SU(2) slave fermion solution of the Kitaev honeycomb lattice model

F. J. Burnell[1,2] and Chetan Nayak[3,4]

[1]*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford OX1 3NP, United Kingdom*
[2]*All Souls College, Oxford, United Kingdom*
[3]*Microsoft Research, Station Q, Elings Hall, University of California, Santa Barbara, California 93106, USA*
[4]*Department of Physics, University of California, Santa Barbara, California 93106, USA*

(Received 16 May 2011; revised manuscript received 21 July 2011; published 13 September 2011)

We apply the SU(2) slave fermion formalism to the Kitaev honeycomb lattice model. We show that both the toric code phase (the A phase) and the gapless phase of this model (the B phase) can be identified with p-wave superconducting phases of the slave fermions, with nodal lines which, respectively, do not or do intersect the Fermi surface. The non-Abelian Ising anyon phase is a $p + ip$ superconducting phase that occurs when the B phase is subjected to a gap-opening magnetic field. We also discuss the transitions between these phases in this language.

DOI: 10.1103/PhysRevB.84.125125   PACS number(s): 75.10.Kt

I. INTRODUCTION

In Ref. 1, Kitaev introduced the following remarkable model of $s = 1/2$ spins on a honeycomb lattice:

$$H = -J_x \sum_{x\text{ links}} S_j^x S_j^x - J_y \sum_{y\text{ links}} S_j^y S_j^y - J_z \sum_{z\text{ links}} S_j^x S_j^z, \quad (1)$$

where the z links are the vertical links on the honeycomb lattice, and the x and y links are at angles $\pm\pi/3$ from the vertical. This model is exactly solvable and has a gapped Abelian topological phase (the A phase), which is equivalent to the toric code.[2] It also has a gapless phase (the B phase), which, when subjected to an appropriate time-reversal symmetry-breaking perturbation, becomes a gapped non-Abelian topological phase supporting Ising anyons.

This model is one of the rare instances of an exactly solvable model of a quantum magnet that does not order in its ground state and, instead, condenses into a topological phase. As such, it is a useful testing ground for theoretical techniques, such as slave fermion representations, which have been applied to approximately solve models of frustrated magnets that are not exactly solvable. Applying these techniques to Eq. (1) can shed light on the physics of this model and, conversely, on the applicability of these techniques.

Kitaev solved the Hamiltonian (1) by introducing a fermionization of the spins in terms of Majorana fermions. By expressing each spin operator as a product of two Majorana fermions, the spin model can be described exactly as a model of Majorana fermions coupled to a Z_2 gauge field. In this description, the effect of the gauge field is particularly transparent: the physical correlators are captured exactly by the fermionic band structure, and the gauge field serves only to enforce the fact that only gauge-invariant observables (e.g., products of spins) are physical.

In this paper, we apply a different fermionization procedure, the SU(2) slave fermion formalism. This representation requires a different projection to eliminate redundancies in the Hilbert space compared to Kitaev's representation in terms of Majorana fermions; therefore, it is interesting to see how the same low-energy degrees of freedom emerge. In the SU(2) slave fermion formalism, the spins are written in terms of standard, rather than Majorana, fermionic spinons. The Hamiltonian of Eq. (1) is then expanded about a resonating valence bond (RVB) mean-field state. We show that this is a stable mean-field theory which captures the physical correlation functions of the exact ground state of Eq. (1). We find that the A phase is a p-wave superconducting state of the slave fermions. The state is fully gapped because the nodes in the order parameter do not intersect the Fermi surface. The Majorana fermions of Kitaev's solution appear as Bogoliubov–de Gennes quasiparticles of the superconducting state. The B phase is a p-wave superconducting state with gapless excitations at the nodal points. These excitations form a single Dirac fermion. When the order parameter develops an ip component, the Dirac fermion acquires a mass, and the system goes into an Ising anyon phase. The transition point between the A phase and the gapless B phase is an interesting quantum critical point, which we describe in terms of superconducting order parameters.

By studying the theory of fluctuations about the mean-field saddle point, we recover the Z_2 gauge field as the unbroken gauge symmetry remaining in the superconducting state. This situates the ground state of the finely tuned Hamiltonian (1) in the broader context of spin liquid[3–12] and superconducting phases, and allows us to understand its phase diagram in terms of these more familiar phases of matter.

II. SU(2) SLAVE FERMION FORMULATION

A. Slave fermion mean-field Hamiltonian

Our starting point is the representation of the spin operators in terms of spinful Dirac fermions, first discussed in Ref. 5 and described in detail in Ref. 7. We thus write the spin operator on site i $\hat{S}_i^\alpha$, $\alpha = x,y,z$, as

$$\hat{S}_i^\alpha = \tfrac{1}{2} f_{i\alpha}^\dagger \sigma_{\alpha\beta}^\alpha f_{i\beta}. \quad (2)$$

Here, we have introduced the fermion operators $f_{i\alpha}$, usually called spinons. For two-spin interactions of the form $\hat{S}_i^\alpha \hat{S}_j^\beta$, one way to treat the resulting Hamiltonian is to use a Hubbard-Stratonovich transformation to decouple the four-fermion interactions, reexpressing them as interactions between a bosonic field Φ (which lives on a link in the lattice) and a pair of fermion operators on the sites i and j bordering this link. One may then study the mean-field solutions, which can be obtained by condensing the bosons. This is often a fruitful way to investigate candidate "spin liquid" ground states, in which the spins are strongly correlated but have no spatial order.

One important caveat in this formulation is that Eq. (2) gives a faithful representation of the Hilbert space only in the subspace of fermionic states for which each site is singly occupied. Thus, at each site (i), we must impose the three (redundant) constraints

$$n_{i\uparrow} + n_{i\downarrow} = 1, \qquad f^\dagger_{i\uparrow} f^\dagger_{i\downarrow} = 0, \quad f_{i\uparrow} f_{i\downarrow} = 0. \tag{3}$$

As explained in Refs. 5,13, and 14 when the Hamiltonian preserves SU(2) spin rotation symmetry, the Lagrange multipliers of these constraints can be viewed as the temporal component of an SU(2) gauge field, leading to a theory of fermions coupled to a fluctuating gauge field. [The spatial components of this gauge field are given by the phases of the fermion kinetic terms, which here arise due to condensation of a bosonic field (see Appendix B 3).] Projection would be enforced by integrating out the gauge fields. In practice, this is typically done approximately using perturbation theory in the fermion gauge-field coupling.[15]

Thus the decoupling (2) leads to a description of the spin model as a theory of fermions (spinons) coupled to an SU(2) gauge field. For the Hamiltonian (1), we will find that the spinons are in a superconducting phase, such that this gauge symmetry is broken down to Z_2. The SU(2) gauge fields are therefore fully gapped, such that the effect of dynamical gauge-field fluctuations on the fermion band structure is minimal. We will nonetheless find that this gauge theory is a useful tool to understand the origin of the various topologically ordered phases described in Ref. 1.

We begin our analysis with the mean-field description of the exact spin-liquid ground state of the Hamiltonian (1). In the case of spin-rotationally invariant Hamiltonians, such as the Heisenberg model, the Hamiltonian simplifies considerably when written in terms of the fermions (2). In the absence of spin-rotational symmetry, as in Eq. (1), the Hamiltonian is more complicated. For instance, the Hamiltonian on x links takes the form

$$\hat{S}^x_i \hat{S}^x_j = -\tfrac{1}{4}[f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} f_{i\downarrow} f_{j\downarrow} + f^\dagger_{i\downarrow} f^\dagger_{j\downarrow} f_{i\uparrow} f_{j\uparrow} + f^\dagger_{i\uparrow} f_{j\uparrow} f^\dagger_{j\downarrow} f_{i\downarrow} + f^\dagger_{i\downarrow} f_{j\downarrow} f^\dagger_{j\uparrow} f_{i\uparrow}] \tag{4}$$

with similar terms on the y links, as detailed in Appendix B. (This form is not unique; using the constraints, it can be rewritten in different forms that are equivalent in the constraint subspace.) In the Heisenberg model, by contrast, the Hamiltonian on each link can be written in the form

$$\hat{S}^x_i \hat{S}^x_j + \hat{S}^y_i \hat{S}^y_j + \hat{S}^z_i \hat{S}^z_j = \sum_{\alpha,\beta} \frac{1}{4} f^\dagger_{i\alpha} f_{i\alpha} f^\dagger_{j\beta} f_{j\beta} - \frac{1}{2} f^\dagger_{i\alpha} f_{j\alpha} f^\dagger_{j\beta} f_{i\beta}.$$

As a result of the more complex form of the Hamiltonian, it is necessary to introduce four Hubbard-Stratonovich fields to decouple the four Fermi interactions. For example, the Lagrangian on the x links can be written in the form

$$\mathcal{L}_x = -\frac{8(|\Phi_1|^2 + |\Phi_2|^2)}{J_x} - \frac{8(|\Theta_1|^2 + |\Theta_2|^2)}{J_x} + \Phi_1(f^\dagger_{i\uparrow} f_{j\uparrow} + f^\dagger_{i\downarrow} f_{j\downarrow}) + i\Phi_2(f^\dagger_{i\uparrow} f_{j\uparrow} - f^\dagger_{i\downarrow} f_{j\downarrow}) + \tilde{\text{H.c.}} + \Theta_1(f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} + f^\dagger_{i\downarrow} f^\dagger_{j\downarrow}) + i\Theta_2(f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} - f^\dagger_{i\downarrow} f^\dagger_{j\downarrow}) + \tilde{\text{H.c.}},$$

where $\tilde{\text{H.c.}}$ is the Hermitian conjugate with all spin directions reversed. The Lagrangian can be decoupled in a similar manner on the y and x links as well, as detailed in Appendix B.

