

Topological insulators and Mott physics from the Hubbard interaction

Stephan Rachel* and Karyn Le Hur†

Department of Physics, Yale University, New Haven, Connecticut 06520, USA

(Received 31 March 2010; published 5 August 2010)

We investigate the Hubbard model on the honeycomb lattice with intrinsic spin-orbit interactions as a paradigm for two-dimensional topological band insulators in the presence of interactions. Applying a combination of Hartree-Fock theory, slave-rotor techniques, and topological arguments, we show that the topological band insulating phase persists up to quite strong interactions. Then we apply the slave-rotor mean-field theory and find a Mott transition at which the charge degrees of freedom become localized on the lattice sites. The spin degrees of freedom, however, are still described by the original Kane-Mele band structure. Gauge-field effects in this region play an important role. When the honeycomb layer is isolated then the spin sector becomes already unstable toward an easy-plane Neel order. In contrast, if the honeycomb lattice is surrounded by extra “screening” layers with gapless spinons, then the system will support a fractionalized topological insulator phase with gapless spinons at the edges. For large interactions, we derive an effective spin Hamiltonian.

DOI: [10.1103/PhysRevB.82.075106](https://doi.org/10.1103/PhysRevB.82.075106)

PACS number(s): 71.30.+h, 71.70.Ej, 73.20.At, 71.10.Fd

I. INTRODUCTION

Topological insulators embody a new class of topological states which have attracted great attention recently.^{1–4} The key for this flourishing development is the understanding that spin-orbit interactions can realize topological insulating phases.^{5–9} The theoretical prediction of such phases in real materials^{10–12} as well as their experimental observations^{13–19} are responsible for the success of this rapidly developing field.

A topological insulator exhibits a bulk energy gap (like ordinary insulators) while the edge (or surface in three dimensions) has gapless states which are protected by time-reversal symmetry. The topological difference between a topological insulator and an ordinary band insulator is characterized by a \mathbb{Z}_2 invariant⁵ which is nonzero in the topological phase. The existence of this topological quantum number as well as the quantized spin Hall conductivity inspired the field,^{20–31} in particular, it was shown that the topological insulator phase—or in two dimensions also called *quantum spin Hall* (QSH) effect—is stable against weak disorder and weak interactions.^{20,26} Inside a topological insulator, Maxwells laws of electromagnetism are altered by an additional topological term with a quantized coefficient, which gives rise to interesting physical effects.^{32,33}

A major role has been played by a simple model introduced by Kane and Mele^{5,6} consisting of a hopping and an intrinsic spin-orbit term on the honeycomb lattice. The Kane-Mele (KM) model (without the Rashba term) essentially consists of two copies with different sign for up and down spins of a model introduced earlier by Haldane.³⁴ Haldane’s pioneering work realizes the quantum Hall effect without an external uniform magnetic field. It breaks, however, time-reversal symmetry (necessary for the quantum Hall effect) which can be restored by taking two copies with different signs for the spins together (as Kane and Mele did). Originally they proposed the model as realization of the QSH effect in graphene^{5,6} and today it should be seen as a paradigm, a perfect theoretical model for topological insulator

phases in two dimensions. The honeycomb lattice is definitely interesting on its own due to the striking development within the graphene community³⁵ but also more exotic phenomena like zero modes^{36,37} have been discovered, for example. The honeycomb lattice has also attracted some attention in relation with exotic phases of light and the Jaynes-Cummings lattice model.³⁸ The additional spin-orbit interactions now make the difference and are responsible for the existence of a topological insulator phase.^{5,6} Consequently, real materials with (strong) spin-orbit interactions have been attracted particular notice.^{39–42} It was also shown that strong nearest- and next-nearest-neighbor repulsions can imitate the intrinsic spin-orbit interactions such that QSH phases are stabilized in the absence of spin-orbit coupling.^{43–45} Beside mercury telluride quantum wells and the Kane-Mele model, topological insulating phases in two dimensions are found to exist in the Kagome lattice⁴⁶ and the decorated honeycomb lattice⁴⁷ provided the presence of spin-orbit interactions.

Other aspects of topological insulators are disorder-induced topological phases as predicted for the HgTe quantum wells^{48,49} and for three-dimensional systems^{50,51} and the proposed existence of axions on the surface of bismuth-tin alloys.³² Axions were postulated more than 30 years ago in the context of the standard model⁵² and their effective action has now been recovered in topological insulators raising hope to detect this dynamical axion field experimentally. Moreover, a QSH phase in ferromagnetic graphene was predicted⁵³ which is protected by the product of charge conjugation and time-reversal symmetry. Most recently, a new family of topological insulators has been discovered^{54,55} in ternary Heusler compounds. Their additional open *f*-shell element might be the key for the realization of exotic topological effects.

Another promising path for the realization of topological phases and, in particular, QSH phases consists of cold atomic gases loaded into optical lattices⁵⁶ which are subjected by a synthetic magnetic field. Such a field has a similar effect on the neutral atoms as a magnetic field coupled to electrons and has been demonstrated experimentally.⁵⁷ A possible ex-

periment to realize a topological insulator was proposed recently.^{58–61} In this spirit, a realization of a topological band insulator (TBI) seems to be feasible in the near future with possibly two major advantages: (i) tuning of the topological insulator band gap or of the details of the engineered Hamiltonian and (ii) availability of onsite interactions (Hubbard model) with tunable interaction strength.

In this paper, we investigate the Hubbard model with intrinsic spin-orbit interactions on the honeycomb lattice which corresponds to the KM model with interactions. Some aspects of the interacting Kane-Mele model were studied in Refs. 62 and 63. A general theory of interaction effects in topological insulators has been proposed introducing a topological order parameter in terms of the full Green's function.⁶⁴ We consider the half-filled case at zero temperature. While the (noninteracting) KM model is known to realize a TBI phase, it is also expected that for sufficiently strong electron-electron interactions magnetic order will take place. Therefore we want to clarify what happens and which phases are present when adding interactions—ranging from very weak to very strong. We focus on the dominant phases at finite spin-orbit coupling. For very weak or no spin-orbit coupling additional (spin liquid) phases might exist but are beyond the scope of this paper. First, we show that interestingly the TBI phase subsists up to quite strong interactions.

Then, applying the slave-rotor mean-field procedure, we investigate the limit of stronger interactions where the charge degrees of freedom form a Mott insulator whereas the spin degrees of freedom are described by a renormalized KM model. In a Mott phase, adding a particle at a given site costs the Hubbard onsite energy U (in contrast, excitations carry a well-defined momentum in the TBI phase). At the mean-field level, this phase has all the properties of a spin liquid (which preserves time-reversal symmetry) with gapless spinon excitations at the edges and is characterized by a hidden order parameter in the spin sector similar to that in the original KM model.^{5,43} On the other hand, one should not underestimate the effect of dynamical compact $U(1)$ gauge fields, especially in two dimensions.⁶⁵

In particular, one predicts⁶⁶ that such a spin liquid phase (with gapless edge spinons) found at the mean-field level can only be stable beyond the mean-field limit if other gapless layers (spinons) are present to screen the gauge field and suppress the gauge fluctuations. Potential candidates can be found in Refs. 67 and 68. Furthermore, Mott physics will suppress the single-particle tunneling at the edges⁶⁹ such that the lowest relevant coupling between layers is the usual spin-spin interaction which may remain irrelevant,⁶⁶ then preserving the gapless edge spinons. Phases exhibiting similar spin-charge separation were also reported in other systems^{40,41,66} and for topological insulators in the presence of a π flux.^{70,71} In contrast, if the honeycomb layer is isolated then the proliferation of instantons will fatally result in a Neel ordering in the XY plane.^{72,73} The two distinct scenarios at the Mott transition are reported in Figs. 1 and 2. Following Ref. 66, fractionalized TI refers to the spin liquid type Mott phase with gapless spinons which preserves time-reversal symmetry and spin-density wave (SDW) in the two figures always refers to the occurrence of a spin-density wave formed in the XY plane.

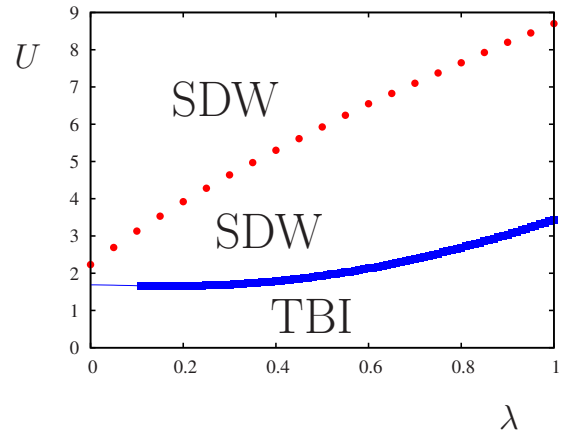


FIG. 1. (Color online) Phase diagram of the isolated honeycomb layer where the proliferation of instantons produces a Neel order in the XY plane already in the entrance of the Mott phase. Above the red dashed line the SDW phase can be described in terms of a mean-field Hartree-Fock theory whereas below the red dashed line the easy-plane Neel order emerges as a result of subtle gauge fluctuations beyond the mean-field solution. (The precise nature of the “transition” associated with the red dashed line is beyond the scope of this paper.)

For very large interactions, applying a conventional Hartree-Fock procedure and deriving an effective spin Hamiltonian, we show that SDW phases with XY ordering are allowed on the honeycomb lattice when adding the spin-orbit term.

The paper is organized as follows. In Sec. II we introduce the KM model, rederive some of its basic properties and introduce the (Hubbard) interaction we consider throughout the paper. In Sec. III we apply the Hartree-Fock method in order to show that a conventional SDW phase with ordering in the XY plane appears at large U . In addition, we derive an effective spin model. Then, in Sec. IV, we use a mean-field approach in momentum space as well as the slave-rotor pic-

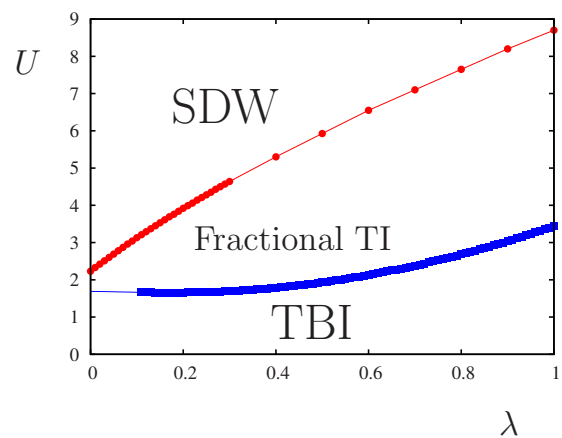


FIG. 2. (Color online) Phase diagram in the presence of additional screening layers (with gapless spinons) allowing to screen the gauge field and therefore stabilize the fractionalized TI phase found at the mean-field level. Here, charge degrees of freedom are in the Mott regime and spin degrees of freedom form a spin liquid with gapless spinons.

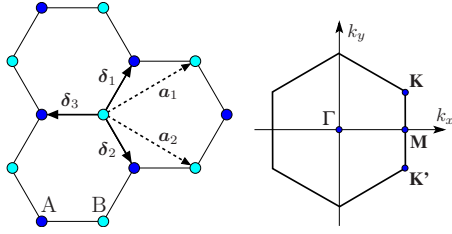


FIG. 3. (Color online) Left: honeycomb lattice consisting of two interpenetrating triangular lattices, A (dark blue dots) and B (cyan dots), with its lattice vectors \mathbf{a}_1 and \mathbf{a}_2 (dashed arrows). In addition, the three nearest-neighbor vectors δ_i ($i=1,2,3$) are shown connecting the two sublattices (solid arrows). Right: corresponding Brillouin zone with the two inequivalent Dirac cones K and K' and the high-symmetry points Γ and M .

ture to argue that the TBI phase as present in the original KM model is stable beyond renormalization-group (RG) results^{6,74} up to moderate interactions. Then, we apply the slave-rotor theory of Florens and Georges⁷⁵⁻⁷⁷ and discuss the intermediate region and the gauge-field effects more thoroughly.

II. MODEL AND GENERAL CONSIDERATIONS

The KM model^{5,6} which might be considered as a spinful version of the Haldane model consists of two parts, a nearest-neighbor hopping term and a second-neighbor hopping spin-orbit term on the honeycomb lattice,

$$\mathcal{H} = -t \sum_{\langle ij \rangle} \sum_{\sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + i\lambda \sum_{\langle\langle ij \rangle\rangle} \sum_{\sigma\sigma'} v_{ij} \sigma_{\sigma\sigma'}^z c_{i\sigma}^{\dagger} c_{j\sigma'}. \quad (1)$$

Here $c_{i\sigma}$ is an electron annihilation operator either on sublattice A or B (then denoted by $a_{i\sigma}$ or $b_{i\sigma}$, respectively) fulfilling the fermionic standard anticommutation relations $\{c_{i\sigma}, c_{j\sigma'}^{\dagger}\} = \delta_{ij} \delta_{\sigma\sigma'}$. As usual t is the hopping integral, λ is the spin-orbit coupling, $\langle ij \rangle$ denotes nearest-neighbor and $\langle\langle ij \rangle\rangle$ next-nearest-neighbor sites, σ^z is the third Pauli matrix, and $v_{ij} = \pm 1$ as discussed below. Throughout the paper we consider the Rashba spin-orbit interaction to be zero. The lattice vectors of the honeycomb lattice are given by

$$\mathbf{a}_1 = \frac{a}{2}(3, \sqrt{3}), \quad \mathbf{a}_2 = \frac{a}{2}(3, -\sqrt{3}) \quad (2)$$

and shown in Fig. 3. The lattice vectors have the length $\sqrt{3}a$ while the lattice spacing a is the distance between neighboring atoms A and B. Note that our notation of the honeycomb lattice is adapted from the review of Castro Neto *et al.*³⁵ We further have the nearest-neighbor vectors

$$\delta_1 = \frac{a}{2}(1, \sqrt{3}), \quad \delta_2 = \frac{a}{2}(1, -\sqrt{3}), \quad \delta_3 = a(-1, 0), \quad (3)$$

which are also shown in Fig. 3. The six next-nearest neighbor vectors δ'_i are given by $\delta'_{1,2} = \pm \mathbf{a}_1$, $\delta'_{3,4} = \pm \mathbf{a}_2$, and $\delta'_{5,6} = \pm(\mathbf{a}_2 - \mathbf{a}_1)$. In what follows we set $a = \hbar = 1$. Throughout the paper N_{Λ} denotes the number of unit cells while N is the number of particles. Hence, the number of lattice sites is $2N_{\Lambda}$

and at half filling $N = 2N_{\Lambda}$. If needed, we refer to the sublattices A and B as Λ_A and Λ_B . The previous definitions imply

$$\sum_{i \in \Lambda} = \sum_{i \in \Lambda_A} = \sum_{i \in \Lambda_B} = \sum_{k \in \text{BZ}} = N_{\Lambda}. \quad (4)$$

As a first step we wish to reproduce the energy bands due to nearest-neighbor hopping and switch to momentum space via

$$c_{i\sigma} = \frac{1}{\sqrt{N_{\Lambda}}} \sum_k e^{ikR_i} c_{k\sigma}, \quad (5)$$

which yields

$$\mathcal{H}_t = -t \sum_{\langle ij \rangle} \sum_{\sigma} (a_{i\sigma}^{\dagger} b_{j\sigma} + \text{H.c.}) \quad (6)$$

$$= \sum_{k\sigma} (a_{k\sigma}^{\dagger}, b_{k\sigma}^{\dagger}) \underbrace{\begin{pmatrix} 0 & -g \\ -g^* & 0 \end{pmatrix}}_{H_k} \begin{pmatrix} a_{k\sigma} \\ b_{k\sigma} \end{pmatrix} \quad (7)$$

$$= \sum_{k\sigma} (l_{k\sigma}^{\dagger}, u_{k\sigma}^{\dagger}) \begin{pmatrix} -|g| & 0 \\ 0 & |g| \end{pmatrix} \begin{pmatrix} l_{k\sigma} \\ u_{k\sigma} \end{pmatrix}. \quad (8)$$

