M.-F. Yang, Phys. Rev. B **69**, 115108 (2004).