Before proceeding, it will be helpful to pick a unit cell for the honeycomb lattice. We will label the two different sites in the unit cell by the index $u = 1,2$ and different unit cells by $\mathbf{R} = n_1\hat{\mathbf{x}} + n_2(\frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}})$. Then, we denote the fermion fields by $f_{\mathbf{R}u\sigma}$. (We will continue to use $f_{i\sigma}$ to denote a fermion operator on site i in either sublattice.) Their Fourier transforms are defined by

$$f_{\mathbf{q}u\sigma} = \frac{1}{\sqrt{N}} \sum_{\vec{R}} e^{i\mathbf{R}\cdot\mathbf{q}} f_{\mathbf{R}u\sigma}, \tag{5}$$

where N is the total number of unit cells.

To proceed, we assume that Φ_i, Θ_i acquire nonzero expectation values. We parametrize these expectation values by $t_{ij,\alpha}$, $\Delta_{ij,\alpha}$, $\alpha = \uparrow, \downarrow$, as explained in Appendix B 2. Unlike in the case of Heisenberg interactions, to describe the Kitaev model, we must condense both hopping and superconducting order parameters or else the mean-field equations will not be satisfied (except in the special case $J_x = J_y = 0$, $J_z \neq 0$), as shown below. (In the Heisenberg case, hopping and d-wave superconducting terms can be rotated into each other by a gauge transformation. This is not true for the p-wave superconducting case considered here.) Because SU(2) spin-rotation invariance is explicitly broken on each link, the superconducting terms $\Delta_\uparrow, \Delta_\downarrow$ are spin polarized. Thus, replacing the fields Φ_i, Θ_i by their expectation values, we obtain the mean-field Hamiltonian

$$H_{\text{MF}} = \frac{1}{2}\sum_{\mathbf{q},\sigma} \psi^\dagger_{\mathbf{q}\sigma} \times \begin{bmatrix} 0 & t_\sigma(\mathbf{q}) & 0 & \Delta_\sigma(\mathbf{q}) \\ t^*_\sigma(\mathbf{q}) & 0 & -\Delta_\sigma(-\mathbf{q}) & 0 \\ 0 & -\Delta^*_\sigma(-\mathbf{q}) & 0 & -t^*_\sigma(-\mathbf{q}) \\ \Delta^*_\sigma(\mathbf{q}) & 0 & -t_\sigma(-\mathbf{q}) \end{bmatrix} \psi_{\mathbf{q}\sigma},$$

$$\psi^\dagger_{\mathbf{q}} = (f^\dagger_{\mathbf{q},1,\sigma} \quad f^\dagger_{\mathbf{q},2,\sigma} \quad f_{-\mathbf{q},1,\sigma} \quad f_{-\mathbf{q},2,\sigma}). \tag{6}$$

(Here, the factor of $\frac{1}{2}$ in the first line compensates for the fact that the expression (6) counts each term in the Hamiltonian twice. Alternatively, we could sum over half the Brillouin zone.) If we write ψ_q in components, it has three indices (in addition to momentum) $\psi_{\mathbf{q}l\sigma a}$, where $u = 1,2$ is a sublattice index, $\sigma = \uparrow, \downarrow$ is a spin index, and $a = \pm$ is a particle-hole index.

Since we will often be using Pauli matrices to act on these indices, we will, to avoid confusion, introduce three different notations for Pauli matrices. We will use $\sigma^{x,y,z}_{\alpha\beta}$ for Pauli matrices acting on spin indices; $\mu^{x,y,z}_{uv}$ for Pauli matrices acting on sublattice indices; and $\tau^{x,y,z}_{ab}$ for Pauli matrices acting on particle-hole indices. (Of course, it is precisely the same three matrices in all three cases.)

By requiring self-consistency of the expectation values, we can express $t_{ij,\alpha}$, $\Delta_{ij,\alpha}$ in terms of $J_{x,y,z}$, as shown in Eq. (B7). At the saddle point of interest, the relevant parameters are

$$\begin{aligned} t_\uparrow(q) &= -\frac{i}{16}(e^{i\vec{q}\cdot\hat{l}_1}J_x + e^{i\vec{q}\cdot\hat{l}_2}J_y),\\ \Delta_\uparrow(q) &= -\frac{i}{16}(e^{i\vec{q}\cdot\hat{l}_2}J_y - e^{i\vec{q}\cdot\hat{l}_1}J_x),\\ t_\downarrow(q) &= -\frac{i}{16}(e^{i\vec{q}\cdot\hat{l}_1}J_x + e^{i\vec{q}\cdot\hat{l}_2}J_y + 2J_z),\\ \Delta_\downarrow(q) &= \frac{i}{16}(e^{i\vec{q}\cdot\hat{l}_1}J_x + e^{i\vec{q}\cdot\hat{l}_2}J_y), \end{aligned} \tag{7}$$

where $\hat{l}_{1,2} = \frac{\sqrt{3}}{2}\hat{y} \pm \frac{1}{2}\hat{x}$ are the lattice vectors.

The band energies and eigenfunctions of $H_{\rm MF}$ reveal the correspondence between this picture and the Majorana fermion decoupling of Ref. 1. The mean-field spectrum consists of three flat bands, with energies

$$\epsilon_{\uparrow x} = \pm\frac{J_x}{8}, \quad \epsilon_{\uparrow y} = \pm\frac{J_y}{8}, \quad \epsilon_{\downarrow z} = \pm\frac{J_z}{8}, \tag{8}$$

and one dispersing band of energy

$$\epsilon_\downarrow(q) = \pm\tfrac{1}{8}|J_x e^{i\vec{q}\cdot\hat{l}_1} + J_y e^{i\vec{q}\cdot\hat{l}_2} + J_z|. \tag{9}$$

[Since we have included an explicit factor of 1/2 in our definition of the spin operators $\vec{S}_i$, our $J_{x,y,z}$ are four times larger than those of Kitaev. There is an additional explicit factor of 4 in his definition of the spectrum in Eqs. (31) and (32) in Ref. 1. This accounts for the factor 16 between our spectra.] The corresponding eigenvectors are naturally expressed in terms of the Majorana fermions

$$\begin{aligned} b^y_{\mathbf{q}u} &= f^\dagger_{\mathbf{q}u\uparrow} + f_{-\mathbf{q}u\uparrow}, \qquad b^x_{\mathbf{q}u} = i(f^\dagger_{\mathbf{q}u\uparrow} - f_{-\mathbf{q}u\uparrow}),\\ b^z_{\mathbf{q}u} &= f^\dagger_{\mathbf{q}u\downarrow} + f_{-\mathbf{q}u\downarrow}, \qquad c_{\mathbf{q}u} = i(f^\dagger_{\mathbf{q}u\downarrow} - f_{-\mathbf{q}u\downarrow}). \end{aligned} \tag{10}$$

We have used the same labels as Ref. 1 for these operators.

However, this is not a unique mapping. For instance, we could, instead, take $c = -(f^\dagger_\uparrow + f_\uparrow)$, $b^x = i(f^\dagger_\downarrow - f_\downarrow)$, $b^y = f^\dagger_\downarrow + f_\downarrow$, $b^z = i(f^\dagger_\uparrow - f_\uparrow)$. Furthermore, the mean-field Hamiltonian has a different expression in terms of these operators than in the mean-field theory of Ref. 1. For example, the bilinears $b^z_{R,1}b^z_{R,2}$ do not commute with the mean-field Hamiltonian. The reason for this is that, if the spin operators are expressed in terms of the $f, f^\dagger$s according to Eq. (2), and then the $f, f^\dagger$s are written in terms of c,b^x,b^y,b^z, according to Eq. (10), then we will not obtain the same representation as in Ref. 1. Only after the constraints are imposed do the operators in Eq. (10) become equivalent to those of Kitaev. This is explained in more detail in Appendix A.