The function g is given by $g \equiv g(\mathbf{k}) = t \sum_{j=1}^3 e^{ik\delta_j}$. Here we used the unitary transformation matrix

$$T_0 = \begin{pmatrix} \frac{1}{\sqrt{2}} \frac{g}{|g|} & \frac{-1}{\sqrt{2}} \frac{g}{|g|} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (9)$$

to diagonalize H_k via $T_0^{\dagger} H_k T_0 = \text{diag}(-|g|, |g|)$. Calculating $\pm|g|$ explicitly results in the well-known tight-binding spectrum of the honeycomb lattice

$$E(\mathbf{k}) = \pm |g| = \pm t [3 + 2 \cos(\sqrt{3}k_y) + 4 \cos(\sqrt{3}k_y/2) \cos(3k_x/2)]^{1/2}, \quad (10)$$

two particle-hole symmetric bands which touch each other at the six corners of the Brillouin zone (BZ) corresponding to two inequivalent points. Expanding around these special points reveals a linear dispersion which gives rise to the name Dirac points. The positions of two inequivalent Dirac points which are shown in Fig. 3 are

$$\mathbf{K} = \left(\frac{2\pi}{3}, \frac{2\pi}{3\sqrt{3}} \right), \quad \mathbf{K}' = \left(\frac{2\pi}{3}, -\frac{2\pi}{3\sqrt{3}} \right). \quad (11)$$

Although it is very convenient to expand around the Dirac points and formulate a Dirac equation on the honeycomb lattice we will keep throughout the paper the full tight-binding model. As a second step, we consider the intrinsic spin-orbit term⁶ of the KM Hamiltonian (1). The expression v_{ij} gives ± 1 depending on the orientation of the sites. A formal definition is

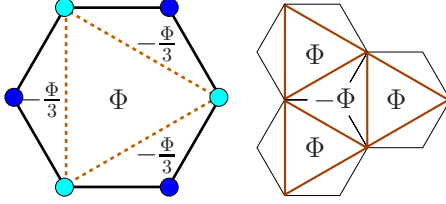


FIG. 4. (Color online) Left: flux configuration per spin and sublattice associated with the intrinsic spin-orbit term. Right: the flux configuration of one sublattice corresponds to a staggered triangular flux lattice.

$$v_{ij} = (\hat{\mathbf{d}}_1 \times \hat{\mathbf{d}}_2)_z, \quad (12)$$

where $\hat{\mathbf{d}}_1$ and $\hat{\mathbf{d}}_2$ are the unit vectors connecting the sites j and i . Essentially, making a left turn yields “−1” while a right turn “+1.” Note that hopping from a site of sublattice A in direction $\hat{\mathbf{d}}_j$ would yield the opposite sign than hopping from a site of sublattice B in the same direction. Hence we should keep in mind that $v_{ij} \propto \tau^z$ (where τ^z is again the third Pauli matrix).

As we will see below the spin-orbit term opens a gap in the bulk. For completeness, we notice that other possible terms which open a gap in the spectrum are different from the spin-orbit term. Such other terms, like a staggered sublattice potential $H_{st} = \sum_i \xi_i c_{i\sigma}^\dagger c_{i\sigma}$, where $\xi_i = 1$ on sublattice A and $\xi_i = -1$ on sublattice B, result in an ordinary band insulator and not in a topological phase since the gap is spin independent. The spin-orbit term preserves the original unit cell. We have shown the corresponding flux configuration per spin and per sublattice in Fig. 4. The net magnetic flux through a plaquette is zero following Haldane’s idea.³⁴ The flux pattern for one of the sublattices corresponds to a triangular staggered flux lattice. Transforming the spin-orbit term to momentum space leads to

$$\begin{aligned} \mathcal{H}_{SO} &= i\lambda \sum_{\langle\langle ij \rangle\rangle} \sum_{\sigma\sigma'} v_{ij} \sigma_{\sigma\sigma'}^z (a_{i\sigma}^\dagger a_{j\sigma'} + b_{i\sigma}^\dagger b_{j\sigma'}) \\ &= 2\lambda \sum_{\sigma\sigma'} \sum_{\mathbf{k}} \sigma_{\sigma\sigma'}^z (a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma'} - b_{\mathbf{k}\sigma}^\dagger b_{\mathbf{k}\sigma'}) \\ &\quad \times \underbrace{(-\sin(\sqrt{3}k_y) + 2 \cos(3k_x/2) \sin(\sqrt{3}k_y/2))}_{\equiv \gamma/2\lambda} \\ &= \sum_{\mathbf{k}} \Psi_{\mathbf{k}}^\dagger \gamma(\mathbf{k}) \sigma^z \tau^z \Psi_{\mathbf{k}}, \end{aligned} \quad (13)$$

where $\Psi_{\mathbf{k}}^\dagger = (a_{\mathbf{k}\uparrow}^\dagger, b_{\mathbf{k}\uparrow}^\dagger, a_{\mathbf{k}\downarrow}^\dagger, b_{\mathbf{k}\downarrow}^\dagger)$ and $\sigma^z \tau^z$ is understood as a 4×4 matrix, $\sigma^z \tau^z = \text{diag}(1, -1, -1, 1)$ (σ^z for spin and τ^z for sublattices).

The bands of the KM model are now obtained by diagonalizing the 4×4 matrix of $\mathcal{H} = \mathcal{H}_t + \mathcal{H}_{SO} = \sum_{\mathbf{k}} \Psi_{\mathbf{k}}^\dagger \mathcal{H}_{\mathbf{k}} \Psi_{\mathbf{k}}$ with the matrix

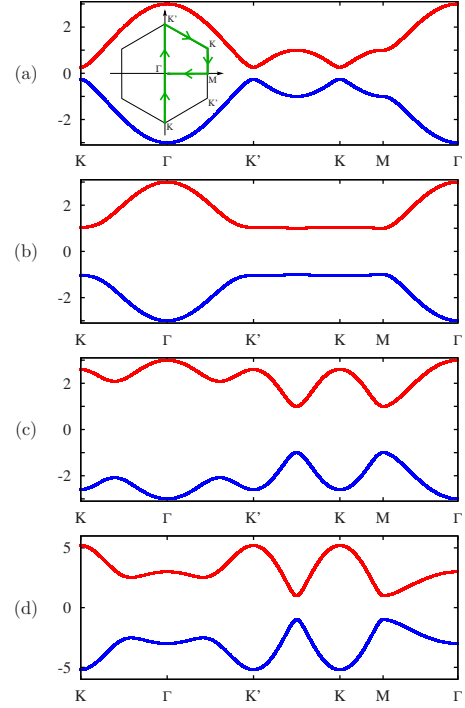


FIG. 5. (Color online) Energy bands ($t=1$) of the Kane-Mele model for (a) $\lambda=0.05$, (b) $\lambda=0.2$, (c) $\lambda=0.5$, and (d) $\lambda=1.0$. The “path” through the Brillouin zone is taken as shown in the inset of (a).

$$\mathcal{H}_{\mathbf{k}} = \begin{pmatrix} \gamma & -g & & \\ -g^* & -\gamma & & \\ & & -\gamma & -g \\ & & -g^* & \gamma \end{pmatrix}. \quad (14)$$

Blank entries should be thought as zeros. As the matrix consists of two independent 2×2 matrices, the diagonalization procedure is identical to Eq. (8) when replacing T_0 by T_\uparrow and T_\downarrow . The exact form of the transformation matrices, T_\uparrow and T_\downarrow , is explicitly given in Sec. IV. The single-particle spectrum of the KM model in the infinite system consists of two twofold-degenerate energy bands (reflecting the Kramers degeneracy),

$$E_{\pm} = \pm \varepsilon(\mathbf{k}) = \pm \sqrt{|g|^2 + \gamma^2}, \quad (15)$$

which are plotted for several values of λ in Fig. 5. The spectrum in Eq. (15) is still particle-hole symmetric. An important feature is that an infinitesimal value of λ opens an infinitesimal gap at the Dirac points. For evaluating the gap size due to the spin-orbit term we know that only the Dirac points \mathbf{K} , \mathbf{K}' as well as the zero-energy lines in γ play a role and it is, hence, sufficient to consider these special points. At the Dirac points, we find

$$\varepsilon(\mathbf{K}) = \varepsilon(\mathbf{K}') = 3\sqrt{3}|\lambda|. \quad (16)$$

At the zero-energy lines of γ , we find (without loss of generality we consider here the line $k_y=0$ only)

$$\varepsilon(k_x, 0) = \sqrt{t^2[5 + 4 \cos(3k_x/2)]} \geq |t|. \quad (17)$$

The minimal value t is reached at the M point of the Brillouin zone. We summarize that the dispersion grows linearly with λ at the Dirac points but immediately when the value of $\lambda = 1/(3\sqrt{3})t$ is reached, the gap remains constant with a gap size

$$\Delta = 2t \quad (\lambda \geq 0.193t). \quad (18)$$

We further consider an ordinary next-nearest-neighbor hopping term without spin-orbit interaction for reasons which will become clear in Sec. V. This term is identical to \mathcal{H}_{SO} when omitting i , λ , and v_{ij} and replacing $\sigma_{\sigma\sigma'}$ by $\delta_{\sigma\sigma'}$. Hence we will find a function g_2 instead of γ ,

$$g_2(\mathbf{k}) = 2 \cos(\sqrt{3}k_y) + 4 \cos(\sqrt{3}k_y/2)\cos(3k_x/2). \quad (19)$$

In contrast to the spin-orbit term, it breaks particle-hole symmetry but does not open a gap at the Dirac points. As already mentioned this term is not present in the KM model (1) but will become relevant in Sec. V.

The aim of this paper is to investigate the effect of a local Hubbard interaction to the KM model: the Hubbard term reads

$$\mathcal{H}_I = U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (20)$$

We consider the case of half filling which allows us in principal to rewrite the Hubbard interaction as

$$\mathcal{H}_I = \frac{U}{2} \sum_i \left(\sum_{\sigma} n_{i\sigma} - 1 \right)^2. \quad (21)$$

This is a particularly convenient formulation for the slave-rotor theory, we will, however, use form (20) as well. Note that Eqs. (20) and (21) are identical at half filling.

III. SDW PHASE AT LARGE U

A. Mean-field arguments

In the past, a transition from the semimetal (Dirac liquid) to a SDW phase has been evidenced in the context of the ordinary Hubbard model ($\lambda=0$) on the honeycomb lattice, when increasing the strength of the onsite interaction. Within the Hartree-Fock approach, this transition takes place⁷⁸ at $\tilde{U}_c = 2.23t$. Within quantum Monte Carlo (QMC) the transition was found⁷⁸ at $\tilde{U}_c \approx 4.5-5t$, and by means of dynamical mean-field theory the transition⁷⁹ occurs even for $\tilde{U}_c > 10t$. While the critical value of the interaction parameter strongly depends on the used method there is no doubt about the existence of the SDW phase for strong interactions. The reason for the occurrence of a SDW phase is simply the bipartite nature of the honeycomb lattice.

In this section, we will apply the Hartree-Fock method to determine the phase boundary $\tilde{U}_c(\lambda)$ at which it becomes favorable to decouple the Hubbard interaction [Eq. (20)] in terms of the sublattice magnetizations $m_i^A = \langle a_{i\uparrow}^\dagger a_{i\uparrow} - a_{i\downarrow}^\dagger a_{i\downarrow} \rangle$ and $m_i^B = \langle b_{i\uparrow}^\dagger b_{i\uparrow} - b_{i\downarrow}^\dagger b_{i\downarrow} \rangle$,

$$\begin{aligned} \mathcal{H}'_I &= \frac{U}{4} \sum_{i \in \Lambda_A \cup \Lambda_B} [(n_{i\uparrow} + n_{i\downarrow})^2 - (n_{i\uparrow} - n_{i\downarrow})^2] \\ &\approx U \sum_{i \in \Lambda_A \cup \Lambda_B} \left[\frac{1}{4} n_i^2 - \frac{1}{2} m_i (n_{i\uparrow} - n_{i\downarrow}) + \frac{1}{4} m_i^2 \right] \\ &= \frac{U}{2} \sum_{i \in \Lambda} [-m^A (n_{i\uparrow}^a - n_{i\downarrow}^a) - m^B (n_{i\uparrow}^b - n_{i\downarrow}^b)] \\ &\quad + \frac{UN_\Lambda}{4} (m^A{}^2 + m^B{}^2) + c, \end{aligned} \quad (22)$$

where $n_i = n_{i\uparrow} + n_{i\downarrow}$, $n_{i\sigma}^a = a_{i\sigma}^\dagger a_{i\sigma}$, and $U \sum_i n_i^2 / 4 = c$ is a constant in the SDW phase. The SDW phase at large U implies a Mott insulating phase for which we can evaluate $\langle n_i^2 \rangle$ explicitly. Indeed, for $U \rightarrow +\infty$, the Mott state is described by the exact wave function

$$|M\rangle = \prod_{i \in \Lambda} a_{i\sigma}^\dagger b_{i\bar{\sigma}}^\dagger |0\rangle, \quad (23)$$

where σ is either \uparrow or \downarrow while $\bar{\sigma}$ ‘‘points’’ in the opposite direction. We find

$$\langle M | n_{i\uparrow}^2 + n_{i\downarrow}^2 + 2n_{i\uparrow} n_{i\downarrow} | M \rangle = 1. \quad (24)$$

Only one of the first two terms contributes depending on the sublattice site i belongs to; the third term is always zero due to the definition of $|M\rangle$. Hence $\langle n_i^2 \rangle = 1$ and c is a constant in the SDW phase. We further assume that $m_i^{(A/B)} = m^{(A/B)}$. In momentum space the mean-field (MF) decoupled Hubbard interaction reads

$$\begin{aligned} \mathcal{H}'_I &= \sum_k \frac{U}{2} (-m^A n_{k\uparrow}^a + m^A n_{k\downarrow}^a - m^B n_{k\uparrow}^b + m^B n_{k\downarrow}^b) \\ &\quad + \frac{UN_\Lambda}{4} (m^A{}^2 + m^B{}^2) + c. \end{aligned} \quad (25)$$

While we could keep m^A and m^B independently, here we will search only for an antiferromagnetic solution. To be more precise we are considering only an Ising-type order parameter. In principal, one could also treat the full spin-rotational problem (e.g., within a σ model representation), we expect, however, no fundamental discrepancies with the simpler approach used here. In order to find the SDW phase we set

$$m \equiv m^A = -m^B. \quad (26)$$

Hence the mean-field Hamiltonian can be written as

$$\mathcal{H}^{\text{HF}} = \sum_k \Psi_k^\dagger [\mathcal{H}_k + Um/2 \text{diag}(-1, 1, 1, -1)] \Psi_k + \frac{U}{2} N_\Lambda m^2. \quad (27)$$

We notice that the mean-field Hamiltonian coincides with the original KM model when replacing γ by $\gamma - Um/2$ up to additional constants. Now we write the mean-field free energy at $T=0$ as

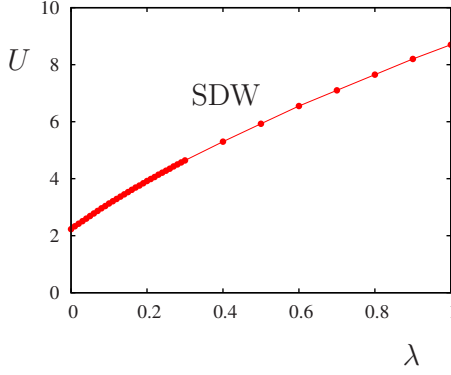


FIG. 6. (Color online) The numerical solution of Eq. (29) is shown ($t=1$). For $\lambda=0$, we confirm that the SDW transition occurs at $\tilde{U}_c=2.23t$ in agreement with the result of Sorella and Tosatti (Ref. 78). With increasing λ , \tilde{U}_c increases up to 8.55 at $\lambda=1.0$. The fact that $\tilde{U}_c(\lambda)$ increases at finite λ may be understood from the effective spin model.

$$F(m) = \sum_k (-2\sqrt{|g|^2 + (\gamma - Um/2)^2}) + \frac{Um^2 N_\Lambda}{2}. \quad (28)$$

Minimizing the free energy with respect to m yields the following self-consistent equation:

$$m = \frac{1}{N_\Lambda} \sum_k \frac{mU/2 - \gamma}{\sqrt{|g|^2 + (\gamma - mU/2)^2}}. \quad (29)$$

The solution of Eq. (29) provides the phase boundary as shown in Fig. 6. For $\lambda=0$ we reproduce the earlier result by Sorella and Tosatti⁷⁸ which is $\tilde{U}_c=2.23t$. With increasing λ we find that $\tilde{U}_c(\lambda)$ increases. The reader may notice that formally we should add a Fermi function in Eqs. (28) and (29) since it may play a role for larger values of m . Here we focus on the phase boundary only which implies small values of m and neglect this point.