The eigenvectors corresponding to the eigenvalues (8) and (9) are given by

$$\begin{aligned} \alpha_{x\pm}(q) &= \tfrac{1}{2}\big(ie^{i\vec{q}\cdot\hat{l}_1}b^x_{q,1} \pm b^x_{q,2}\big),\\ \alpha_{y\pm}(q) &= \tfrac{1}{2}\big(ie^{i\vec{q}\cdot\hat{l}_2}b^y_{q,1} \pm b^y_{q,2}\big),\\ \alpha_{z\pm}(q) &= \tfrac{1}{2}\big(ib^z_{q,1} \pm b^z_{q,2}\big),\\ \alpha_{0\pm}(q) &= \tfrac{1}{2}(ie^{i\theta_q}c_{q,1} \pm c_{q,2}), \end{aligned} \tag{11}$$

where $\theta_q = {\rm Arg}(J_x e^{i\vec{q}\cdot\hat{l}_1} + J_y e^{i\vec{q}\cdot\hat{l}_2} + J_z)$, and in all cases, plus sign corresponds to the negative-energy solution. The $b^\alpha_{q,i}$ therefore lie in the three flat bands, and are localized on x, y, and z links, respectively, and c is the dispersing Majorana mode identified by Ref. 1.

Hence, the saddle point (7) reproduces exactly the description of Ref. 1, with the precise mapping between the fermions $f_{qu\sigma}$ and Kitaev's Majorana fermions given by Eq. (10). The only difference is that Ref. 1 does not include the energy of the flat bands, so that $b^{x,y,z}$ enter only in determining the band structure of the remaining Majorana mode c. The fermionic mean-field energy we obtain per unit cell at half-filling is

$$-\frac{1}{8}(J_x + J_y + J_z) - \frac{2}{N}\sum_q \epsilon_q. \tag{12}$$

However, the first term is canceled by the zero-point energy arising from terms of the form $\frac{|\Phi_i|^2}{J_{x,y,z}}$, $\frac{|\Theta_i|^2}{J_{x,y,z}}$ in the Hubbard-Stratonovich Hamiltonian, so we are left with precisely the same energy as in Kitaev's solution.

Superficially, we have obtained an eight-band mean-field theory from a model of spinful fermions on a lattice with a two-site unit cell. Readers might, thus, justifiably be concerned that we have in fact obtained double the degrees of freedom that we would have expected. However, we have combined $f_{\mathbf{q}u\sigma}$ and $f^\dagger_{-\mathbf{q}u\sigma}$ into the same spinor; consequently, we should restrict $\mathbf{q}$ to half the Brillouin zone to avoid double counting.

B. Slave fermion band structure

To understand the physics of this model, it is useful to focus on the band structure of the down-spin fermions. It suffices to consider the case $J_x = J_y = J$:

$$\epsilon_\downarrow(q) = \pm\frac{J}{8}\left\{\left(\frac{J_z}{J} + 2\cos\frac{q_x}{2}\cos\frac{\sqrt{3}q_y}{2}\right)^2 + 4\left(\cos\frac{q_x}{2}\sin\frac{\sqrt{3}q_y}{2}\right)^2\right\}^{1/2}. \tag{13}$$

This describes a pair of bands that cross at either 0 or 2 distinct points in the Brillouin zone, as shown in Fig. 1. Following Ref. 1, we will call the former case, which occurs for $|J_z| > 2|J|$, the A phase. In the A phase, the spectrum is fully gapped. When $|J_z| < 2|J|$, there are two Majorana cones in the spectrum or, equivalently, a single Dirac cone. This is the B phase. Our objective here is to understand how this band structure arises in the slave fermion superconductor, and use this analogy to understand the transitions between these phases.

We begin with a more scrupulous analysis of the nature of the superconducting state. Since the character of the phase is determined by the dispersing fermion band, we will focus on the mean-field Hamiltonian for the down spins. If we combine the down-spin fermions on the two sublattices into the spinor

$$\Psi_q = \begin{pmatrix} f_{q1\downarrow} \\ f_{q2\downarrow} \end{pmatrix}, \tag{14}$$

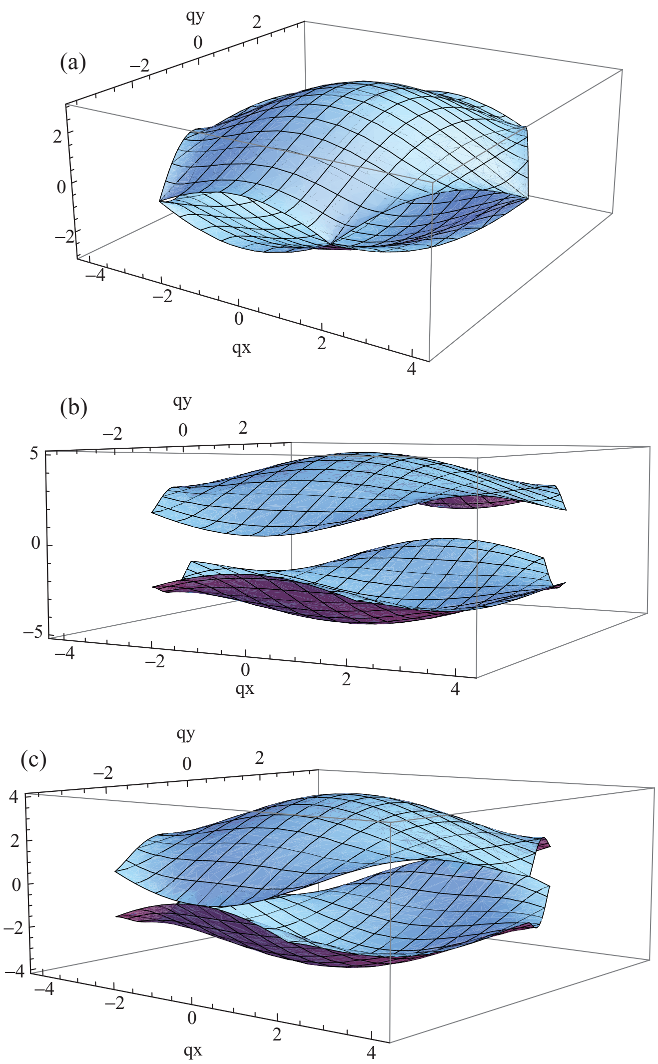

FIG. 1. (Color online) Fermionic spectra for the dispersing mode of the fermionized Hamiltonian, in the gapless B phase (a), and gapped A phase (b). (The remaining 6 bands of the Hamiltonian (6) are flat, and lie at energies $\pm J_x$, $\pm J_y$, and $\pm J_z$). In the gapless phase the Fermi surface consists of nodes at two distinct points in the hexagonal Brillouin zone. The loci of these nodes varies with the relative magnitudes of J_x, J_y, and J_z. The transition to the gapped A phase occurs when the pair of nodes come together and annihilate, as shown in (c).

then the Hamiltonian has the general form

$$\begin{aligned} H_{\rm down} &= \Psi_q^\dagger \left[\epsilon_q^{(x)}\mu_x + \epsilon_q^{(y)}\mu_y\right]\Psi_q \\ &\quad + \Psi_q^\dagger\left(\Delta_q^{(s)}\mu_y + \Delta_q^{(t)}\mu_x\right)(\Psi_{-q}^\dagger)^T + \text{H.c.} \\ &\quad + \frac{J_z}{8}\left(2 - \frac{J}{J_z}\right)\Psi_q^\dagger\mu_y\Psi_q, \end{aligned} \tag{15}$$

where we have taken $J_x = J_y = J$, and

$$\begin{aligned} \epsilon_q^{(x)} &= \frac{J}{8}\cos\frac{q_x}{2}\sin\frac{\sqrt{3}q_y}{2}, \\ \epsilon_q^{(y)} &= \frac{J}{16}\left(1 + 2\cos\frac{q_x}{2}\cos\frac{\sqrt{3}q_y}{2}\right) \end{aligned} \tag{16}$$

represent the kinetic energy for fermions hopping on the honeycomb lattice. The third line corresponds to an in-plane "magnetic field" in pseudospin space due to the enhanced hopping along the z links. This term shifts the positions of the Majorana cones, but is otherwise unremarkable.

The second line is a superconducting pairing term along the x and y links. Both

$$\begin{aligned} \Delta_q^{(s)} &= \frac{J}{8}\cos\frac{\sqrt{3}q_y}{2}\cos\frac{q_x}{2}, \\ \Delta_q^{(t)} &= -\frac{J}{8}\sin\frac{\sqrt{3}q_y}{2}\cos\frac{q_x}{2} \end{aligned} \tag{17}$$

are nonvanishing in the mean-field state. The superscipts (s) and (t) refer to the fact that these are pseudospin-singlet and pseudospin-triplet superconducting order parameters.

If we linearize about the nodes (we work at the isotropic point $J = J_z$ for simplicity), then the Hamiltonian for down spins takes the form

$$\begin{aligned} H_{\rm down} &= \Psi_p^\dagger\left[-\frac{J\sqrt{3}}{32}p_y\mu_x + \frac{J\sqrt{3}}{32}p_x\mu_y - \frac{J}{16}\mu_y\right]\Psi_p \\ &\quad - \frac{J}{16}\Psi_p^\dagger\mu_y(\Psi_{-p}^\dagger)^T + \text{H.c.} \\ &\quad + \frac{J\sqrt{3}}{32}\Psi_p^\dagger[p_y\mu_x - p_x\mu_y](\Psi_{-p}^\dagger)^T + \text{H.c.} \end{aligned} \tag{18}$$

Here, $\vec{p}$ is the distance from the node $(4\pi/3,0)$. This Hamiltonian has four eigenvalues, the two nondispersing ones $\pm J_z/8$ and the two dispersing ones in Eq. (13).

It is helpful to isolate the dispersing band. [The Hamiltonian (15) contains both the dispersing and nondispersing down-spin bands]. To this end, we form the Dirac fermion

$$\eta_q = e^{i\pi/4}(c_{q1} - ic_{q2}). \tag{19}$$

The mean-field Hamiltonian for η_q is (up to a constant)

$$\tilde{H} = \frac{1}{2}\sum_q(\epsilon_q\eta_q^\dagger\eta_q + \Delta_q\eta_q^\dagger\eta_{-q}^\dagger + \text{H.c.}), \tag{20}$$

where

$$\epsilon_q = \frac{1}{8}\left(J_z + 2J\cos\frac{q_x}{2}\cos\frac{\sqrt{3}}{2}q_y\right), \tag{21}$$

$$\Delta_q = \frac{1}{4}J\cos\frac{q_x}{2}\sin\frac{\sqrt{3}}{2}q_y. \tag{22}$$

To understand this Hamiltonian better, it is useful to momentarily imagine that $\Delta_q = 0$ and focus on ϵ_q. The Hamiltonian now describes spinless fermions on the honeycomb lattice with dispersion ϵ_q. First, consider $J_z > 2J$. We see that there is no Fermi surface: ϵ_q is never equal to zero. Consider the minimum energy excitation, which occurs at $\vec{q} = (0, \frac{2\pi}{\sqrt{3}})$ and has energy $J_z - 2J$. Near the minimum, the band is approximately quadratic. There are no excitations near zero energy because the effective "Fermi energy" lies below the bottom of the band. Superconductivity does not change this picture very much, other than to break U(1) symmetry (which is very important when we go beyond the mean field). When superconductivity is turned back on, there are no nodes or nodal excitations because there is no Fermi surface.

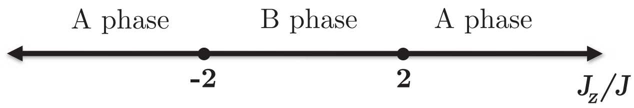

FIG. 2. Schematic phase diagram of the Hamiltonian (1).

For $J_z < 2J$, there is a Fermi surface that surrounds the point $(0,\frac{2\pi}{\sqrt{3}})$. Strictly speaking, for the usual Brillouin zone, this point sits on its boundary, so half the Fermi surface encircles $(0,\frac{2\pi}{\sqrt{3}})$, while the other half encircles the equivalent point $(0,-\frac{2\pi}{\sqrt{3}})$, which differs by a reciprocal lattice vector. Of course, we could take a different unit cell for the reciprocal lattice that only includes one of these two points; then, the Fermi surface will surround this point. We now restore the superconducting gap Δ_q. This opens a gap on the Fermi surface, except at the points on the Fermi surface that intersect the nodal line $(q_y = \frac{2\pi}{\sqrt{3}})$. [The nodal line $q_y = 0$ does not intersect the Fermi surface, except for the point $(4\pi/3,0)$, which is equivalent to $(2\pi/3,2\pi/\sqrt{3})$ under translation by a reciprocal lattice vector.] For $2 - J_z/J \ll 1$ small, the Fermi surface is approximately circular. Let us expand momenta about $(0,\frac{2\pi}{\sqrt{3}})$ so that $(q_x,q_y) \approx (0,\frac{2\pi}{\sqrt{3}}) + (2p_x,2p_y/\sqrt{3})$. Then $\epsilon_p \approx J(p_x^2 + p_x^2) - \mu$, where the Fermi energy μ is given by $\mu = 2J - J_z$, and $\Delta_p = J\,p_y$. Thus, the Hamiltonian in the B phase looks like that of a p_y superconductor, which has nodes at $p_y = 0$. As J_z is decreased and the system moves toward the isotropic point, the nodes move toward the corners of the Brillouin zone, eventually reaching the graphene spectrum at the isotropic point.

III. MEAN-FIELD PHASE DIAGRAM IN THE ABSENCE OF TIME-REVERSAL SYMMETRY-BREAKING PERTURBATIONS

We will now apply the mean-field description outlined in the preceding section to understanding the phase diagram of (1) in terms of its fermionic band structure and superconducting gap. For reference, a schematic mean-field phase diagram is shown in Fig. 2. As we shall see, the principal advantage of the spinful mean-field decoupling is that it allows us to better understand the system's behavior away from the exactly solvable point, both in terms of proximate phases and the fate of physical quantities such as the spin-spin correlation functions as we perturb the Hamiltonian (1). At the end of this section, we also describe at mean-field level the nature of the phase transition separating the gapped A phase and gapless B phase.

A. The A phase

We begin by studying the A phase, for which $J_z > 2J$ and the band structure (13) is fully gapped. In this phase, superconductivity, which couples fermions along the x and y links, competes with dimerization along the z links, as is evident from the two-band Hamiltonian (15). In the A phase, the dimerization term dominates, leading to a fully gapped band structure. In the extreme limit $J = 0$, $J_z \neq 0$, dimerization leads to a gap, even in the absence of superconductivity. (Indeed, many fruitful explorations of the A phase treat it as an effective theory of such interacting dimers.[16–19])

As seen at the end of the preceding section, we may view the A phase as a spin-polarized p-wave superconductor with chemical potential which lies below the conduction band. One amusing consequence of this is that the topological order of this phase is, as explained in Ref. 20, that of a Z_2 gauge theory. Its topological nature stems from the fact that, in the condensed phase, the only remnant of the interactions between gauge fields and matter is a "statistical" interaction[21] due to the Berry phase of π accrued by a charge if it encircles a vortex of flux $\frac{\hbar}{2e}$. This provides an alternative perspective on the well-documented fact[1,16] that the A phase is smoothly connected to the so-called toric code,[2] a model of Ising spins that realizes a topological Z_2 gauge theory with matter. In particular, this highlights that the topological order of the A phase is not restricted to the set of exactly solvable Hamiltonians described by (1), but is that of a garden-variety s-wave superconductor.

If we only cared about the single-particle gap, then we could close the superconducting gap entirely without closing the total fermion gap. However, the gauge symmetry of the problem would not be broken down to Z_2 in this case, so there would be gapless gauge-field fluctuations about the mean-field solution. [In the dimerized limit where $J_z = J_y = 0$, these gapless modes are absent since the gauge field cannot propagate, even though the U(1) gauge symmetry is unbroken.]

Because the A phase is fully gapped, it is stable to weak perturbations away from the solvable point discussed here. For instance, we could add a weak magnetic field and/or Heisenberg interaction without changing the qualitative features of this phase. Since the system is fully gapped, perturbation theory can be used and the effect will be small, so long as the perturbation is weak. This is in contrast to the B phase, which, as we will see, is unstable in the face of appropriately chosen perturbations.

B. The nodal B phase

We now briefly describe the B phase, for which $J_z < 2J$. Now Eq. (20) is the band structure of a p-wave superconductor, whose nodes of which intersect the Fermi surface at two distinct points in the Brillouin zone.