For the usual Hubbard model (i.e., $\lambda=0$) on the honeycomb lattice the occurrence of the SDW phase is not surprising since we know that on any bipartite lattice an antiferromagnetic insulator, the SDW phase, will be favored in the limit of large U (at least at half filling). The effective Hamiltonian which describes the low-energy behavior of the Hubbard model for large values of U/t is the antiferromagnetic Heisenberg model. It is, however, rather unclear whether the Mott transition, i.e., the phase transition from a semimetal into a gapped insulator phase, occurs simultaneously with the transition from a disordered spin state into the SDW phase. This unclarity is reflected in a current debate.^{80–85}

While in the early works⁷⁸ no indication was found for two separate phase transitions (Mott-Hubbard and magnetic phase transition), Lee and Lee⁸⁰ reported a possible realization of the nodal spin liquid directly at the Mott transition. While Herbut favored a direct semimetal-SDW transition using a large N approach,⁸¹ Hermele⁸² proposed the stability of the $SU(2)$ algebraic spin liquid in a small region followed by a valence bond solid phase. A recent QMC investigation⁸⁴

rather predicts the existence of a resonating valence bond (RVB) phase between the Dirac-liquid and the SDW phases.

Within the Hartree-Fock procedure presented above we cannot answer the question since the SDW order already implies the Mott phase and does not tell anything about the question where the phase transition into the Mott phase or in a spin liquid phase occurs. It is, however, clear that a possible spin liquid phase must be somewhere below the transition $\tilde{U}_c=2.23t$. The SDW phase is definitely the upper boundary of such a scenario. The fact that $\tilde{U}_c(\lambda)$ increases at finite λ may be understood from the effective spin model which we will discuss now.

B. Effective spin model

In what follows we will investigate the behavior of the spin-orbit term in the strong-coupling limit $U \rightarrow \infty$ with t and λ keeping fixed. Similarly to the usual Hubbard model, we expand the Hamiltonian in powers of t/U . The spin model can be derived in a systematic way as shown, e.g., in Ref. 86 but essentially we have to consider the second-order process of the Hamiltonian $\mathcal{H}_{SO} = \sum_{\langle ij \rangle} (\mathcal{H}_{SO})_{ij}$ with the additional prefactor $-2/U$. The minus sign respects second-order perturbation theory which always lowers the energy,

$$\begin{aligned} (\mathcal{H}_{SO})_{ij} (\mathcal{H}_{SO})_{ji} &= -\nu_{ij} \nu_{ji} \lambda^2 (a_{i\uparrow}^\dagger a_{j\uparrow} - a_{i\downarrow}^\dagger a_{j\downarrow}) (a_{j\uparrow}^\dagger a_{i\uparrow} - a_{j\downarrow}^\dagger a_{i\downarrow}) \\ &= \lambda^2 \left(2S_i^x S_j^x + 2S_i^y S_j^y - 2S_i^z S_j^z - \frac{1}{2} n_i n_j + n_i \right). \end{aligned} \quad (30)$$

Without loss of generality we have considered the hopping process on sublattice A but there is no difference with the equivalent process on sublattice B. At half filling and for $U \rightarrow \infty$, we can assume $n_i \equiv n_{i\uparrow} + n_{i\downarrow} = 1$ and neglect the last terms which are constant. Together with the mentioned factor $-2/U$ we find the effective spin model

$$\tilde{\mathcal{H}} = |J_2| (-S_i^x S_j^x - S_i^y S_j^y + S_i^z S_j^z) \quad (31)$$

with the exchange coupling $J_2 = 4\lambda^2/U$ on a triangular lattice (where we assume that the sum counts every nearest-neighbor pair only once). The spin model then consists of a XY term which favors ferromagnetic order and a Z term which favors antiparallel alignment of the spins. From the nearest-neighbor hopping term we obtain an isotropic antiferromagnetic Heisenberg term with $J_1 = 4t^2/U$. This term stabilizes the antiferromagnetic order, i.e., the SDW. The J_2 term (more precisely, its z component) competes with the J_1 term. This tends to explain the increase in the critical interaction to reach the SDW order in the phase diagram, Fig. 6.

On the other hand, the xy component of the J_2 term favors ferromagnetic order in the XY plane on each sublattice. While the ordering vector might point in any direction when $\lambda=0$, we assume that ordering within the XY plane is preferred immediately when $\lambda \neq 0$. This can also be seen from energetic arguments. Once the ground state is XY ordered, “ \uparrow ” and “ \downarrow ” in $|M\rangle$ refer to the x component of spin. In other words, S_i^z acts like a spin-flip operator. Consequently, any

operator containing S_i^z has a zero expectation value; it particularly implies that for $U \rightarrow +\infty$, $\langle M | \sum_k \gamma \Psi_k^\dagger \sigma^z \tau^z \Psi_k | M \rangle = 0$.

This suggests that the SDW phase (with preferable XY order) might even persist for $J_2 > J_1$ which is beyond the scope of this paper. We conclude this part by making a brief comparison with the ordinary J_1 - J_2 model on the honeycomb lattice with $J_{1/2} > 0$. For weak values of J_2/J_1 , a Néel order is present. For moderate values of $J_2/J_1 \approx 0.3$ a RVB phase was proposed.⁸⁷ For stronger frustrations, a valence bond crystal exhibiting a spin gap is reasonable.⁸⁷

IV. STABILITY OF TBI PHASE

In this section, we provide several arguments establishing that the topological band insulator phase present for $\lambda > 0$ in the original KM model is stable (against Mott physics) not only for weak interactions^{6,8,20,26} but also for moderate interactions $U \sim t$. We first consider a mean-field approach in momentum space which shows that the effect of $U < 2t$ for $\lambda > 0.2t$ does not affect the insulator phase and, hence, should be irrelevant for the topological band insulator phase. Then we introduce the slave-rotor theory of Florens and Georges^{75,76} and argue that this provides a rigorous proof concerning the stability of the TBI phase beyond the perturbative regime. We will discuss both approaches to show that the obtained results are valid beyond the renormalization-group method.⁶ In addition, we will briefly discuss the \mathbb{Z}_2 topological invariant for the present situation.

A. Mean-field arguments

The tight-binding approach starts—after Fourier transformation—with the momentum operators for both sublattices ($a_{k\sigma}, b_{k\sigma}$). Diagonalization of the tight-binding matrix \mathcal{H}_k introduces a new set of operators associated with the bands. As the spin remains a good quantum number [see, e.g., Eq. (14)], we call the new operators $u_{k\sigma}$ and $l_{k\sigma}$, where u and l refer to upper and lower bands.

The explicit transformation matrices between the sublattice basis ($a_{k\sigma}, b_{k\sigma}$) and the band basis ($l_{k\sigma}, u_{k\sigma}$) are given by

$$\begin{pmatrix} a_{k\uparrow} \\ b_{k\uparrow} \end{pmatrix} = \begin{pmatrix} -\alpha_- & -\alpha_+ \\ \beta_- & \beta_+ \end{pmatrix} \begin{pmatrix} l_{k\uparrow} \\ u_{k\uparrow} \end{pmatrix} \equiv T_\uparrow \begin{pmatrix} l_{k\uparrow} \\ u_{k\uparrow} \end{pmatrix} \quad (32)$$

and

$$\begin{pmatrix} a_{k\downarrow} \\ b_{k\downarrow} \end{pmatrix} = \begin{pmatrix} \alpha_+ & \alpha_- \\ \beta_+ & \beta_- \end{pmatrix} \begin{pmatrix} l_{k\downarrow} \\ u_{k\downarrow} \end{pmatrix} \equiv T_\downarrow \begin{pmatrix} l_{k\downarrow} \\ u_{k\downarrow} \end{pmatrix}. \quad (33)$$

We define the functions

$$\alpha_\pm \equiv \alpha_\pm(\mathbf{k}) = \mathcal{N}_\pm d_\pm, \quad (34)$$

$$\beta_\pm \equiv \beta_\pm(\mathbf{k}) = \mathcal{N}_\pm, \quad (35)$$

where

$$d_\pm = \frac{g(\gamma \pm \varepsilon)}{|g|^2} \quad (36)$$

and

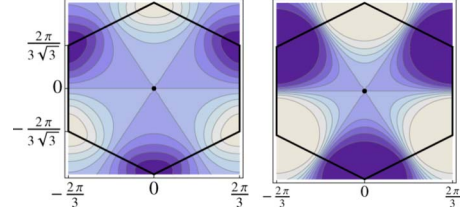


FIG. 7. (Color online) Exemplarily the function $|\alpha_-(\mathbf{k})|^2$ is shown in the BZ for $\lambda/t=0.2$ (left) and $\lambda/t=1.0$ (right). The dark blue region corresponds to a value of zero while the white region to a value of 1. The function $|\alpha_+(\mathbf{k})|^2$ is identical to $|\alpha_-(\mathbf{k})|^2$ by interchange in white and dark blue regions. The black lines mark the boundary of the BZ and the black dot in the center the Γ point.

$$\mathcal{N}_\pm = \frac{|g|}{\sqrt{|g|^2 + (\gamma \pm \varepsilon)^2}}. \quad (37)$$

Note that g , γ , ε and then also d_\pm , α_\pm , β_\pm , and \mathcal{N}_\pm are \mathbf{k} dependent. But for the sake of clarity we omit the \mathbf{k} dependence. To give the reader a better idea about α_\pm and β_\pm , we have plotted $|\alpha_-(\mathbf{k})|^2$ for $\lambda/t=1.0$ and for $\lambda/t=0.2$ in Fig. 7. Let us mention that d_\pm is a complex function and hence α_\pm while β_\pm is real. Nonetheless, to make the equations more symmetric, we will use the complex conjugate of β_\pm as well. Technical aspects and mathematical considerations associated with the change in basis are presented in Appendix A. There, we show useful formulas like $|\alpha_\pm|^2 = |\beta_\mp|^2$ and obtain the important result,

$$\sum_{\mathbf{k} \in \text{BZ}} |\alpha_\pm(\mathbf{k})|^2 = \sum_{\mathbf{k} \in \text{BZ}} |\beta_\pm(\mathbf{k})|^2 = \frac{N_\Lambda}{2}. \quad (38)$$

Now we want to investigate the effect of the Hubbard term on the topological band insulator state more deeply. For that purpose, we transform the interaction term into the band basis. Assuming $\lambda > 0.2t$ we know from Eq. (18) that the gap size $\Delta = 2t$ is large and hence we can neglect all terms containing operators of the upper band (since it costs roughly the energy Δ to make any process between lower and upper bands). Then we decompose the remaining term in a standard way. Eventually the Hubbard term reduces to a chemical-potential term, as we will see, where the chemical potential is given by $U/2$. The whole procedure does not use any additional assumptions and works for any interaction strength U .

$$\mathcal{H}'_I = U \sum_i n_{i\uparrow} n_{i\downarrow}$$

$$\begin{aligned} &= \frac{U}{N_\Lambda} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} a_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger a_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger a_{\mathbf{k}'\downarrow} a_{\mathbf{k}\uparrow} + b_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger b_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger b_{\mathbf{k}'\downarrow} b_{\mathbf{k}\uparrow} \\ &\approx \frac{U}{N_\Lambda} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} [\mathcal{A}(\mathbf{k}, \mathbf{k}', \mathbf{q}) + \mathcal{B}(\mathbf{k}, \mathbf{k}', \mathbf{q})] l_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger l_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger l_{\mathbf{k}'\downarrow} l_{\mathbf{k}\uparrow}, \quad (39) \end{aligned}$$

where we have suppressed all terms containing operators $u_{k\sigma}$ or $u_{k\sigma}^\dagger$ in the last line. The prefactors \mathcal{A} and \mathcal{B} are given by

$$\mathcal{A}(\mathbf{k}, \mathbf{k}', \mathbf{q}) = \alpha_-^*(\mathbf{k} + \mathbf{q}) \alpha_+^*(\mathbf{k}' - \mathbf{q}) \alpha_+(\mathbf{k}') \alpha_-(\mathbf{k}),$$

$$\mathcal{B}(\mathbf{k}, \mathbf{k}', \mathbf{q}) = \beta_-^*(\mathbf{k} + \mathbf{q}) \beta_+^*(\mathbf{k}' - \mathbf{q}) \beta_+(\mathbf{k}') \beta_-(\mathbf{k}). \quad (40)$$

Now the Hubbard term will be decomposed as follows:

$$\begin{aligned} l_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger l_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger l_{\mathbf{k}'\downarrow} l_{\mathbf{k}\uparrow} &\approx \delta_{q0} \langle l_{\mathbf{k}\uparrow}^\dagger l_{\mathbf{k}'\downarrow} \rangle l_{\mathbf{k}'\downarrow} l_{\mathbf{k}\uparrow} + l_{\mathbf{k}\uparrow}^\dagger l_{\mathbf{k}'\downarrow} \delta_{q0} \langle l_{\mathbf{k}'\downarrow}^\dagger l_{\mathbf{k}\uparrow} \rangle \\ &\quad - \delta_{q0} \langle l_{\mathbf{k}\uparrow}^\dagger l_{\mathbf{k}'\downarrow} \rangle \langle l_{\mathbf{k}'\downarrow}^\dagger l_{\mathbf{k}\uparrow} \rangle. \end{aligned} \quad (41)$$

Then the prefactors reduce to

$$\mathcal{A}(\mathbf{k}, \mathbf{k}', \mathbf{q}) \delta_{q0} = |\alpha_-(\mathbf{k})|^2 |\alpha_+(\mathbf{k}')|^2,$$

$$\mathcal{B}(\mathbf{k}, \mathbf{k}', \mathbf{q}) \delta_{q0} = |\beta_-(\mathbf{k})|^2 |\beta_+(\mathbf{k}')|^2. \quad (42)$$