To simplify the algebra, we will consider the symmetric point $J_x = J_y = J_z \equiv J$. The energies of the dispersing Majorana bands are then exactly those of free fermions in a honeycomb lattice. The spectrum is gapless at the points $\vec{q} = (\pm\frac{2\pi}{3},\frac{2\pi}{\sqrt{3}})$ [and at the equivalent points $(\pm\frac{4\pi}{3},0)$, $(\frac{2\pi}{3},-\frac{2\pi}{\sqrt{3}})$, which differ from the first two by reciprocal lattice vectors]. These nodes account for two distinct cones in the energy spectrum, as in graphene. However, unlike in graphene, the band structure (13) is that of a pair of bands of dispersing *Majorana* fermions. In the vicinity of these nodal points, it is useful to rewrite the Hamiltonian (20) in terms of the spinor

$$\chi_q = \begin{pmatrix} \eta_q \\ \eta_{-q}^\dagger \end{pmatrix}, \tag{23}$$

where $\vec{q}$ is restricted to lie in half of the Brillouin zone to avoid double counting, e.g., over $q_x > 0$. In terms of this spinor, the Hamiltonian can be written in the form

$$H = \frac{1}{2}\sum_{q_x>0,q_y} \chi_q^\dagger[\Delta_q\tau_x + \epsilon_q\tau_z]\chi_q. \tag{24}$$

In the vicinity of the nodes (at the isotropic point $J_z = J$), we can expand $\vec{q} = (\frac{4\pi}{3},0) + (p_x, p_y)$ and write

$$\tilde{\chi}_p = \begin{pmatrix} \eta_{(\frac{4\pi}{3},0)+\vec{p}} \\ \eta^\dagger_{-(\frac{4\pi}{3},0)-\vec{p}} \end{pmatrix}, \tag{25}$$

and $\vec{p}$ now ranges unrestricted over small $\vec{p}$ (e.g., over $|\vec{p}| < \Lambda$, for some cutoff Λ), i.e., near the nodes. Expanding $\epsilon = \frac{\sqrt{3}J}{16} p_y, \Delta = \frac{\sqrt{3}J}{16} p_x$, we can write

$$\begin{aligned} H &= \sum_{\vec{p}} \tilde{\chi}_p^\dagger \left[\frac{\sqrt{3}J}{32} p_y \tau_x + \frac{\sqrt{3}J}{32} p_x \tau_z \right] \tilde{\chi}_p \\ &= v \int d^2x\, \tilde{\chi}^\dagger [i\partial_y \tau_y + i\partial_x \tau_z] \tilde{\chi} \end{aligned} \tag{26}$$

with $v = \frac{\sqrt{3}}{32} J$. Thus, these two Majorana fermions combine to form a single Dirac fermion. This Dirac cone is formed by combining the two nodes of a p_y superconductor. This single Dirac cone does not violate the usual fermion doubling arguments since the gauge symmetry is broken. We will see presently, however, that it is central to the non-Abelian statistics of the gapped B* phase.

We now consider some of the correlation functions of the B phase. Since there are gapless excitations, the energy density will certainly have power-law correlations. How about the spin-spin correlation function? At the soluble point, this is short ranged. Consider, for instance, the S^z-S^z correlation. In terms of the slave fermions, $S_i^z = (f^\dagger_{i\uparrow} f_{i\uparrow} - f^\dagger_{i\downarrow} f_{i\downarrow})/2$. Since up and down spins decouple,

$$\langle S_i^z S_j^z \rangle = \tfrac{1}{4} \langle f^\dagger_{i\uparrow} f_{i\uparrow} f^\dagger_{j\uparrow} f_{j\uparrow} \rangle + \tfrac{1}{4} \langle f^\dagger_{i\downarrow} f_{i\downarrow} f^\dagger_{j\downarrow} f_{j\downarrow} \rangle. \tag{27}$$

The first term vanishes since it only involves b^x and b^y, and these create and annihilate fermions in the up-spin flat bands. Here, b^x and b^y are defined in terms of $f_\downarrow$, $f_\downarrow^\dagger$ according to Eq. (10). (It is important to remember that, although they play the same role in our analysis as the operators with the same labels in Ref. 1, they are not identical, in spite of the obvious similarity.) Thus, we are left with

$$\begin{aligned} \langle S_i^z S_j^z \rangle &= \langle f^\dagger_{i\downarrow} f_{i\downarrow} f^\dagger_{j\downarrow} f_{j\downarrow} \rangle / 4 \\ &= \left(1 + \langle i b_i^z c_i \rangle + \langle i b_j^z c_j \rangle - \langle b_i^z c_i\, b_j^z c_j \rangle \right)/16 = 0. \end{aligned} \tag{28}$$

At the mean-field level, this is a free fermion problem, so we can evaluate these correlation functions. The Hamiltonian does not mix b^z with c, so $\langle i b_i^z c_i \rangle = 0$ and $\langle b_i^z c_i\, b_j^z c_j \rangle = \langle b_i^z b_j^z \rangle \langle c_j c_i \rangle$. Since b^z creates a fermion in a flat, nondispersing band, $\langle b_i^z b_j^z \rangle = 0$ unless i and j are the same or neighboring sites.

One of the appealing features of the formalism we use is that correlation functions in the presence of small perturbations to the Hamiltonian (1) can be calculated with relative ease. For instance, suppose we consider a weak magnetic field in the z direction, as in Ref. 22. This adds a perturbation to the Hamiltonian

$$H_{\text{pert}} = \frac{1}{2} h_z \sum_i (f^\dagger_{i\uparrow} f_{i\uparrow} - f^\dagger_{i\downarrow} f_{i\downarrow}). \tag{29}$$

For small h_z, this perturbation does not spoil the basic structure of the spectrum: there are still three gapped bands and one gapless one. The up-spin gapped band will still be nondispersing and will be at the same energy, but the corresponding eigenoperators will mix b^x and b^y [unlike the eigenoperators (11) in the unperturbed Hamiltonian]. The down-spin gapped band will now disperse weakly, but will remain gapped. However, the eigenoperators for the down-spin bands will now mix b^z and c. Thus, when we compute the $\langle S_i^z S_j^z \rangle$ correlation function, b^z will have a small amplitude, proportional to h_z for small h_z, to create a dispersing fermion. Thus, this correlation function will have power-law falloff.

To see this more precisely, we add the magnetic-field term to the down-spin Hamiltonian

$$\begin{aligned} H_{\text{down}} &= \Psi_p^\dagger \left[-\frac{J\sqrt{3}}{32} p_y \mu_x + \frac{J\sqrt{3}}{32} p_x \mu_y - \frac{J}{16} \mu_y \right] \Psi_p \\ &\quad + \frac{J\sqrt{3}}{32} \Psi_p^\dagger [p_y \mu_x - p_x \mu_y] (\Psi_{-p}^\dagger)^T + \text{H.c.} \\ &\quad - \frac{J}{16} \Psi_p^\dagger \mu_y (\Psi_{-p}^\dagger)^T + \text{H.c.} - \frac{1}{2} h_z \Psi_p^\dagger \Psi_p. \end{aligned} \tag{30}$$

When we diagonalize this Hamiltonian, we find a new set of eigenoperators $\tilde{\alpha}_{z\pm}$, $\tilde{\alpha}_{0\pm}$. The eigenoperator $\tilde{\alpha}_{z+}$ creates a fermion in a weakly dispersing gapped band and has short-ranged correlation functions. The eigenoperator $\tilde{\alpha}_{0+}$ creates a fermion in a gapless band and has power-law correlation functions. For small h_z (and, for simplicity, small momentum k), we can express the fermions $\alpha_{z\pm} = (i b^z_{q,1} \pm b^z_{q,2})/2$, $\alpha_{0\pm} = (i e^{i\theta_q} c_{q,1} \pm c_{q,2})/2$, in terms of these new eigenoperators as

$$\begin{aligned} \alpha_{z\pm} &= \tilde{\alpha}_{z\pm} \pm \frac{h_z}{2} \tilde{\alpha}_{0\pm}, \\ \alpha_{0\pm} &= \mp \frac{h_z}{2} \tilde{\alpha}_{z\pm} + \tilde{\alpha}_{0\pm}. \end{aligned} \tag{31}$$

Thus, we now have

$$\begin{aligned} \langle b_i^z b_j^z \rangle &= -\langle (\alpha_{z+,i} + \alpha_{z-,i})(\alpha_{z+,j} + \alpha_{z-,j}) \rangle \\ &= -\langle [\tilde{\alpha}_{z+,i} + \tilde{\alpha}_{z-,i} + h_z(\tilde{\alpha}_{0+,i} - \tilde{\alpha}_{0-,i})/2] \\ &\quad \times [\tilde{\alpha}_{z+,j} + \tilde{\alpha}_{z-,j} + h_z(\tilde{\alpha}_{0+,j} - \tilde{\alpha}_{0-,j})/2] \rangle \\ &= \langle \tilde{\alpha}_{z+,i} \tilde{\alpha}_{z+,j} \rangle + \langle \tilde{\alpha}_{z-,i} \tilde{\alpha}_{z-,j} \rangle \\ &\quad + \frac{h_z^2}{4} (\langle \tilde{\alpha}_{0+,i} \tilde{\alpha}_{0+,j} \rangle + \langle \tilde{\alpha}_{0-,i} \tilde{\alpha}_{0-,j} \rangle). \end{aligned} \tag{32}$$

Here, we have assumed that, for the sake of concreteness and simplicity, the sites i and j are on the 1 sublattice. From the Hamiltonian (30), we have, for large separation $|\mathbf{x} - \mathbf{y}|$ and to zeroth order in h_z,

$$\begin{aligned} &\langle \tilde{\alpha}_{0+,\mathbf{x}} \tilde{\alpha}_{0+,\mathbf{y}} \rangle + \langle \tilde{\alpha}_{0-,\mathbf{x}} \tilde{\alpha}_{0-,\mathbf{y}} \rangle \\ &\quad = \int \frac{d\omega}{2\pi} \frac{d^2k}{(2\pi)^2} \frac{\frac{J\sqrt{3}}{16}(k_y + i k_x/2)\, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}}{\omega^2 - \left(\frac{J\sqrt{3}}{16}\right)^2 \left(k_x^2 + 4k_y^2\right)}. \end{aligned} \tag{33}$$

Therefore, at long distances,

$$\langle \tilde{\alpha}_{0\pm,\mathbf{x}} \tilde{\alpha}_{0\pm,\mathbf{y}} \rangle \sim \frac{1}{|\mathbf{x} - \mathbf{y}|^2}. \tag{34}$$

Combined with the $\langle c_i c_j \rangle$, which has the same power law, this gives an $\langle S_i^z S_j^z \rangle$ correlation function that falls off as $1/r^4$ in the presence of a small magnetic field, in agreement with the results of Ref. 22.