We find the MF-decoupled interaction term in the new basis where we assume $\langle n_{\mathbf{k}\sigma}^l \rangle = \langle l_{\mathbf{k}\sigma}^\dagger l_{\mathbf{k}\sigma} \rangle = 1$ for $\sigma = \uparrow, \downarrow$ since the lower band is completely filled. Therefore, we get

$$\begin{aligned} \mathcal{H}_t' &\approx \sum_{\mathbf{k}\mathbf{k}'} \frac{U}{N_\Lambda} [\mathcal{A}(\mathbf{k}, \mathbf{k}', 0) n_{\mathbf{k}\downarrow}^l + \mathcal{A}(\mathbf{k}', \mathbf{k}, 0) n_{\mathbf{k}\uparrow}^l + \mathcal{B}(\mathbf{k}, \mathbf{k}', 0) n_{\mathbf{k}\downarrow}^l + \mathcal{B}(\mathbf{k}', \mathbf{k}, 0) n_{\mathbf{k}\uparrow}^l - \mathcal{B}(\mathbf{k}, \mathbf{k}', 0) - \mathcal{A}(\mathbf{k}, \mathbf{k}', 0)] \\ &= U \sum_{\mathbf{k}} \left(|\alpha_-(\mathbf{k})|^2 \frac{1}{N_\Lambda} \sum_{\mathbf{k}'} |\alpha_+(\mathbf{k}')|^2 n_{\mathbf{k}\uparrow}^l + |\alpha_+(\mathbf{k})|^2 \frac{1}{N_\Lambda} \sum_{\mathbf{k}'} |\alpha_-(\mathbf{k}')|^2 n_{\mathbf{k}\downarrow}^l \right. \\ &\quad \left. + |\beta_-(\mathbf{k})|^2 \frac{1}{N_\Lambda} \sum_{\mathbf{k}'} |\beta_+(\mathbf{k}')|^2 n_{\mathbf{k}\uparrow}^l + |\beta_+(\mathbf{k})|^2 \frac{1}{N_\Lambda} \sum_{\mathbf{k}'} |\beta_-(\mathbf{k}')|^2 n_{\mathbf{k}\downarrow}^l \right) - UN_\Lambda/2 \\ &= \frac{U}{2} \sum_{\mathbf{k}} \{ [|\alpha_-(\mathbf{k})|^2 + |\beta_-(\mathbf{k})|^2] n_{\mathbf{k}\uparrow}^l + [|\alpha_+(\mathbf{k})|^2 + |\beta_+(\mathbf{k})|^2] n_{\mathbf{k}\downarrow}^l \} = \sum_{\mathbf{k}} \mu^{\text{eff}} n_{\mathbf{k}}^l \end{aligned} \quad (43)$$

with the effective chemical potential $\mu^{\text{eff}} = U/2$ and $n_{\mathbf{k}}^l = n_{\mathbf{k}\uparrow}^l + n_{\mathbf{k}\downarrow}^l$. In the last line of Eq. (43), we have dropped the constant term. For $\lambda > 0.2t$, as the gap always spans from $-t$ to t , we conclude that as long as $U < 2t$ the effective chemical potential lies in the gap, i.e., the U term does not affect the TBI gap. Thus we can assume that the physics in the KM model will be unchanged and consequently the edge modes persist (they will be described by the helical Luttinger liquid theory as a result of interactions²⁴).

B. General slave-rotor arguments

Now we will apply the slave-rotor formalism which allows to address correlated electron systems at weak up to moderate interactions. The method has been introduced by Florens and Georges;^{75,76} we also refer the reader to the review by Zhao and Paramekanti.⁷⁷

Within this approach,^{75–77} the original fermion operators will be rewritten by a product of a fermionic operator $f_{i\sigma}$, the spinon or auxiliary fermion, and a phase factor $e^{i\theta_i}$, the rotor,

$$c_{i\sigma} = e^{i\theta_i} f_{i\sigma}. \quad (44)$$

The idea is that the original fermions c_σ are represented by a collective phase degree of freedom θ (conjugate to the total charge) and auxiliary fermions f_σ . Introducing an additional variable, the angular momentum $L \propto i\partial_\theta$ associated with a quantum O(2) rotor θ , simplifies then the original quartic interaction between the fermions as it is replaced by a simple kinetic term $\propto L^2$. State vectors in the new Hilbert space should have the form $|\Psi\rangle = |\Psi_f\rangle |\Psi_\theta\rangle$. The price we have to pay for the whole rewriting of the original problem is that the new Hilbert space is enlarged compared to the original one

since unphysical states are present. To resolve this problem we have to impose a constraint,

$$\sum_{\sigma} f_{i\sigma}^\dagger f_{i\sigma} + L_i = 1. \quad (45)$$

Since the original fermion operators fulfill anticommutation relations also the spinon (or auxiliary fermion) operators do so. The reader may notice that in the rotor condensate phase the original electron and spinon operators are proportional, and one will find a situation where the spinon band structure describes physical electrons. In contrast, when the rotor is uncondensed, there is spin-charge separation and the spinons are emergent charge-neutral quasiparticles carrying spin only. The term “spinon” should not imply that the particles associated with the new f operators obey fractional statistics in the spirit of the elementary excitation of the one-dimensional Heisenberg antiferromagnet.

Rewriting the hopping term \mathcal{H}_t yields

$$\mathcal{H}_t = -t \sum_{\langle ij \rangle} \sum_{\sigma} (f_{i\sigma}^{a\dagger} f_{j\sigma}^b e^{-i\theta_{ij}} + \text{H.c.}), \quad (46)$$

where $f^{a/b}$ refers to sublattice A/B and $\theta_{ij} \equiv \theta_i - \theta_j$. The spin-orbit term \mathcal{H}_{SO} has a similar form,

$$\mathcal{H}_{\text{SO}} = i\lambda \sum_{\langle\langle ij \rangle\rangle} \sum_{\sigma\sigma'} v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^\dagger f_{j\sigma'} e^{-i\theta_{ij}}. \quad (47)$$

Rewriting the Hubbard term [Eq. (21)] which is local yields

$$\mathcal{H}_I = \frac{U}{2} \sum_i \left(\sum_{\sigma} n_{i\sigma}^f - 1 \right)^2 = \frac{U}{2} \sum_i L_i^2, \quad (48)$$

where we have used the constraint and $n_{i\sigma}^f = f_{i\sigma}^\dagger f_{i\sigma}$.

Without proceeding further, the introduced formalism already allows to read off the following results. The Hubbard interaction U affects the rotor sector only. As long as $U < t$ the rotors will condense favoring the uniform ansatz $\theta_{ij} = 0$. It implies that the auxiliary fermions $f_{i\sigma}$ behave like the original electrons since $\exp(\pm i\theta_{ij}) = 1$ far away from the phase transition. Superfluid or Bose-condensed phases are known to be robust (roughly up to $U \sim t$). We can assume, hence, that the superfluid phase of the rotors also persists against moderate interactions before the phase transition in the insulating phase occurs. Both approaches presented in this section are beyond RG arguments (the perturbative regime). Here, we have chosen to rewrite the Hubbard term in the rotor variables.

In the seminal paper by Kane and Mele⁶ the stability of the TBI phase is briefly discussed. The derived RG equations indicate that additional Coulomb interactions increase the spin-orbit gap size and does not destroy the TBI phase. The RG procedure, however, is only applicable in the perturbative regime where $t \gg U, \lambda$. In contrast, the arguments presented here provide evidence that for $\lambda > 0.2t$ the TBI phase is stable up to $U \sim t$ (against the Mott phase) reaching the strongly interacting regime. As a last point we shall mention that the shown stability of the TBI phase implies also the stability of the edge modes (which are described by the helical Luttinger liquid theory as a result of interactions²⁴).

In Appendix B we present the slave-rotor approach using a simple approximation to decompose Eq. (46). We will restrict ourselves to the case $\lambda = 0$. When performing the mean-field procedure it turns out that the used approximation leads to results which are not so reliable in finite dimensions. Therefore, in Sec. V, we will apply the σ model representation of the slave-rotor theory to find the transition for finite spin-orbit coupling where the rotors undergo a quantum phase transition from superfluid to Mott insulating phase. Before we address this issue, we briefly discuss the stability of the TBI phase in the context of \mathbb{Z}_2 invariants.

C. \mathbb{Z}_2 invariants

For the (noninteracting) KM model considered here, there exists in principle two nontrivial topological invariants, the spin-Chern number of Sheng *et al.*²¹ and the \mathbb{Z}_2 invariant proposed by Kane and Mele.^{5,6} The spin-Chern number can be evaluated by integrating the Berry curvature of a fiber bundle obtained by imposing twisted boundary conditions.²¹ The procedure demonstrated in Ref. 21 implied that the spin-Chern number is a robust topological invariant. Essentially the idea is that the system conserving S^z decouples into two independent Hamiltonians for the up and down spins, each Hamiltonian is characterized by a Chern integer. While the sum of the Chern integers is zero due to time-reversal symmetry, its difference defines a quantized spin Hall conductivity.^{2,21} The \mathbb{Z}_2 invariant is then given by half the difference of the Chern integers (i.e., the spin-Chern number)

modulo 2. Here we will choose another way and calculate the \mathbb{Z}_2 invariant directly for the KM model. We follow Fu and Kane¹¹ and briefly adapt the calculation of the invariant for the sake of completeness.

To compute the \mathbb{Z}_2 invariant it is important to keep the full tight-binding model; in this section, we use the notations of Ref. 11. We express the matrix \mathcal{H}_k of Eq. (14) in terms of Γ matrices,

$$\mathcal{H}_k = \sum_{a=1}^5 d_a(\mathbf{k}) \Gamma^a, \quad (49)$$

where the five Γ^a matrices are given by $\Gamma^1 = \tau^x \otimes I$, $\Gamma^2 = \tau^y \otimes I$, $\Gamma^3 = \tau^z \otimes \sigma^x$, $\Gamma^4 = \tau^z \otimes \sigma^y$, and $\Gamma^5 = \tau^z \otimes \sigma^z$. The coefficients d_1 and d_2 are essentially real and imaginary parts of g [more precise, real and imaginary parts of $\exp(-i\mathbf{k} \cdot \boldsymbol{\delta}_3)g$], respectively, and $d_5 = \gamma$. d_3 and d_4 are both zero. In addition to the five Γ^a matrices, there are also their ten commutators $\Gamma^{ab} = [\Gamma^a, \Gamma^b] / (2i)$. Their anticommutators $\Gamma^a \Gamma^b + \Gamma^b \Gamma^a = 2\delta_{ab}$ obey the Clifford algebra. The parity operator is defined as

$$\mathcal{P} = \tau^x \otimes I = \Gamma^1. \quad (50)$$

Obviously, inversion \mathcal{P} interchanges the sublattices (τ) but not the spin (σ). The time-reversal operator \mathcal{T} is defined by

$$\mathcal{T} = i(I \otimes \sigma^y)K, \quad (51)$$

where K denotes complex conjugation. The Dirac matrices are chosen to be even under \mathcal{PT} , $\mathcal{PT}\Gamma^a(\mathcal{PT})^{-1} = \Gamma^a$ while the commutators are odd under \mathcal{PT} . Note that all Γ^a are odd under \mathcal{P} and \mathcal{T} except Γ^1 which is even. Time-reversal and inversion symmetries imply that the product \mathcal{PT} commutes with the Hamiltonian. The only time-reversal invariant points of the BZ which have to fulfill for a reciprocal-lattice vector $-\Gamma_i = \Gamma_i + \mathbf{G}$ are given by

$$\Gamma_i = \frac{1}{2}(n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2) \quad (52)$$

with $n_i = 0, 1$. We define Γ_1 as the Γ point of the BZ [i.e., $\mathbf{k} = (0, 0)$], $\Gamma_2 = \frac{1}{2}\mathbf{b}_1$, $\Gamma_3 = \frac{1}{2}\mathbf{b}_2$, and $\Gamma_4 = \frac{1}{2}(\mathbf{b}_1 + \mathbf{b}_2)$, where the \mathbf{b}_i are the reciprocal-lattice vectors of the honeycomb lattice as shown in Fig. 8. The latter three points are usually referred to as M points. The \mathbb{Z}_2 invariants characterizing the occupied band are determined¹¹ by

$$\delta_i = -\text{sign}[d_1(\Gamma_i)]. \quad (53)$$

The \mathbb{Z}_2 invariant $\nu = 0, 1$ which distinguishes a topological band insulator in two dimensions from a conventional band insulator is given by the product of all δ_i ,

$$(-1)^\nu = \prod_{i=1}^4 \delta_i. \quad (54)$$

Since we can write $d_1(\mathbf{k}) = t[1 + \cos(\mathbf{k} \cdot \mathbf{a}_1) + \cos(\mathbf{k} \cdot \mathbf{a}_2)]$ and use $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$, we find $d_1(\Gamma_1) = 3t$, $d_1(\Gamma_2) = d_1(\Gamma_3) = t$, and $d_1(\Gamma_4) = -t$. Thus $\delta_1 = \delta_2 = \delta_3 = -1$ and $\delta_4 = +1$ which implies by virtue of Eq. (54)

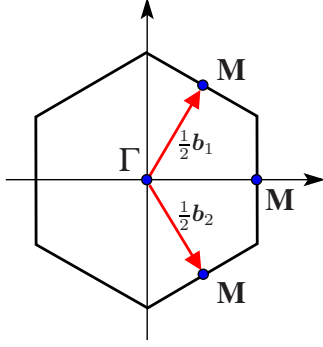


FIG. 8. (Color online) Brillouin zone with the four time-reversal invariant momenta Γ and M . In addition, the reciprocal-lattice vectors (with half of their length) are shown (red arrows).

$$\nu = 1. \quad (55)$$

In fact, a nonzero \mathbb{Z}_2 invariant implies a topological band insulator phase provided there is an energy gap throughout the BZ. In particular, the argument with invariants is a single-particle picture argument.²⁸

As the mean-field interacting Hamiltonian has been reduced to a single-particle Hamiltonian (which leaves d_1 unchanged) the argument based on \mathbb{Z}_2 invariants is applicable as long as the spin-orbit gap is present. The band Hartree-Fock approach has precisely shown that the Hubbard term behaves as a chemical potential which lies between the bands as long as $|U| < 2t$. Thus we have substantiated our earlier statement that the TBI phase and the presence of edge modes will be stable up to a region beyond the perturbative regime.

V. MOTT TRANSITION

Now we will use the slave-rotor mean-field theory^{75,76} to find the transition where the charge degrees of freedom form a Mott insulating state (and not a band insulator).

A. Self-consistency equations

In contrast to the approximation in Appendix B we will use a more sophisticated approach^{75,76,80} in order to find the transition to the Mott phase. We replace the phase field representing the $O(2)$ degree of freedom by a complex bosonic field $X = e^{i\theta}$ which is constrained by

$$|X(\tau)|^2 = 1. \quad (56)$$

The associated Lagrange multiplier is called ρ . The derivation of the Lagrangian and the decomposition of hopping and spin-orbit terms is shown in Appendix C. The mean-field parameters associated with the decomposition are given by

$$Q_X = \left\langle \sum_{\sigma} f_{i\sigma}^{a*} f_{j\sigma}^b \right\rangle, \quad (57)$$

$$Q_f = \langle \exp(-i\theta_{ij}) \rangle, \quad (58)$$

for the hopping term and

$$Q'_X = \left\langle \sum_{\sigma\sigma'} i\nu_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^* f_{j\sigma'} \right\rangle, \quad (59)$$

$$Q'_f = \langle \exp(-i\theta_{ij}) \rangle, \quad (60)$$

for the spin-orbit term. Finally we find the imaginary time Green's function for the f_{σ}^f fields,

$$G_{f\sigma} = \frac{1}{i\omega_n - \Sigma_k}, \quad (61)$$

and for the X fields,

$$G_X = \frac{1}{\nu_n^2 U + \rho + \xi_k}. \quad (62)$$

Here we introduced the renormalized KM spectrum for the spinon sector,

$$\Sigma_k = \sqrt{(Q_f |g|)^2 + (Q'_f \gamma)^2}, \quad (63)$$

and defined

$$\xi_k = -Q_X |g(\mathbf{k})| + Q'_X \lambda g_2(\mathbf{k}), \quad (64)$$

with g_2 from Eq. (19). Note that we consider only the half-filled case here which allowed us to set $\mu = h = 0$. In the absence of spin-orbit coupling, $\lambda \rightarrow 0$, we find $\Sigma_k \rightarrow Q_f |g|$ and $\xi_k \rightarrow -Q_X |g|$ and recover, hence, the Green's function from Florens and Georges⁷⁶ (when setting $\epsilon \equiv -|g|$). From there, we find directly the five self-consistency equations,

$$1 = \frac{1}{N_{\Lambda}} \sum_{\mathbf{k}} \frac{1}{\beta} \sum_n G_X(\mathbf{k}, i\nu_n) = \frac{U}{N_{\Lambda}} \sum_{\mathbf{k}} \frac{1}{\sqrt{\Delta_g^2 + 4U[\xi - \min(\xi_k)]}}. \quad (65)$$

In Eq. (65) we performed the evaluation of the Matsubara sum at zero temperature (see Appendix D) and we introduced the insulating gap

$$\Delta_g = 2\sqrt{U[\rho + \min(\xi_k)]}. \quad (66)$$

While Δ_g is nonzero in the insulating phase, directly at the phase transition it will vanish since the rotors condense. Before we can use Eq. (65) to find the transition line, we have to know the explicit form of ξ_k and hence of Q_X and Q'_X . The latter two mean-field parameters are determined by use of the (second and third) self-consistency equations. We start with Q_X ,

$$\begin{aligned} t \sum_{j=1}^3 \sum_{\sigma} \langle f_{i\sigma}^{a\dagger} f_{j\sigma}^b \rangle &= \frac{1}{N_{\Lambda}} \sum_{\mathbf{k}\sigma} g(\mathbf{k}) \langle f_{k\sigma}^{a\dagger} f_{k\sigma}^b \rangle \\ &= \frac{1}{N_{\Lambda}} \sum_{\mathbf{k}} g(-\alpha_-^* \beta_- + \alpha_+^* \beta_+) \langle f_{k\sigma}^{a\dagger} f_{k\sigma}^b \rangle \\ &= \frac{1}{N_{\Lambda}} \sum_{\mathbf{k}} \frac{|g|^2}{\sqrt{|g|^2 + \gamma^2}}. \end{aligned} \quad (67)$$

Due to the lattice symmetry, the sum over the three-nearest neighbors, $\sum_{j=1}^3$, just appears as a factor 3 in the final expression. Thus we find the mean-field parameter

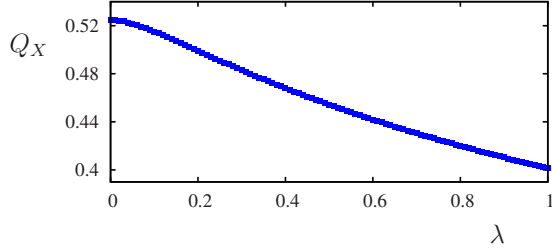


FIG. 9. (Color online) Numerical solution of the mean field Eq. (68). The behavior of $Q_X(\lambda)$ is shown ($t=1$).

$$Q_X = \frac{1}{3tN_\Lambda} \sum_k \frac{|g|^2}{\sqrt{|g|^2 + \gamma^2}}. \quad (68)$$

We have plotted Q_X as a function of λ in Fig. 9.

In a similar way we proceed in order to find Q'_X ,

$$\begin{aligned} \lambda \sum_{j=1}^6 Q'_X &= \left\langle \lambda \sum_j \sum_{\sigma\sigma'} i\nu_{ij} \sigma^z f_{i\sigma}^* f_{j\sigma} \right\rangle \\ &= \frac{1}{N_\Lambda} \sum_k \gamma \langle \Psi_k^\dagger \sigma^z \tau^x \Psi_k \rangle \\ &= \frac{1}{N_\Lambda} \sum_k \gamma (|\alpha_-|^2 - |\alpha_+|^2 - |\beta_-|^2 + |\beta_+|^2) \\ &= -\frac{1}{N_\Lambda} \sum_k \frac{2\gamma^2}{\sqrt{|g|^2 + \gamma^2}}. \end{aligned} \quad (69)$$

Again the lattice symmetry is responsible for the fact that the sum over the next-nearest neighbors, $\sum_{j=1}^6$, can be replaced by a factor 6. Then the self-consistency equation reads

$$\begin{aligned} Q'_X &= \left\langle \sum_{\sigma\sigma'} i\nu_{ij} \sigma^z f_{i\sigma}^* f_{j\sigma} \right\rangle \\ &= -\frac{1}{3\lambda N_\Lambda} \sum_k \frac{\gamma^2}{\sqrt{|g|^2 + \gamma^2}} = -|Q'_X(\lambda)|. \end{aligned} \quad (70)$$

We have plotted Q'_X as a function of λ in Fig. 10. With the knowledge of Q_X and Q'_X finally the rotor spectrum ξ_k of Eq. (64) is well defined and we can proceed with Eq. (65). If one

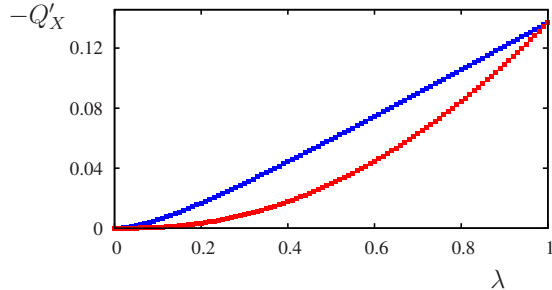


FIG. 10. (Color online) Numerical solution of $-\lambda Q'_X$ (red curve) and $-Q'_X$ (blue curve) as a function of λ ($t=1$) is shown.

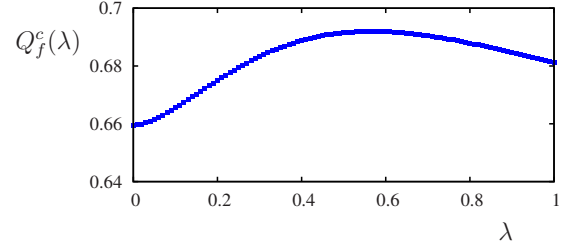


FIG. 11. (Color online) Numerical solution of $Q_f^c(\lambda)$, i.e., Q_f along the line $U_c(\lambda)$, as a function of λ ($t=1$) is shown.

moves toward the transition from the Mott insulator to the superfluid phase of the rotors, the rotor gap Δ_g must close. It yields

$$U_c(\lambda) = \left[\frac{1}{2N_\Lambda} \sum_{k'} \frac{1}{\sqrt{\xi_k - \min(\xi_k)}} \right]^{-2}, \quad (71)$$

which defines the transition line between TBI and the EMI phase as shown in the phase diagram, Fig. 1. The sum over k' means that formally the lowest bound corresponds to $k \rightarrow k_{\min} + \eta$, $\eta \ll 1$, and k_{\min} is associated with the minimum of ξ_k . Hence, no divergence appears in the sum. A formal justification to cut the sum can be given by switching to “energy space” and considering the density of states. The same line of argument applies to Eqs. (73) and (75).

As a last point we have to consider Q_f and Q'_f and its behavior along the line $U_c(\lambda)$. Applying the same line of reasoning as for Q_X we directly find

$$\begin{aligned} Q_f &= \langle X_i^* X_j \rangle |_{ijmn} = \frac{1}{N_\Lambda} \sum_k e^{ik\delta_\mu} \langle X_k^{a*} X_k^b \rangle \\ &= \frac{1}{N_\Lambda} \sum_k \frac{|g|}{6t} \frac{1}{\beta} \sum_n G_X(\mathbf{k}, i\nu_n) \\ &= \frac{1}{N_\Lambda} \sum_k \frac{|g|}{6t} \frac{U}{2\sqrt{U(\rho + \xi_k)}}. \end{aligned} \quad (72)$$

Here δ_μ denotes one of the three nearest-neighbor vectors. Along the transition line we have $\Delta_g=0$ and obtain

$$Q_f^c(\lambda) = \frac{\sqrt{U_c(\lambda)}}{6tN_\Lambda} \sum_{k'} \frac{|g|}{\sqrt{\xi_k - \min(\xi_k)}}. \quad (73)$$

It turns out that Q_f^c is a slowly varying function of λ . We have plotted it in Fig. 11.

The last self-consistency equation determines Q'_f .

$$\begin{aligned} Q'_f &= \langle X_i^* X_j \rangle |_{ijnmn} \\ &= \frac{1}{N_\Lambda} \sum_k e^{ik\delta'_\mu} \langle X_k^{(a/b)*} X_k^{(a/b)} \rangle \\ &= \frac{1}{N_\Lambda} \sum_k e^{ik\delta'_\mu} \frac{1}{2\beta} \sum_n G_X(\mathbf{k}, i\nu_n) \end{aligned}$$

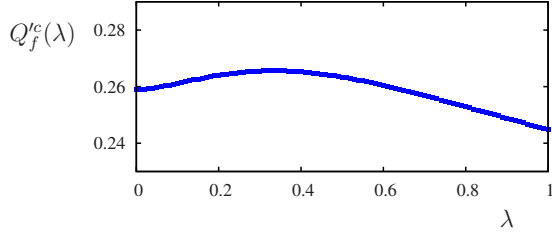


FIG. 12. (Color online) Numerical solution of $Q_f^c(\lambda)$ as a function of λ ($t=1$) is shown.

$$= \frac{U}{2N_\Lambda} \sum_k e^{ik\delta'_\mu} \frac{1}{2\sqrt{U(\rho + \xi_k)}}, \quad (74)$$

with δ'_μ being one of the six next-nearest-neighbor vectors. Thus we find Q_f^c along the Mott transition,

$$Q_f^c(\lambda) = \frac{\sqrt{U_c(\lambda)}}{2N_\Lambda} \sum_{k'} \frac{e^{ik\delta'_\mu}}{\sqrt{\xi_k - \min(\xi_k)}}, \quad (75)$$

which we have plotted in Fig. 12. From Figs. 11 and 12 we see that Q_f^c and Q_f^c behave similarly but Q_f^c is (roughly) three times larger than Q_f^c .

B. Discussion

In Fig. 13, we show the spectrum of the noninteracting KM model for $\lambda=0.2$ as well as the renormalized spinon spectrum for $U_c(\lambda=0.2)$. The geometry we considered is a stripe with 14 unit cells in y direction while the stripe is infinitely long in x direction. Here we see that interactions contribute through Q_f and Q_f^c such that the bulk spin gap is decreased compared to the TBI phase.

At the *mean-field* level, spin-charge separation will occur and, while the charge is frozen in the Mott insulating state, the spin degrees of freedom will exhibit a Hamiltonian reminiscent of the KM model. In particular, this implies the existence of gapless edge spinons. In this sense this gives rise to the fractionalized TI mentioned in Sec. I (Ref. 66) or a “topological Mott insulator” as Pesin and Balents⁴¹ did. Note that the topological Mott insulator phase has a different

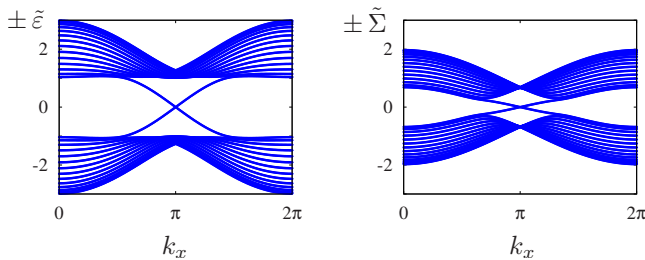


FIG. 13. (Color online) Left: spectrum $\pm\tilde{\varepsilon}(k_x)$ of the KM model on a stripe geometry as explained in the text for $\lambda=0.2$ and $t=1$. Right: spectrum $\pm\tilde{\Sigma}(k_x)$ of the spinon sector for the same parameters at the critical line $U_c(\lambda=0.2)$. The hopping and spin-orbit amplitudes are renormalized with $Q_f^c=0.68$ and $Q_f^c=0.26$. The tildes refer to the finite stripe geometry in contrast to the spectra of the infinite system.

meaning than in the work of Raghu *et al.*⁴³ where the topological band insulator phase was caused by strong interactions. On the other hand, as already mentioned in Sec. I, $U(1)$ gauge fields associated with the lattice theory,⁸⁰ see also Appendix E, cannot be ignored in two dimensions. In particular, to stabilize the fractionalized TI beyond the mean-field level, one requires extra layers supporting gapless spinons allowing to screen the gauge field.⁶⁶ (The stability of spinon excitations at the edges results from the fact that single-particle tunneling is suppressed as a result of Mott physics.⁶⁹) This also implies that in the context of an isolated (single) honeycomb layer the fractionalized TI is unstable to instanton proliferation and to easy-plane Neel ordering.^{72,73}

Let us assume that conditions are realized such that the fractionalized TI is stable against gauge fluctuations. Then, we can be even more precise when focusing on the spinon bulk sector. We can write for the corresponding ground-state wave function,

$$|\Psi_f\rangle = \prod_{k\sigma=\uparrow,\downarrow} f_{k\sigma}^\dagger |0\rangle, \quad (76)$$

i.e., the lower band $-\sum_k$ [see Eq. (63)] is completely filled. The explicit knowledge of $|\Psi_f\rangle$ allows us to calculate the expectation value of the z component of spin,

$$\begin{aligned} \langle \Psi_f | \frac{1}{2} (n_{i\uparrow}^f - n_{i\downarrow}^f) | \Psi_f \rangle &= \frac{1}{2N_\Lambda} \sum_k \langle \Psi_f | (|\alpha_-|^2 f_{k\uparrow}^\dagger f_{k\uparrow}^f - |\alpha_+|^2 f_{k\downarrow}^\dagger f_{k\downarrow}^f) | \Psi_f \rangle \\ &= \frac{1}{2N_\Lambda} \sum_k (|\alpha_-|^2 - |\alpha_+|^2) = 0. \end{aligned} \quad (77)$$

In the last line we have used Eq. (38). In the same way, we can easily check that $\langle S_i^x \rangle = 0$. At the mean-field level, we can also calculate the spin-spin correlation functions $\langle S_i^+ S_j^- \rangle$ for i and j on the same sublattice or on different sublattices. We show the former case explicitly,

$$\begin{aligned} \langle S_i^+ S_j^- \rangle &= \langle \Psi_f | f_{i\uparrow}^{a\dagger} f_{i\downarrow}^{a\dagger} f_{j\downarrow}^a f_{j\uparrow}^a | \Psi_f \rangle \\ &= \left(\frac{1}{N_\Lambda} \sum_{k_1} e^{-ik_1(\mathbf{R}_i - \mathbf{R}_j)} |\alpha_-(\mathbf{k}_1)|^2 \right) \\ &\quad \times \left(\frac{1}{N_\Lambda} \sum_{k_2} e^{ik_2(\mathbf{R}_i - \mathbf{R}_j)} |\alpha_+(\mathbf{k}_2)|^2 \right), \end{aligned} \quad (78)$$

which we can evaluate numerically and find that the correlations decay to zero on very short distances. In fact, when the distance $|\mathbf{R}_i - \mathbf{R}_j|$ reaches roughly four unit cells, the correlations are already smaller than 10^{-6} . A similar calculation for both i and j on sublattice B as well as i and j on different sublattices reveals comparable results. We expect that the one-dimensional character of spinon excitations at the edges gives rise to power-law spin-correlation functions. Other exotic spin liquid phases may be found in the vicinity of a Mott state.⁸⁸

The only ‘‘hidden’’ order which seems to be present is reflected in $\langle \Psi_f | \sum_k \gamma \Phi_k^\dagger \sigma^z \tau^z \Phi_k | \Psi_f \rangle \neq 0$ being a remnant of the TBI phase; here we have introduced the corresponding vector $\Phi_k^\dagger = (f_{k\uparrow}^{a\dagger}, f_{k\uparrow}^{b\dagger}, f_{k\downarrow}^{a\dagger}, f_{k\downarrow}^{b\dagger})$.