In the face of perturbations that are not quadratic in the fermions, such explicit calculations are more difficult in general. However, as is frequently the case in spin-liquid models,[8] the structure of the Fermi surface (here a pair of Dirac cones) is protected by symmetries of the mean-field state. Thus, small perturbations which do not break any symmetries of the problem can not open a gap in the spectrum.

C. Transition between A and B phases

As we move within the gapless B phase, from the isotropic point $J_x = J_y = J_z$ toward the boundary to the A phase, the two nodal points move together and, at the phase-transition point, merge. The nodes then annihilate as the phase boundary is crossed. In this section, we focus on the transition point.

As discussed in Sec. II, the dispersing spin-down band can be rewritten as a model of spinless fermions with p_y superconducting order, as in Eq. (20). At the boundary between the A and B phases, the Fermi surface has shrunk to a point because the effective chemical potential is precisely at the bottom of the band. When the effective chemical potential is at the bottom of the band, the spectrum is quadratic in the absence of superconductivity. Superconductivity with p_y pairing symmetry leaves the spectrum gapless but makes the spectrum linear in one direction. We now examine this in more detail. Expanding about the bottom of the band $(q_x,q_y) \approx (0,\frac{2\pi}{\sqrt{3}}) + (2p_x,2p_y/\sqrt{3})$, we can write the Hamiltonian (20) in the form

$$\begin{aligned}\tilde{H} &= \frac{1}{2}\sum_p \left[\frac{J}{8}p^2\,\eta_p^\dagger\eta_p - \frac{J}{4}p_y\,(\eta_p^\dagger\eta_{-p}^\dagger - \eta_p\eta_{-p})\right]\\ &= \frac{1}{2}\sum_{p_x>0,p_y}\chi_{-p}^T\left[-\frac{J}{4}p_y I - \frac{J}{8}p^2 i\tau_y\right]\chi_p,\end{aligned} \tag{35}$$

where

$$\chi_p = \begin{pmatrix}\eta_p\\ \eta_{-p}^\dagger\end{pmatrix}. \tag{36}$$

If we go to a Majorana basis

$$\varphi_p = \frac{1}{\sqrt{2}}\begin{pmatrix}\eta_p + \eta_{-p}^\dagger\\ (\eta_p - \eta_{-p}^\dagger)/i\end{pmatrix}, \tag{37}$$

this can be rewritten as

$$\begin{aligned}H &= \frac{1}{2}\sum_{p_x>0,p_y}\varphi_{-p}^T\left[-\frac{J}{4}p_y\tau_z - \frac{J}{8}p^2\tau_y\right]\varphi_p\\ &= \frac{1}{2}\int d^2x\,\varphi^T\left[-\frac{J}{4}i\partial_y\tau_z - \frac{J}{8}\partial^2\tau_x\right]\varphi.\end{aligned} \tag{38}$$

Therefore, the low-energy theory can be called a single gapless Majorana fermion, albeit an anisotropic and nonrelativistic one.

IV. BEYOND MEAN-FIELD THEORY

Thus far, we have found a consistent mean-field solution of (1) using the fermionization (2), which reproduces exactly the Majorana fermion band structure and phase diagram of the exact solution proposed by Ref. 1. We next ask what can be said about its fate upon including fluctuations of the various bosonic fields. The answer is not obvious since, unlike the decoupling used by Ref. 1, the product $b_i^\alpha b_{i+1}^\alpha$ on each link does not commute with the full unprojected fermion Hamiltonian (although it does commute with the quadratic Hamiltonian $H_{\rm MF}$). Here, we first establish that these fluctuations do not alter the results of the previous sections. Second, we demonstrate that, at long wavelengths, these bosonic modes lead to precisely the $\mathbb{Z}_2$ gauge theory of Ref. 1. Together, these facts cement the equivalence between the fermionization (2) and Kitaev's exact solution.

The underlying reason for this stability is that the unprojected mean-field wave functions we obtain can be mapped via Eq. (10) onto unprojected wave functions in the Majorana fermionization of Ref. 1. Enforcing the SU(2) gauge constraints to reduce the model back to the physical Hilbert space amounts to two things: First, it eliminates the distinction between different possible mappings between $f_\sigma, f_\sigma^\dagger$ and $b_{x,y,z},c$. Second, it imposes a condition that is equivalent to the $\mathbb{Z}_2$ constraint required for the fermionization of Ref. 1. Thus, when expressed in the Majorana basis given by (10), the effect of this projection will be to apply the projector relevant to Kitaev's Majorana fermionization. In this way, both fermionizations lead to the same wave functions after projection.

A. Symmetries and robustness of the mean-field solution

First, we will show that, for the solvable Hamiltonian (1), the model's unusually large number of symmetries protect the exact fermionic band structure. The mean-field solution is thus exact in that it correctly describes all correlators of the physical spin degrees of freedom, in spite of the apparent violence done to the wave function by Gutzwiller projection.

We begin by listing the symmetries that are relevant to this discussion. The Hamiltonian (1) has the following discrete symmetries:

$$\hat{\mathbf{C}}\ :\ S_i^{x,y,z} \to s_{x,y,z}S_i^{x,y,z}, \tag{39}$$

where the sign $s_{x,y,z} = \pm 1$ can be chosen independently for x, y, and z spin operators. In the fermionic description, this leads to two discrete symmetries preserved by the mean-field Hamiltonian:

$$\begin{aligned}\hat{\mathbf{C}} &: f_{qu\sigma} \to f_{-qu\sigma}^\dagger,\\ \hat{\mathbf{S}} &: f_{q1\sigma} \to f_{-q1\sigma},\, f_{q2\sigma} \to -f_{-q2\sigma}.\end{aligned} \tag{40}$$

Here, the charge-conjugation symmetry $\hat{\mathbf{C}}$ is unitary [it is simply $\psi_{qu\sigma a} \to (\tau^x)_{ab}\,\psi_{qu\sigma b}$], while the sublattice symmetry $\hat{\mathbf{S}}$ is an antiunitary symmetry. Thus, in the mean-field Hamiltonian, $\hat{\mathbf{C}}$ takes $\Delta_{ij},t_{ij} \to \Delta_{ij},t_{ij}$ while $\hat{\mathbf{S}}$ takes $\Delta_{ij},t_{ij} \to \Delta_{ij}^*,t_{ij}^*$. Quadratic Hamiltonians invariant under $\hat{\mathbf{C}}$ have eigenstates that are diagonal in the Majorana basis (10). These symmetries impose an important restriction on t_{ij} and Δ_{ij}. $\hat{\mathbf{C}}$ is preserved as long as t_{ij},Δ_{ij} are purely imaginary. $\hat{\mathbf{S}}$ is preserved so long as there are no terms directly coupling fermions on the same sublattice.

Time-reversal symmetry is also respected by the model and its mean-field solution

$$\hat{\mathbf{T}}: f_{qu\uparrow} \to f_{-qu\downarrow},\; f_{qu\downarrow} \to -f_{-qu\uparrow}. \tag{41}$$

Single-spin terms (i.e., a magnetic field) and three-spin interactions break this symmetry. However, not all $\hat{\mathbf{T}}$-breaking perturbations will open a gap in the B phase: only those perturbations that break $\hat{\mathbf{S}}$ will open a gap in the spectrum, as we will see below. For example, the magnetic field discussed in Sec. III B breaks $\hat{\mathbf{T}}$ and $\hat{\mathbf{C}}$, but not $\hat{\mathbf{S}}$. As shown explicitly above, this does not gap the B phase and indeed results in power-law spin-spin correlations.