In the phase diagram of Figs. 1 and 2 the region above the TBI phase has to be handled with care for $0 < \lambda < 0.1t$ and is beyond the scope of this paper. In particular, in the absence of spin-orbit coupling ($\lambda=0$), mean-field slave-rotor techniques predict a Mott insulator with gapless spin excitations.⁸⁰ In the limit $\lambda \rightarrow 0$ we recover the earlier result of Lee and Lee,⁸⁰ i.e., $U_c = 1.68t$. On the other hand, in a recent QMC study it was shown⁸⁴ that the intermediate phase for $\lambda=0$ is a RVB spin liquid in contrast to Refs. 80 and 82.

While we have clarified the question what the effect of a Hubbard onsite interaction is one could also consider nearest- and next-nearest-neighbor repulsion with amplitudes V_1 and V_2 . Such a model in absence of spin-orbit coupling was investigated by Raghu *et al.*⁴³ From the band Hartree-Fock approach presented in Sec. IV we see, however, that the effect for small V_1 and V_2 is negligible. This is in correspondence with Ref. 43 where strong nearest- and next-nearest-neighbor interactions are required to reach QSH phases while weak interactions leave the semimetal unchanged. As the intrinsic spin-orbit interaction already opened a gap, V_1 and V_2 should be of the same order to affect the phase diagram. As large values of V_1 and V_2 are in the model under consideration somewhat unphysical, we can conclude that additional weak nearest- and next-nearest-neighbor interactions are negligible for the KM model.

VI. CONCLUSION

We have investigated the Kane-Mele model in the presence of a Hubbard interaction as a paradigm for two-dimensional topological insulators with interactions. Using a mean-field procedure and arguments from the slave-rotor theory, we have shown that the TBI phase characterized by a Z_2 topological invariant is stable (against Mott phases) up to moderate interactions which are beyond the perturbative regime. The topological band insulator phase is separated from a Mott insulating region through $U_c(\lambda)$.

At the mean-field level, charge constituents become frozen in the Mott state while the spin constituents form a quantum spin liquid with gapless edge spinons (preserving the time-reversal symmetry). Even though this fractionalized TI phase is unstable against gauge fluctuations in the isolated honeycomb lattice system,^{72,73} the vicinity of other screening layers exhibiting gapless spinon excitations⁶⁶ would allow to stabilize such a phase of matter in (quasi-)two-dimensional systems. For very large onsite interactions, the fractionalized TI phase inevitably turns into a SDW phase with XY ordering.

For very weak spin-orbit interactions, other insulating phases reminiscent of the ‘‘gapless Mott insulator’’ phase of Pesin and Balents⁴¹ might exist. It remains an open question if such a possible phase might be connected with the expected spin liquid phase for $\lambda \rightarrow 0$.

ACKNOWLEDGMENTS

We acknowledge discussions with L. Balents, M. Her-

mele, Y.-B. Kim, J. Moore, G. Murthy, and R. Shankar. This work was supported by NSF under Grant No. DMR-0803200 and by the Yale Center for Quantum Information Physics (Grant No. DMR-0653377). K.L.H. also acknowledges support from CNRS at LPS Orsay, Grant No. UMR 8502, and S.R. from the Deutsche Forschungsgemeinschaft (DFG) under Grant No. RA 1949/1-1.

APPENDIX A: BAND BASIS

In this appendix, we present some supplementary material concerning the band basis which was introduced in Sec. IV. First, let us check that $T_\uparrow^\dagger T_\uparrow = T_\uparrow T_\uparrow^\dagger = 1$ and $T_\downarrow^\dagger T_\downarrow = T_\downarrow T_\downarrow^\dagger = 1$. We find the diagonal elements $|\alpha_\pm|^2 + |\beta_\pm|^2 = 1$ which can be easily checked. The off-diagonal elements are given by

$$\begin{aligned} \alpha_-^* \alpha_+ + \beta_-^* \beta_+ &= \mathcal{N}_- \mathcal{N}_+ (d_-^* d_+ + 1) \\ &= \mathcal{N}_- \mathcal{N}_+ \left(\frac{\gamma^2 - \varepsilon^2 + |g|^2}{|g|^2} \right) = 0. \end{aligned} \quad (\text{A1})$$

We will further need the following expressions:

$$\mathcal{N}_+ \mathcal{N}_- = \frac{|g|}{2\varepsilon} \quad \text{and} \quad \mathcal{N}_\pm^2 = \frac{|g|^2}{2(\varepsilon^2 \pm \gamma\varepsilon)}. \quad (\text{A2})$$

In the above transformations the limit $g \rightarrow 0$ should be handled with care as the original eigenvectors diverge. This is a consequence of the fact that the matrix \mathcal{H}_k is already diagonal for $t=0$. To consider the case with $\lambda=0$, we recover the transformation matrix (9) from Sec. II,

$$\lim_{\gamma \rightarrow 0} T_\uparrow = \lim_{\gamma \rightarrow 0} T_\downarrow = T_0. \quad (\text{A3})$$

In order to prepare the following part of the section, we have to show explicitly that $\sum_k |\alpha_\pm(\mathbf{k})|^2 = \sum_k |\beta_\pm(\mathbf{k})|^2 = N_\Lambda/2$. One can show that $|\alpha_\pm|^2 = |\beta_\mp|^2$ (which implies $|\alpha_+|^2 + |\alpha_-|^2 = 1$) and it is sufficient to consider $|\alpha_\pm|^2$ in the following. First we show that $|\alpha_\pm(\mathbf{k})|^2 + |\alpha_\pm(-\mathbf{k})|^2 = 1$ and then we argue that the result follows directly. By using $|g(-\mathbf{k})| = |g(\mathbf{k})|$, $\varepsilon(-\mathbf{k}) = \varepsilon(\mathbf{k})$, and $\gamma(-\mathbf{k}) = -\gamma(\mathbf{k})$ we could calculate $|\alpha_\pm(\mathbf{k})|^2 + |\alpha_\pm(-\mathbf{k})|^2 = 1$ explicitly. This is not necessary since the only thing we have to show is

$$|\alpha_-(\mathbf{k})|^2 = |\alpha_+(-\mathbf{k})|^2. \quad (\text{A4})$$

By looking at the definition of α_\pm and using $\gamma(-\mathbf{k}) = -\gamma(\mathbf{k})$, Eq. (A4) turns out to be correct. Now we have to divide the BZ into two parts, e.g., as follows: $\text{BZ} = \mathcal{K}_1 \cup \mathcal{K}_2$ with $\mathcal{K}_1 = \{\mathbf{k} \in [-\frac{2\pi}{3}, 0] \times [-\frac{4\pi}{3\sqrt{3}}, 0]\}$ and $\mathcal{K}_2 \in [0, \frac{2\pi}{3}] \times [-\frac{4\pi}{3\sqrt{3}}, 0]$ while \mathcal{K}_2 contains the remainder of the BZ such that $\mathcal{K}_1 \cap \mathcal{K}_2 = 0$. This ensures that $\mathbf{k} \in \mathcal{K}_1$ implies $-\mathbf{k} \in \mathcal{K}_2$ and vice versa.

Hence we can split the sum over the BZ as

$$\begin{aligned} \sum_{\mathbf{k} \in \text{BZ}} |\alpha_\pm(\mathbf{k})|^2 &= \sum_{\mathbf{k} \in \mathcal{K}_1} |\alpha_\pm(\mathbf{k})|^2 + \sum_{\mathbf{k} \in \mathcal{K}_2} |\alpha_\pm(\mathbf{k})|^2 \\ &= \sum_{\mathbf{k} \in \mathcal{K}_1} (|\alpha_\pm(\mathbf{k})|^2 + |\alpha_\pm(-\mathbf{k})|^2) \\ &= \sum_{\mathbf{k} \in \mathcal{K}_1} 1 = N_\Lambda/2. \end{aligned} \quad (\text{A5})$$

APPENDIX B: SIMPLE SLAVE-ROTOR MEAN-FIELD THEORY

In this appendix, we perform a simple mean-field decomposition for the KM model with Hubbard interactions which is rewritten in spinons and rotors. Starting from Eqs. (46) and (47), adding the rewritten interaction term $H_I = U/2 \sum_i L_i^2$, we assume that state vectors in the Hilbert space should have the form $|\Psi\rangle = |\Psi_f\rangle |\Psi_\theta\rangle$. Decoupling the rotor and fermion variables and treating the constraint by introducing a Lagrange multiplier h , we obtain effective Hamiltonians for spinon and rotor sector,

$$\mathcal{H}^f = - \sum_{\langle ij \rangle \sigma} t_{ij}^{\text{eff}} f_{i\sigma}^\dagger f_{j\sigma}^b + i \sum_{\langle\langle ij \rangle\rangle} \lambda_{ij}^{\text{eff}} \sum_{\sigma\sigma'} v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^\dagger f_{j\sigma'} + (\mu + h) \sum_{i\sigma} f_{i\sigma}^\dagger f_{i\sigma}, \quad (\text{B1})$$

$$\mathcal{H}^\theta = - \sum_{\langle ij \rangle} \mathcal{J}_{ij}^{\text{eff}} \cos(\theta_{ij}) + \sum_{\langle\langle ij \rangle\rangle} \mathcal{G}_{ij}^{\text{eff}} \cos(\theta_{ij}) + \sum_i \left(\frac{U}{2} L_i^2 + h L_i \right). \quad (\text{B2})$$

In the spinon sector we recover the original KM model expressed in f_σ operators rather than c_σ operators while the rotor sector corresponds to a quantum-XY-like model for the phase variables. The effective amplitudes are determined by the following self-consistent equations:

$$t_{ij}^{\text{eff}} = t \langle \cos(\theta_{ij}) \rangle_\theta, \quad (\text{B3})$$

$$\lambda_{ij}^{\text{eff}} = \lambda \langle \cos(\theta_{ij}) \rangle_\theta, \quad (\text{B4})$$

$$\mathcal{J}_{ij}^{\text{eff}} = t \sum_{\sigma} \langle f_{i\sigma}^\dagger f_{j\sigma}^b \rangle_f, \quad (\text{B5})$$

$$\mathcal{G}_{ij}^{\text{eff}} = \lambda \sum_{\sigma\sigma'} \langle i v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^\dagger f_{j\sigma'} \rangle_f. \quad (\text{B6})$$

The expectation values are taken with respect to $|\Psi_f\rangle$ or $|\Psi_\theta\rangle$, respectively. In the following, we show the main steps in finding the self-consistency equations and the phase transition within this simple approach. In fact, it is the analogous calculation for the honeycomb lattice to Secs. IIIB2 and IIIB3 of Ref. 76 for cubic lattices. We restrict ourselves to the case $\lambda=0$; the Hamiltonian \mathcal{H}^f reads in momentum space

$$\mathcal{H}^f = - \sum_{k\sigma} (f_{k\sigma}^\dagger f_{k\sigma}^b) \begin{pmatrix} \mu + h & Zg \\ Zg^* & \mu + h \end{pmatrix} \begin{pmatrix} f_{k\sigma}^a \\ f_{k\sigma}^b \end{pmatrix}. \quad (\text{B7})$$

Here we decomposed the expectation value simply as $\langle \cos(\theta_i - \theta_j) \rangle \approx \langle \cos \theta \rangle^2 = Z$. By diagonalization of the matrix in Eq. (B7) we obtain the upper and lower bands,

$$\epsilon_{\pm} = \pm Z|g| - (\mu + h). \quad (\text{B8})$$

We stress again that our constraint is differently chosen compared to Ref. 76 where it is given by $\sum_{\sigma} f_{i\sigma}^\dagger f_{i\sigma} - L = 1$. In the constraint used in this paper the L term appears with a positive sign. The difference is due to the different definition of the fermions in term of spinons and rotors, see Eq. (6) of

Ref. 76. Now, the Lagrange multiplier h is defined from the following constraint equation:

$$\langle L \rangle_\theta = - \sum_{\sigma} \left(\langle f_{i\sigma}^\dagger f_{i\sigma} \rangle_f - \frac{1}{2} \right). \quad (\text{B9})$$

To proceed further, we treat \mathcal{H}^θ , which corresponds to a quantum-XY model, at the mean-field level. The applied approximation, $\cos(\theta_{ij}) \approx 2 \cos(\theta_i) \langle \cos(\theta_j) \rangle - \text{const}$, reduces the rotor Hamiltonian to a mean-field Hamiltonian of independent sites,

$$H_{\text{MF}}^\theta = H_0 + H_I = \sum_i \left(\frac{U}{2} L_i^2 - h L_i \right) + \sum_i K \cos \theta_i. \quad (\text{B10})$$

Here the coupling constant K is given by

$$K = -2t \sum_{\sigma} \sum_j \langle f_{i\sigma}^\dagger f_{j\sigma}^b \rangle_f \langle \cos \theta_j \rangle_\theta. \quad (\text{B11})$$

As long as we are in the rotor-condensed phase, we can assume $\langle \cos \theta_j \rangle \equiv \langle \cos \theta \rangle$ which allows us to evaluate K ,

$$\begin{aligned} K &= 2 \langle \cos \theta \rangle \sum_{\sigma} \frac{1}{N_\Lambda} (-t) \sum_k (g/t) \langle f_{k\sigma}^\dagger f_{k\sigma}^b \rangle \\ &= 4 \langle \cos \theta \rangle \frac{1}{N_\Lambda} \sum_k -g \frac{g^*}{2|g|} \langle f_{k\sigma}^\dagger f_{k\sigma}^b \rangle \\ &= 2 \frac{1}{N_\Lambda} \sum_k -|g| \langle \cos \theta \rangle \equiv 2\bar{\epsilon}(1/2) \langle \cos \theta \rangle. \end{aligned} \quad (\text{B12})$$

The matrix elements in the previous equation are easily calculated using the definitions of Sec. IV,

$$\langle f_{k\sigma}^\dagger f_{k\sigma}^b \rangle = \begin{cases} -\alpha_-^* \beta_- \langle l_{k\uparrow}^\dagger l_{k\uparrow} \rangle & \gamma \rightarrow 0 \\ \alpha_+^* \beta_+ \langle l_{k\downarrow}^\dagger l_{k\downarrow} \rangle & \gamma \rightarrow 0 \end{cases} = \begin{cases} \frac{1}{2} \frac{g^*}{|g|} \\ \frac{1}{2} \frac{g^*}{|g|} \end{cases}. \quad (\text{B13})$$

We define the half bandwidth $D=3t$ for the honeycomb lattice and find numerically the result

$$D \sum_{\sigma} \langle f_{i\sigma}^\dagger f_{j\sigma}^b \rangle = |\bar{\epsilon}(1/2)| \approx 1.57. \quad (\text{B14})$$

Now let us consider Eq. (B9),

$$\begin{aligned} \langle L \rangle &= - \sum_{\sigma} \left(\langle f_{i\sigma}^\dagger f_{i\sigma} \rangle - \frac{1}{2} \right) \\ &= - \sum_{\sigma} \left[\frac{1}{N_\Lambda} \sum_k \left(\frac{1}{2} \langle f_{k\sigma}^\dagger f_{k\sigma}^b \rangle + \frac{1}{2} \langle f_{k\sigma}^\dagger f_{k\sigma}^a \rangle \right) - \frac{1}{2} \right] \\ &= -2 \left(\frac{1}{2} - \frac{1}{2} \right) = 0, \end{aligned} \quad (\text{B15})$$

where we assumed that the lower band is completely filled while the upper band is empty. Also, since we still assume

the half-filled case, Eq. (40) of Ref. 76 is easily

$$n = \frac{1}{2} \sum_{\sigma} \langle f_{i\sigma}^{\dagger} f_{i\sigma} \rangle = \frac{1}{2} \sum_{\sigma} \frac{1}{2} = \frac{1}{2}. \quad (\text{B16})$$

This is equivalent to set $\mu+h=0$. Thus, we can also introduce $\mu_0(n)$ which is defined as

$$\mu_0(n) = \frac{h + \mu^{\text{half filling}}}{Z} \rightarrow 0. \quad (\text{B17})$$

Similar to Ref. 76 we obtain the Green's function

$$G_{fl}(\mathbf{k}, i\omega_n)^{-1} = i\omega_n - Z\varepsilon_{\mathbf{k}}. \quad (\text{B18})$$

As a last step $\langle \cos \theta \rangle$ has to be calculated. Following Ref. 76 we calculate it in first-order perturbation theory in K . We start with $\langle \cos \theta \rangle = \langle \psi_n^{(1)} | \cos \theta | \psi_n^{(1)} \rangle$, where

$$|\psi_n^{(1)}\rangle = |l_n\rangle + \sum_{l \neq l_n} \frac{\langle l | \cos \theta | l_n \rangle}{E_l - E_l} |l\rangle. \quad (\text{B19})$$

To first order in K only the ‘‘mixed’’ element contributes

$$\langle \cos \theta \rangle = 2K \sum_{l \neq l_n} \frac{|\langle l | \cos \theta | l_n \rangle|^2}{E_l - E_l}. \quad (\text{B20})$$

The energies are given by $E_l = 1/(2U)(UL+h)^2 + \text{const}$ and the matrix elements by $\frac{1}{4} |\langle l | e^{i\theta} + e^{-i\theta} | l_n \rangle|^2 = \frac{1}{4} (\delta_{l, l_{n-1}} + \delta_{l, l_{n+1}})$. Altogether we find the result

$$\langle \cos \theta \rangle = -\frac{2K}{U}, \quad (\text{B21})$$

which is in agreement with Ref. 76 at half filling. Then we substitute $\langle \cos \theta \rangle$ in Eq. (B12) and obtain finally

$$U_c = -4\bar{\varepsilon}(1/2) = 4|\bar{\varepsilon}(1/2)|. \quad (\text{B22})$$

By means of Eq. (B14) we find the phase transition at

$$U_c^{\infty} \simeq 6.30t, \quad (\text{B23})$$

which should be considered as the correct result in $d=\infty$ dimensions. In this appendix, we used a simple mean-field approximation with the severe restriction $Z = \langle \cos(\theta_{ij}) \rangle$. This approximation might be justified in large dimensions. Therefore, we can assume that the result U_c^{∞} is exact in $d=\infty$.

APPENDIX C: DERIVATION OF GREEN'S FUNCTIONS

In this appendix, we pedagogically show all the relevant steps starting from the slave-rotor Hamiltonian to the Green's functions. We have omitted this part in Sec. V for the sake of clarity. The slave-rotor Hamiltonian reads

$$\begin{aligned} \mathcal{H} = & -t \sum_{\langle ij \rangle} \sum_{\sigma} (f_{i\sigma}^{a\dagger} f_{j\sigma}^b e^{-i\theta_{ij}} + \text{H.c.}) \\ & + i\lambda \sum_{\langle\langle ij \rangle\rangle} \sum_{\sigma\sigma'} v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^{\dagger} f_{j\sigma'} e^{-i\theta_{ij}} - \mu \sum_{i,\sigma} f_{i\sigma}^{\dagger} f_{i\sigma} + \frac{U}{2} \sum_i L_i^2, \end{aligned} \quad (\text{C1})$$

where we still have to fulfill the constraint Eq. (45) with the

Lagrange multiplier h . Then, the action is built from

$$S_0 \equiv \int_0^{\beta} d\tau [-iL\partial_{\tau}\theta + \mathcal{H} + f^{\dagger} \partial_x f], \quad (\text{C2})$$

where the first two terms correspond to the Legendre transform between \mathcal{H} and \mathcal{L} and we are switching from phase and angular momentum operator (θ, L) to fields $(\theta, \partial_{\tau}\theta)$. Here L and $\partial_{\tau}\theta$ are related as follows:

$$i\partial_{\tau}\theta = \frac{\partial \mathcal{H}}{\partial L}, \quad (\text{C3})$$

which yields $L = (i/U)\partial_{\tau}\theta$. We obtain the action

$$\begin{aligned} S_0 = & \int_0^{\beta} d\tau \left[\sum_{i\sigma} f_{i\sigma}^{\dagger} (\partial_{\tau} - \mu + h_i) f_{i\sigma} + \frac{1}{2U} \sum_i (\partial_{\tau}\theta_i + ih_i)^2 \right. \\ & + \sum_i \left(-h_i + \frac{h_i^2}{2U} \right) - t \sum_{\langle ij \rangle} \left(\sum_{\sigma} f_{i\sigma}^{a\dagger} f_{j\sigma}^b e^{-i\theta_{ij}} + \text{c.c.} \right) \\ & \left. + \lambda \sum_{\langle\langle ij \rangle\rangle} \left(\sum_{\sigma\sigma'} i v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^{\dagger} f_{j\sigma'} e^{-i\theta_{ij}} \right) \right]. \end{aligned} \quad (\text{C4})$$

Now we have the choice to decompose the hopping and spin-orbit terms either in a standard way (as Florens and Georges did⁷⁶) or in a more elaborate way (as Lee and Lee did⁸⁰) to obtain an effective theory. Since we are mainly interested in the transition line to the Mott phase we restrict ourselves to the first way for the moment. Thus we will use again the decomposition $\alpha\beta \approx \langle \alpha \rangle \beta + \alpha \langle \beta \rangle - \langle \alpha \rangle \langle \beta \rangle$ with

$$\alpha_{ij} = \sum_{\sigma} f_{i\sigma}^{a\dagger} f_{j\sigma}^b, \quad \langle \alpha_{ij} \rangle \equiv Q_X,$$

$$\beta_{ij} = \exp(-i\theta_{ij}), \quad \langle \beta_{ij} \rangle \equiv Q_f,$$

for the hopping term and

$$\alpha'_{ij} = \sum_{\sigma\sigma'} i v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^{\dagger} f_{j\sigma'}, \quad \langle \alpha'_{ij} \rangle \equiv Q'_X,$$

$$\beta'_{ij} = \exp(-i\theta_{ij}), \quad \langle \beta'_{ij} \rangle \equiv Q'_f,$$

for the spin-orbit term. Then we replace the exponentials $\exp(i\theta_i)$ by complex bosonic fields $X(\tau)$ which are constraint via $|X_i|^2 = 1$. This constraint is imposed by a complex Lagrange multiplier ρ_i . Then the decomposed Lagrangian has the form

$$S = S' + S'' + S''', \quad (\text{C5})$$

where S' contains the hopping term, S'' the spin-orbit term, and S''' the other terms. S' is given by

$$\begin{aligned}
S' &\approx \int_0^\beta d\tau \left[-|t|Q_X \sum_{\langle ij \rangle} X_i^* X_j + \text{c.c.} \right. \\
&\quad \left. - |t|Q_f \sum_{\langle ij \rangle} \sum_{\sigma} f_{i\sigma}^* f_{j\sigma} + \text{c.c.} + |t| \sum_{ij} Q_f Q_X \right] \\
&= \int_0^\beta d\tau [\mathcal{L}_X + \mathcal{L}_f + \dots]. \tag{C6}
\end{aligned}$$

The second term S'' is given by

$$\begin{aligned}
S'' &\approx \int_0^\beta d\tau \left[\lambda Q'_X \sum_{\langle\langle ij \rangle\rangle} X_i^* X_j + \text{c.c.} \right. \\
&\quad \left. + \lambda Q'_f \sum_{\langle\langle ij \rangle\rangle} \sum_{\sigma\sigma'} i v_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^* f_{j\sigma'} + \lambda \sum_{\langle\langle ij \rangle\rangle} Q'_f Q'_X \right] \\
&= \int_0^\beta d\tau [\mathcal{L}'_X + \mathcal{L}'_f + \dots]. \tag{C7}
\end{aligned}$$

Introducing the X -field changes the Hubbard term such as

$$(\partial_\tau \theta_i + h_i)^2 = [(i\partial_\tau + h_i)X_i^*][(-i\partial_\tau + h_i)X_i] \tag{C8}$$

and the term S''' becomes the form

$$\begin{aligned}
S''' &= \frac{1}{2U} \sum_i [(i\partial_\tau + h_i)X_i^*][(-i\partial_\tau + h_i)X_i] + \sum_i \rho_i |X_i|^2 \\
&\quad + \sum_{i\sigma} f_{i\sigma}^* (\partial_\tau - \mu + h_i) f_{i\sigma} + \dots \\
&= \int_0^\beta d\tau [\mathcal{L}''_X + \mathcal{L}''_f + \dots]. \tag{C9}
\end{aligned}$$

Here and in the previous equations the ellipsis corresponds to the other terms which are independent of f_σ and X . The Fourier transform of \mathcal{L}_X and \mathcal{L}_f leads in the (u, l) basis to the bands obtained earlier,

$$\begin{aligned}
\mathcal{L}_X &= Q_X \sum_{\mathbf{k}} [-g(\mathbf{k})] X_{\mathbf{k}}^{a*} X_{\mathbf{k}}^b + \text{c.c.} \\
&= Q_X \sum_{\mathbf{k}} (-|g|) X_{\mathbf{k}}^{l*} X_{\mathbf{k}}^l + |g| X_{\mathbf{k}}^{u*} X_{\mathbf{k}}^u, \tag{C10}
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_f &= Q_f \sum_{\mathbf{k}\sigma} [-g(\mathbf{k})] f_{\mathbf{k}\sigma}^{a*} f_{\mathbf{k}\sigma}^b + \text{c.c.} \\
&= Q_f \sum_{\mathbf{k}\sigma} (-|g|) f_{\mathbf{k}\sigma}^{l*} f_{\mathbf{k}\sigma}^l + |g| f_{\mathbf{k}\sigma}^{u*} f_{\mathbf{k}\sigma}^u. \tag{C11}
\end{aligned}$$

For the X part of the spin-orbit term \mathcal{L}'_X we find the following expression:

$$\begin{aligned}
\mathcal{L}'_X &= Q'_X \sum_{\mathbf{k}} \lambda g_2(\mathbf{k}) (X_{\mathbf{k}}^{a*} X_{\mathbf{k}}^a + X_{\mathbf{k}}^{b*} X_{\mathbf{k}}^b) \\
&= Q'_X \sum_{\mathbf{k}} \lambda g_2(\mathbf{k}) (X_{\mathbf{k}}^{l*} X_{\mathbf{k}}^l + X_{\mathbf{k}}^{u*} X_{\mathbf{k}}^u), \tag{C12}
\end{aligned}$$

where $g_2(\mathbf{k})$ is the usual next-nearest-neighbor hopping contribution as defined in Eq. (19). The last term which must be

transformed into momentum space is \mathcal{L}'_f which clearly produces the γ term. Therefore we will add \mathcal{L}_f to \mathcal{L}'_f in the (f_σ^a, f_σ^b) basis and then transform both terms to the (f_σ^l, f_σ^u) basis as we did with the original bands of the KM model,

$$\begin{aligned}
\mathcal{L}_f + \mathcal{L}'_f &= \sum_{\mathbf{k}\sigma} Q_f (-g(\mathbf{k}) f_{\mathbf{k}\sigma}^{a*} f_{\mathbf{k}\sigma}^b - g(\mathbf{k})^* f_{\mathbf{k}\sigma}^{b*} f_{\mathbf{k}\sigma}^a) \\
&\quad + \sum_{\mathbf{k}\sigma\sigma'} \sigma_{\sigma\sigma'}^z Q'_f \gamma (f_{\mathbf{k}\sigma}^{a*} f_{\mathbf{k}\sigma'}^a - f_{\mathbf{k}\sigma}^{b*} f_{\mathbf{k}\sigma'}^b) \\
&= \sum_{\mathbf{k}\sigma} -\sum_{\mathbf{k}l} f_{\mathbf{k}\sigma}^{l*} f_{\mathbf{k}\sigma}^l + \sum_{\mathbf{k}l} f_{\mathbf{k}\sigma}^{u*} f_{\mathbf{k}\sigma}^u. \tag{C13}
\end{aligned}$$

Here we have introduced the renormalized KM spectrum for the spinon sector,

$$\Sigma_{\mathbf{k}} = \sqrt{(Q_f |g|)^2 + (Q'_f \gamma)^2}. \tag{C14}$$

Finally we find the imaginary time Green's function for the f_σ^l fields,

$$G_{fl} = \frac{1}{i\omega_n - \Sigma_{\mathbf{k}}} \tag{C15}$$

and for the X fields,

$$G_X = \frac{1}{v_n^2/U + \rho + \xi_{\mathbf{k}}}, \tag{C16}$$

where we defined

$$\xi_{\mathbf{k}} = -Q_X |g(\mathbf{k})| + Q'_X \lambda g_2(\mathbf{k}). \tag{C17}$$

APPENDIX D: MATSUBARA SUM

In the self-consistency Eqs. (65), (72), and (74) we had to evaluate the following Matsubara sum:

$$\begin{aligned}
\frac{1}{\beta} \sum_n G_X(\mathbf{k}, i\nu_n) &= \frac{U}{\beta} \sum_n \frac{1}{v_n^2 + U(\rho + \xi_{\mathbf{k}})} \\
&= \frac{U}{\beta} \sum_n \frac{1}{(i\nu_n + A)(-i\nu_n + A)}, \tag{D1}
\end{aligned}$$

where $A = \sqrt{U(\rho + \xi_{\mathbf{k}})}$. By taking the corresponding contour, we find

$$\begin{aligned}
0 &= \oint_C \frac{n_B(z)}{(z+A)(z-A)} dz \\
&= 2\pi i \frac{1}{\beta} \sum_{n=0, \pm 1, \pm 2, \dots} \frac{1}{(i\nu_n + A)(-i\nu_n + A)} \\
&\quad - 2\pi i \left[\frac{n_B(A)}{2A} - \frac{n_B(-A)}{2A} \right], \tag{D2}
\end{aligned}$$

where $n_B(z) = [\exp(\beta z) - 1]^{-1}$ is the Bose function. The last equation then implies

$$\frac{1}{\beta} \sum_{n \in \mathbb{Z}} \frac{1}{(i\nu_n + A)(-i\nu_n + A)} = \frac{n_B(A) - n_B(-A)}{2A}$$

$$= \frac{\coth(\beta A/2)}{2A} \xrightarrow{T \rightarrow 0} \frac{1}{2A}.$$

At zero temperature we find the result

$$\frac{1}{\beta} \sum_n G_X(\mathbf{k}, i\nu_n) = \frac{U}{2\sqrt{U(\rho + \xi_k)}}. \quad (\text{D3})$$