The relation between these symmetries is

$$\hat{\mathbf{T}} = \hat{\mathbf{S}}\hat{\mathbf{G}}^x\hat{\mathbf{C}}, \tag{42}$$

where the symmetry $\hat{\mathbf{G}}^x$ is given by

$$\begin{aligned}\hat{G}^{x,y}:\quad & f_{i\uparrow} \to f^\dagger_{i\downarrow}\\ & t_{ij,\uparrow},\Delta_{ij,\uparrow} \to t_{ij,\downarrow},\Delta_{ij,\downarrow}\\ & t_{ij,\downarrow},\Delta_{ij,\downarrow} \to t_{ij,\uparrow},\Delta_{ij,\uparrow},\end{aligned} \tag{43}$$

which are a discrete subset of the off-diagonal SU(2) rotations interchanging up and down spins. In the mean-field solution, these are no longer local symmetries. However, they remain global symmetries of the theory, the effect of which is to rotate between different possible mappings between the four Majorana fermions $(c, b^{x,y,z})$ and the four self-adjoint combinations $f^\dagger_{i\sigma} + f_{i\sigma}, i(f^\dagger_{i\sigma} - f_{i\sigma})$ of the spinful fermions. Thus, $\hat{\mathbf{T}}$ is a *projective* symmetry, i.e., a symmetry that maps the system to a different but gauge-equivalent saddle point. Such projective symmetries are important to classifying the phases of spin-liquid systems.[8]

Besides these more generic discrete symmetries [Eq. (1)] represents a somewhat special point in a more extended space of similar spin Hamiltonians: there is a product of spin operators on each plaquette that commutes with H. This is

$$\mathcal{P} = \prod_{i=1}^{6} S^{e(i)}(i) = \pm\frac{1}{2^6}, \tag{44}$$

where $e(i) = z$ for a vertex that sits between x and y links on the plaquette, y for a vertex that sits between x and z links on a plaquette, and x for a vertex that sits between y and z links on a plaquette (see Fig. 3). In the ground state, the value of this operator is positive on each plaquette.[1]

In terms of the fermionic operators, $\mathcal{P}$ can be written as

$$\mathcal{P}_f \equiv P_0\left(\prod_{i=1}^{6} b^\alpha_i b^\alpha_{i+1}\right)P_0, \tag{45}$$

where P_0 denotes Gutzwiller projection onto singly occupied states, $\alpha = x,y,z$ on x,y, and z links, respectively, and b^α_i are the Majorana fermions defined in Eq. (10). [Since the quantity in parentheses is not SU(2) gauge invariant, the projection operator is necessary in this case.] In the mean-field state, each species of Majorana fermion is localized on the appropriate links, with $\langle b^\alpha_i b^\alpha_{i+1}\rangle_{\rm MF} = 1/2$. Terms annihilated by P_0 do not contribute since $\langle f^\dagger_{i\uparrow} f^\dagger_{i\downarrow}\rangle_{\rm MF} = \langle f_{i\uparrow} f_{i\downarrow}\rangle_{\rm MF} = 0$. Hence, we find that the mean-field value

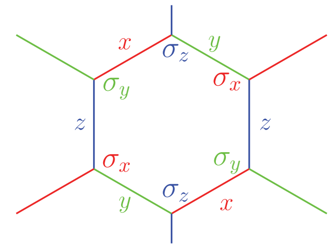

FIG. 3. (Color online) The product of spin operators conserved separately on each plaquette by the Kitaev Hamiltonian (1). The Hamiltonian (1) distinguishes between three types of links on the honeycomb lattice, which we call x, y, and z links (color-coded red, green, and blue, respectively, here). On x links the spin-spin interaction term is $S^{(x)}_i S^{(x)}_j$, and similarly for y and z links. The product of spin operators shown here—a product around a plaquette of the spin variable associated with the external edge at each vertex—commutes with the spin Hamiltonian.

$$\mathcal{P}_f = \langle b^x_1 b^x_2\rangle\langle b^y_2 b^y_3\rangle\langle b^z_3 b^z_4\rangle\langle b^x_4 b^x_5\rangle\langle b^y_5 b^y_6\rangle\langle b^z_6 b^z_1\rangle = \frac{1}{2^6} \tag{46}$$

is precisely that of the exact solution.

We now show that, combined with the discrete symmetries mentioned above, conservation of $\mathcal{P}_f$ prevents fluctuations about mean field from altering the fermionic band structure in any way. We will first establish that the symmetries forbid any terms other than those in Eq. (46) from contributing to $\mathcal{P}_f$. If there can be no further contributions to $\mathcal{P}_f$ induced by fluctuations, however, then also no spectral weight can be transferred from the equal-time correlation functions of the b^α, as otherwise we would not arrive at the correct value for $\mathcal{P}$. This means that all further-neighbor correlators must vanish exactly.

By Wick's theorem, we need only consider the possibility of other pairings of the fermionic operators that give a nonzero contribution to $\mathcal{P}_f$. The only possibility allowed by $\hat{\mathbf{C}}$ and $\hat{\mathbf{P}}\hat{\mathbf{S}}$ is to give a nonvanishing expectation value to terms of the form $\langle b^x_1 b^x_4\rangle, \langle b^y_2 b^y_5\rangle$, etc. Thus, we consider

$$\langle b^x_1 b^x_4\rangle\langle b^y_2 b^y_5\rangle\langle b^z_3 b^z_6\rangle\langle b^x_2 b^x_5\rangle\langle b^y_3 b^y_6\rangle\langle b^z_4 b^z_1\rangle. \tag{47}$$

However, the interacting Hamiltonian for the up spins decouples exactly into separate Hamiltonians for each chain of x-y links in the lattice. In particular, the full Hamiltonian contains no interaction term coupling b^x_1 and b^x_4, as they lie on different chains. Hence interactions can not shift $\langle b^x_1 b^x_4\rangle$ from its mean-field value of 0. As (47) is the only extra contribution to $\mathcal{P}_f$ not explicitly forbidden by symmetry, we conclude that Eq. (46) must remain valid in the full solution, and that consequently no fermion bilinears can be shifted from their mean-field values.

B. Gauge theory of fluctuations about mean field

Thus far, we have shown how to reproduce Kitaev's mean-field portrait of the exact spin-liquid ground state using the fermionization (2), and argued that including fluctuations about mean field will not change the fermionic band structure. Hence, we have obtained an alternative mean-field description of the ground state of (1), which reproduces faithfully the spin correlators of the exact ground state.

Although the mean-field solutions describe identical physics, however, the fermionization (2) differs quite dramatically from that of Ref. 1 in the nature of the bosonic variables, and consequently the theory of fluctuations about mean field. After Hubbard-Stratonovich transforming the four-fermion interactions, we obtain bosonic fields which condense to give both the hopping and superconducting order parameters as well as the SU(2) gauge fields associated with the constraint (3). One might therefore wonder why these do not lead to significantly different physical theories after fluctuations about mean field have been accounted for. Here we address this question, allowing us to posit that (6) describes a gapped spin-liquid phase that exists even away from the exactly solvable limit of the Hamiltonian (1).

The bosonic fluctuations about mean field can be separated into the following degrees of freedom. There are three scalar fields describing fluctuations in the amplitudes of the various kinetic and superconducting terms. All of these are massive, and as we shall see, two of them can be interpreted as Higgs fields for the broken SU(2) symmetry. In addition, there are three independent fields associated with phase fluctuations of the various link variables. These can be identified as an SU(2) gauge field (describing phase fluctuations of the spin-symmetric hopping term) and two Goldstone bosons associated with the phases of the order parameters breaking the SU(2) symmetry. We will briefly discuss each type in turn; a more detailed analysis is presented in Appendix B 3.

We begin with the scalar fields describing fluctuations in the amplitude of the various bosonic order parameters that fix the mean-field fermionic band structure. The general form of the Hubbard-Stratonovich action ensures that all of the scalar fields are massive, with energy gaps of order $\frac{1}{J}$ at the isotropic coupling point. Because of this mass gap, fluctuations in the amplitudes of the mean-field parameters are not generally expected to have an important effect on the fermions. The notable exception to this[23] is in cases when they destabilize the spin-liquid saddle point in favor of a "dimerized" state with spins hopping predominantly along a subset of links in the lattice. As we discuss in Sec. III A, an analog of the dimerized phase does occur for anisotropic $J_{x,y,z}$; in general, we may therefore conjecture that, away from the solvable point, this phase boundary may be shifted, but that fluctuations of the mean-field hopping and superconducting amplitudes will not qualitatively alter the phase diagram.

Next, we consider the impact of phase fluctuations described by the SU(2) gauge theory. Naively, the gauge theory is strongly fluctuating since there is no small parameter in the problem. However, the ground state of (1) is a Higgs phase, so the gauge field is massive. (Importantly, this explains why the gauge theory is not confined.)