APPENDIX E: DYNAMICAL GAUGE FIELD

In this appendix, we decompose the action Eq. (C4) in a different way compared to Appendix C and show, how the effective-field theory and the gauge field a_{ij} emerge. Following Lee and Lee⁸⁰ we decompose the hopping term via Hubbard-Stratonovich decomposition, $\alpha_{ij} = \sum_{\sigma} f_{i\sigma}^{\alpha} f_{j\sigma}^{\beta}$ and $\beta_{ij} = e^{i\theta_{ij}}$,

$$\int d\eta_{ij} d\eta_{ij}^* d\eta_{ji} d\eta_{ji}^* e^{-\ell[|\eta_{ij}|^2 + |\eta_{ji}|^2 - \eta_{ij}^* \beta_{ij} - \eta_{ij} \alpha_{ij} - \eta_{ji}^* \beta_{ji} - \eta_{ji} \alpha_{ji}]} = \frac{\pi^2}{\ell^2} e^{\ell[\alpha_{ij} \beta_{ij} + \alpha_{ji} \beta_{ji}]}. \quad (\text{E1})$$

The same equation holds for α' and β' in order to decouple the spin-orbit term,

$$\alpha'_{ij} = \sum_{\sigma\sigma'} i\nu_{ij} \sigma_{\sigma\sigma'}^z f_{i\sigma}^* f_{j\sigma'},$$

$$\beta'_{ij} = \exp(-i\theta_{ij}).$$

In Eq. (E1), ℓ is given by $\Delta\tau$ times the hopping or the spin-orbit amplitude, respectively. We further follow Lee and Lee and change the variables of integration by $\eta_{ij} = |\chi_{ij}| e^{-w_{ij} + i(a_{ij}^+ - a_{ij})}$ and $\eta_{ji} = |\chi_{ij}| e^{-w_{ij} + i(a_{ij}^+ - a_{ij})}$. Note that η_{ij} and η_{ji} are independent complex variables, and hence w_{ij} and a_{ij}^+ are independent and necessary. At this point, we replace again $\exp(i\theta_{ij})$ by the bosonic X_i field with the constraint $|X_i|^2 = 1$ imposed by the Lagrange multiplier ρ_i . Then we find the action which coincides with Eq. (4) of Ref. 80 apart from two terms coming from the spin-orbit interaction and the slightly different notation of the rotor variables. We replace the variables by their saddle-point values plus fluctuations (see for details Ref. 80), neglect the massive modes and we can integrate out the ρ_i field to restore the θ field. Finally we obtain the effective Lagrangian (similar to Ref. 80),

$$L' = \sum_{i\sigma} f_{i\sigma}^* (\partial_\tau - ia_i^\tau + i\tilde{h}_i - \mu) f_{i\sigma} + \frac{1}{2U} \sum_i (\partial_\tau \theta_i - a_i^\tau - \tilde{h}_i)^2 - \sum_{\langle ij \rangle, \sigma} |t| \tilde{\chi}_{ij}^X e^{ia_j^j} f_{j\sigma}^* f_{i\sigma} - \sum_{\langle ij \rangle} |t| \tilde{\chi}_{ij}^f e^{-(\theta_i - \theta_j - a_j^j)} - \sum_{\langle\langle ij \rangle\rangle} \sum_{\sigma\sigma'} |\lambda| \tilde{\chi}_{ij}^{X'} e^{ia_j^j} i\nu_{ij} f_{j\sigma}^* \sigma_{\sigma\sigma'}^z - \sum_{\langle\langle ij \rangle\rangle} |\lambda| \tilde{\chi}_{ij}^{f'} e^{-i(\theta_i - \theta_j - a_j^j)}. \quad (\text{E2})$$

Here a_i^τ and a_i^j are the temporal and spatial gauge fields coming from the fluctuations from h_i and a_{ij} , respectively. h_i is the Lagrange multiplier associated with the global constraint introduced earlier. The quantities with tildes are the saddle-point values and are identical to the mean-field parameters which we have evaluated in Sec. V. We notice that both spinons and rotors couple to the U(1) gauge field. Since we assume weak gauge fluctuations, we take the saddle-point approximation, i.e., $a_i^j = 0$; we recover for the spinons the same terms which resulted in Sec. V in

the renormalized KM spectrum ($\tilde{\chi}_{ij}^X \rightarrow Q_X$ and $\tilde{\chi}_{ij}^{X'} \rightarrow Q_X'$). We should also mention that the spinons are still coupled to the gauge field through the first term in Eq. (E2) which contains $\sim f_{i\sigma}^* a_i^\tau f_{i\sigma}$. In principle, although we have set $a_i^j = 0$, we could allow for small deviations and expand $\exp(ia_j^j) \approx 1 + ia_j^j$; thus the spinons couple to both temporal and spatial gauge fields. On the other hand, we know that the rotors are gapped in the Mott phase and can be integrated out. This generates the Maxwellian term (see, e.g., Refs. 72, 73, and 80).

*stephan.rachel@yale.edu

†karyn.lehur@yale.edu

¹X.-L. Qi and S.-C. Zhang, *Phys. Today* **63**(1), 33 (2010).

²M. Hasan and C. Kane, arXiv:1002.3895 (unpublished).

³J. Moore, *Nat. Phys.* **4**, 270 (2008); **5**, 378 (2009); *Nature (London)* **464**, 194 (2010).

⁴M. Büttiker, *Science* **325**, 278 (2009).

⁵C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).

⁶C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).

⁷L. Fu, C. L. Kane, and E. J. Mele, *Phys. Rev. Lett.* **98**, 106803 (2007).

⁸J. E. Moore and L. Balents, *Phys. Rev. B* **75**, 121306(R) (2007).

⁹R. Roy, *Phys. Rev. B* **79**, 195322 (2009).

¹⁰B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, *Science* **314**,

- 1757 (2006).
- ¹¹L. Fu and C. L. Kane, *Phys. Rev. B* **76**, 045302 (2007).
 - ¹²H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, *Nat. Phys.* **5**, 438 (2009).
 - ¹³M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, *Science* **318**, 766 (2007).
 - ¹⁴D. Hsieh, D. Qian, L. Wray, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Nature (London)* **452**, 970 (2008).
 - ¹⁵D. Hsieh, Y. Xia, L. Wray, D. Qian, A. Pal, J. H. Dil, J. Osterwalder, F. Meier, G. Bihlmayer, C. L. Kane, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Science* **323**, 919 (2009).
 - ¹⁶Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Nat. Phys.* **5**, 398 (2009).
 - ¹⁷Y. L. Chen, J. G. Analytis, J.-H. Chu, Z. K. Liu, S.-K. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fang, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z.-X. Shen, *Science* **325**, 178 (2009).
 - ¹⁸P. Roushan, J. Seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, and A. Yazdani, *Nature (London)* **460**, 1106 (2009).
 - ¹⁹D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, J. Osterwalder, L. Patthey, J. G. Checkelsky, N. P. Ong, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Nature (London)* **460**, 1101 (2009).
 - ²⁰C. Xu and J. E. Moore, *Phys. Rev. B* **73**, 045322 (2006).
 - ²¹D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, *Phys. Rev. Lett.* **97**, 036808 (2006).
 - ²²X.-L. Qi, Y.-S. Wu, and S.-C. Zhang, *Phys. Rev. B* **74**, 045125 (2006).
 - ²³L. Fu and C. L. Kane, *Phys. Rev. B* **74**, 195312 (2006).
 - ²⁴C. Wu, B. A. Bernevig, and S.-C. Zhang, *Phys. Rev. Lett.* **96**, 106401 (2006).
 - ²⁵J. C. Y. Teo, L. Fu, and C. L. Kane, *Phys. Rev. B* **78**, 045426 (2008).
 - ²⁶S.-S. Lee and S. Ryu, *Phys. Rev. Lett.* **100**, 186807 (2008).
 - ²⁷M. Levin and A. Stern, *Phys. Rev. Lett.* **103**, 196803 (2009).
 - ²⁸E. Prodan, *Phys. Rev. B* **80**, 125327 (2009).
 - ²⁹Z. Wang, N. Hao, and P. Zhang, *Phys. Rev. B* **80**, 115420 (2009).
 - ³⁰D.-H. Lee, *Phys. Rev. Lett.* **103**, 196804 (2009).
 - ³¹C. Xu, *Phys. Rev. B* **81**, 020411(R) (2010).
 - ³²X.-L. Qi, T. L. Hughes, and S.-C. Zhang, *Phys. Rev. B* **78**, 195424 (2008).
 - ³³X. L. Qi, R. D. Li, J. D. Zang, and S. C. Zhang, *Science* **323**, 1184 (2009).
 - ³⁴F. D. M. Haldane, *Phys. Rev. Lett.* **61**, 2015 (1988).
 - ³⁵A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).
 - ³⁶P. Ghaemi and F. Wilczek, [arXiv:0709.2626](https://arxiv.org/abs/0709.2626) (unpublished).
 - ³⁷D. L. Bergman and K. Le Hur, *Phys. Rev. B* **79**, 184520 (2009).
 - ³⁸J. Koch and K. Le Hur, *Phys. Rev. A* **80**, 023811 (2009).
 - ³⁹G. Jackeli and G. Khaliullin, *Phys. Rev. Lett.* **102**, 017205 (2009).
 - ⁴⁰A. Shitade, H. Katsura, J. Kunes, X.-L. Qi, S.-C. Zhang, and N. Nagaosa, *Phys. Rev. Lett.* **102**, 256403 (2009).
 - ⁴¹D. A. Pesin and L. Balents, *Nat. Phys.* **6**, 376 (2010).
 - ⁴²B. Yang and Y. Kim, [arXiv:1004.4630](https://arxiv.org/abs/1004.4630) (unpublished).
 - ⁴³S. Raghu, X.-L. Qi, C. Honerkamp, and S.-C. Zhang, *Phys. Rev. Lett.* **100**, 156401 (2008).
 - ⁴⁴C. Weeks and M. Franz, *Phys. Rev. B* **81**, 085105 (2010).
 - ⁴⁵J. Wen, A. Rüegg, C.-C. J. Wang, and G. A. Fiete, [arXiv:1005.4061](https://arxiv.org/abs/1005.4061) (unpublished).
 - ⁴⁶H.-M. Guo and M. Franz, *Phys. Rev. B* **80**, 113102 (2009).
 - ⁴⁷A. Rüegg, J. Wen, and G. A. Fiete, *Phys. Rev. B* **81**, 205115 (2010).
 - ⁴⁸J. Li, R.-L. Chu, J. K. Jain, and S.-Q. Shen, *Phys. Rev. Lett.* **102**, 136806 (2009).
 - ⁴⁹C. W. Groth, M. Wimmer, A. R. Akhmerov, J. Tworzydło, and C. W. J. Beenakker, *Phys. Rev. Lett.* **103**, 196805 (2009).
 - ⁵⁰P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, *Phys. Rev. Lett.* **105**, 036803 (2010).
 - ⁵¹R. Biswas and A. Balatsky, *Phys. Rev. B*, **81**, 233405 (2010).
 - ⁵²F. Wilczek, *Phys. Rev. Lett.* **58**, 1799 (1987).
 - ⁵³Q.-F. Sun and X. C. Xie, *Phys. Rev. Lett.* **104**, 066805 (2010).
 - ⁵⁴H. Lin, L. A. Wray, Y. Xia, S. Jia, R. J. Cava, A. Bansil, and M. Z. Hasan, *Nat. Mater.* **9**, 546 (2010).
 - ⁵⁵S. Chadov, X. Qi, J. Kübler, G. H. Fecher, C. Felser, and S.-C. Zhang, *Nat. Mater.* **9**, 541 (2010).
 - ⁵⁶I. Bloch, J. Dalibard, and W. Zwerger, *Rev. Mod. Phys.* **80**, 885 (2008).
 - ⁵⁷Y.-J. Lin, R. L. Compton, K. Jiménez-García, J. V. Porto, and I. B. Spielman, *Nature (London)* **462**, 628 (2009).
 - ⁵⁸N. Goldman, I. Satija, P. Nikolic, A. Bermudez, M. Martin-Delgado, M. Lewenstein, and I. Spielman, [arXiv:1002.0219](https://arxiv.org/abs/1002.0219) (unpublished).
 - ⁵⁹C. Wu, *Phys. Rev. Lett.* **101**, 186807 (2008).
 - ⁶⁰T. D. Stanescu, V. Galitski, J. Y. Vaishnav, C. W. Clark, and S. Das Sarma, *Phys. Rev. A* **79**, 053639 (2009).
 - ⁶¹T. D. Stanescu, V. Galitski, and S. Das Sarma, *Phys. Rev. A* **82**, 013608 (2010).
 - ⁶²J. Goryo, N. Maeda, and K.-I. Imura, [arXiv:0905.2296](https://arxiv.org/abs/0905.2296) (unpublished).
 - ⁶³J. Goryo and N. Maeda, [arXiv:1007.1507](https://arxiv.org/abs/1007.1507), *J. Phys. Soc. Jpn.* (to be published).
 - ⁶⁴Z. Wang, X.-L. Qi, and S.-C. Zhang, [arXiv:1004.4229](https://arxiv.org/abs/1004.4229) (unpublished).
 - ⁶⁵A. M. Polyakov, *Phys. Lett.* **59B**, 82 (1975).
 - ⁶⁶M. W. Young, S.-S. Lee, and C. Kallin, *Phys. Rev. B* **78**, 125316 (2008).
 - ⁶⁷X.-G. Wen, *Phys. Rev. B* **65**, 165113 (2002).
 - ⁶⁸M. Fisher, O. Motrunich, and D. Sheng, [arXiv:0812.2955](https://arxiv.org/abs/0812.2955) (unpublished).
 - ⁶⁹K. Le Hur, *Phys. Rev. B* **63**, 165110 (2001).
 - ⁷⁰Y. Ran, A. Vishwanath, and D.-H. Lee, *Phys. Rev. Lett.* **101**, 086801 (2008).
 - ⁷¹X.-L. Qi and S.-C. Zhang, *Phys. Rev. Lett.* **101**, 086802 (2008).
 - ⁷²Y. Ran, A. Vishwanath, and D. Lee, [arXiv:0806.2321](https://arxiv.org/abs/0806.2321) (unpublished).
 - ⁷³M. Hermele, Y. Ran, P. A. Lee, and X.-G. Wen, *Phys. Rev. B* **77**, 224413 (2008).
 - ⁷⁴R. Shankar, *Rev. Mod. Phys.* **66**, 129 (1994).
 - ⁷⁵S. Florens and A. Georges, *Phys. Rev. B* **66**, 165111 (2002).
 - ⁷⁶S. Florens and A. Georges, *Phys. Rev. B* **70**, 035114 (2004).
 - ⁷⁷E. Zhao and A. Paramekanti, *Phys. Rev. B* **76**, 195101 (2007).
 - ⁷⁸S. Sorella and E. Tosatti, *Europhys. Lett.* **19**, 699 (1992).
 - ⁷⁹S. A. Jafari, *Eur. Phys. J. B* **68**, 537 (2009).
 - ⁸⁰S.-S. Lee and P. A. Lee, *Phys. Rev. Lett.* **95**, 036403 (2005).
 - ⁸¹I. F. Herbut, *Phys. Rev. Lett.* **97**, 146401 (2006).
 - ⁸²M. Hermele, *Phys. Rev. B* **76**, 035125 (2007).

- ⁸³H. Feldner, Z. Meng, A. Honecker, D. Cabra, S. Wessel, and F. Assaad, *Phys. Rev. B* **81**, 115416 (2010).
- ⁸⁴Z. Y. Meng, T. C. Lang, S. Wessel, F. Assaad, and A. Muramatsu, *Nature (London)* **464**, 847 (2010).
- ⁸⁵G. Sun and S. Kou, [arXiv:0911.3002](https://arxiv.org/abs/0911.3002) (unpublished).
- ⁸⁶A. Auerbach, *Interacting Electrons and Quantum Magnetism* (Springer, New York, 1994).
- ⁸⁷J. Fouet, P. Sindzingre, and C. Lhuillier, *Eur. Phys. J. B* **20**, 241 (2001).
- ⁸⁸K. Le Hur and T. M. Rice, *Ann. Phys.* **324**, 1452 (2009).