To see that the model (1) is in a Higgs phase, we view the mean-field solution (6) as a condensate of two independent order parameters in the adjoint representation of SU(2). As explained in detail in Appendix B 3, the combination of superconducting and spin-antisymmetric hopping terms break the SU(2) gauge symmetry. This leaves only the residual Z_2 gauge-symmetry group one normally finds in a superconductor:

$$f_{i\sigma},\ f_{i\sigma}^{\dagger} \rightarrow -f_{i\sigma},\, -f_{i\sigma}^{\dagger}, \quad t_{ij,\sigma},\Delta_{ij,\sigma} \rightarrow t_{ij,\sigma},\Delta_{ij,\sigma} \tag{48}$$

comprising the residual Z_2 symmetry of the U(1) subgroup broken by superconductivity. As a result of the Anderson-Higgs phenomenon, the dynamical fluctuations in the gauge field are suppressed at long wavelengths, so that gauge-field fluctuations are not expected to substantially alter the fermionic band structure. (Here the gauge field results from the constraints of the purely two-dimensional system, and consequently is fully gapped unlike the electromagnetic gauge field in thin-film superconductors.) However, the gauge field makes itself felt in the interesting topological structure of the spin-liquid phase.

An alternative route for a gauge field to acquire a mass is through the generation of a Chern-Simons term. We will return to this possibility when we consider perturbations breaking $\hat{\mathbf{T}}$ in Sec. V, where we shall see that it plays an important role in the topological nature of the theory.

In summary, we can understand the exact ground state of Eq. (1), i.e., a phase whose propagating degrees of freedom consist of Majorana fermions coupled to a Z_2 gauge field, as a rather special incarnation of the Z_2 spin liquid: a spin-polarized p-wave superconductor. In this description, we arrive at Majorana fermions not by expressing the spins directly in a Majorana basis, but rather by starting with Dirac fermions coupled to an SU(2) gauge field and choosing a mean-field solution, which breaks the gauge symmetry. The Z_2 flux is thus the superconducting vortex, while the Z_2 charge carried by the Majorana fermions reflects the fact that the superconducting state conserves charge modulo 2.

V. $\hat{\mathbf{T}}$-BREAKING PERTURBATIONS: THE GAPPED B* PHASE

In Sec. III C, we showed that one way to open a gap in the B phase, i.e., by merging the two nodes, can be understood as a transition between a nodal and nodeless superconductor. This drives the system into the A phase. There is, however, a second way to open a gap: we may add another pairing term to the effective Hamiltonian (15), which will fully gap the spectrum provided that the corresponding gap does not vanish at the Dirac points. Here we focus on this latter gapped phase, and discuss its topological properties.

As noted in Sec. IV A, this second gapped phase necessarily breaks one of the two discrete symmetries of the mean-field solution—and hence the physical time-reversal symmetry of the spin model—since we must include couplings between sites on the same sublattice. Here we will focus on the case of broken $\hat{\mathbf{S}}$, as this can be realized by adding a three-spin interaction that commutes with the Hamiltonian (1).

A. Mean-field theory with $\hat{\mathbf{T}}$ -breaking terms

In terms of the original spin degrees of freedom, the $\hat{\mathbf{T}}$ -breaking term we must add to enter the B* phase is

$$\frac{J'}{2}\left(\sum_{\vec{r}_{ik}=\hat{x}} S_i^x S_j^z S_k^y + \sum_{\vec{r}_{ik}=\hat{l}_1} S_i^z S_j^y S_k^x + \sum_{\vec{r}_{ik}=\hat{l}_2} S_i^z S_j^x S_k^y\right), \tag{49}$$

where $\hat{l}_{1,2}$ are the lattice vectors of the honeycomb lattice (see Fig. 4). It is easy to see that this commutes with the plaquette product of spins (44),[24] and hence preserves the Z_2 vorticity on each plaquette. Hence, it also commutes with the full Hamiltonian, although not individually with the spin bilinears on each edge.

Expressing the spins in terms of Dirac fermions yields a six-fermion interaction. Although we can not perform the analog of an exact Hubbard-Stratonovich transformation for the resulting action, which contains both four and six fermion terms, at small J' it is possible to evaluate its effect on the mean-field solution in a controlled way (see Appendix C). We find that (consistent with the treatment of Ref. 1) the effect of such a term is to induce second-neighbor hopping and superconducting terms, without altering the rest of the band structure (except for an overall rescaling of the bandwidth).

We therefore begin by studying the resulting mean-field Hamiltonian. The three-spin interaction introduces the following quadratic fermion terms for the down-spin band:

$$H^{(1)}_{\rm MF} = \frac{J'}{8}(-\sin q_x + \sin \vec{q}\cdot\hat{l}_1 - \sin\vec{q}\cdot\hat{l}_2)[-\Psi^\dagger_q \mu_z \Psi_q + \Psi^\dagger_q(\Psi^\dagger_{-q})^T + \text{H.c.}], \tag{50}$$

where Ψ was defined in Eq. (14). As shown in Appendix C, the perturbation (49) does not alter the mean-field Hamiltonian of the up spins, which therefore maintain their flat band structure and remain localized on x and y links. In addition, the new couplings do not disrupt the pair of flat spin-down bands. Thus, the basic structure of the initial mean-field solution is preserved, and the only effect of the interaction (49) at mean field is to alter the structure of the dispersing spin-down band.

The new effective mean-field Hamiltonian for the spin-down fermions therefore has the form

$$\begin{aligned} H_{\rm down} &= \Psi^\dagger_q\left[\epsilon^{(x)}_q\mu_x + \epsilon^{(y)}_q\mu_y + \epsilon^{(z)}_q\mu_z\right]\Psi_q \\ &\quad + \Psi^\dagger_q\left(\Delta^{(s)}_q\mu_y + \Delta^{(t)}_q\mu_x\right)(\Psi^\dagger_{-q})^T + \text{H.c.} \\ &\quad + \tilde{\Delta}^{(p)}_q\Psi^\dagger_q(\Psi^\dagger_{-q})^T + \text{H.c.} + \frac{J_z}{8}\left(2 - \frac{J}{J_z}\right)\Psi^\dagger_q\mu_y\Psi_q \end{aligned} \tag{51}$$

with $\epsilon^{(x,y)}_q, \Delta^{(s,p)}_q$ given in Eqs. (16) and (17), and

$$\epsilon_z = \tilde{\Delta}^{(p)} = \frac{J'}{8}\left(-\sin q_x + 2\sin\frac{q_x}{2}\cos\frac{\sqrt{3}q_y}{2}\right). \tag{52}$$

In the vicinity of the Dirac cone, for $J_{x,y,z} \equiv J$, this gives

$$\begin{aligned} H_{\rm down} &= -\Psi^\dagger_q\left[\frac{\sqrt{3}}{32}Jq_y\mu_x - \frac{\sqrt{3}}{32}Jq_x\mu_y + \frac{J}{16}\mu_y \right.\\ &\quad \left. + \left(\frac{3\sqrt{3}}{64}J'q^2 - \frac{3\sqrt{3}}{16}J'\right)\mu_z\right]\Psi_q \\ &\quad + \frac{J\sqrt{3}}{32}\Psi^\dagger_p[p_y\mu_x - p_x\mu_y](\Psi^\dagger_{-p})^T + \text{H.c.} \\ &\quad - \frac{J}{16}\Psi^\dagger_p\mu_y(\Psi^\dagger_{-p})^T + \text{H.c.} \\ &\quad + (3/8\sqrt{3}J'^2q^2 + 3\sqrt{3}J'/2)\Psi^\dagger_q(\Psi^\dagger_{-q})^T + \text{H.c.}, \end{aligned} \tag{53}$$

which we can view as a mixed s-wave and chiral p-wave superconductor. This term opens a gap at the Dirac cone, so that the system is now fully gapped. We discuss the consequences in the next section.

B. Topological features of the gapped B* phase

Thus far, we have established that adding the spin interaction (49) has the effect, at mean field, of breaking $\hat{\mathbf{S}}$ and opening a gap in the spectrum of the dispersing Majorana mode (c), while leaving the band structure of the localized Majorana modes ($b^{x,y,z}$) unchanged. We will now see how this perturbation leads to a topological phase with zero-energy Majorana fermions bound to vortices, exactly as in the spinless $p + ip$ superconductor of Read and Green.[25]

The simplest way to identify the nature of the B* phase is to consider the Hamiltonian (20), where the B phase is a p_y superconductor. The perturbation modifies the Hamiltonian according to

$$\begin{aligned} \Delta_q &\to \Delta_q - i\frac{J'}{4}\sin\frac{q_x}{2}\left(\cos\frac{\sqrt{3}}{2}q_y - \cos\frac{q_x}{2}\right) \\ &\approx -i\,\text{sgn}(q_x)\frac{J'}{4}\left(1 + \frac{J_z}{2J}\right)\sqrt{1 - \left(\frac{J_z}{2J}\right)^2}. \end{aligned} \tag{54}$$

In the second line, we have approximated Δ_q by its value in the vicinity of the nodes. From this expression, we see that this is an ip_x superconducting gap, which opens up a gap at the nodes.

As noted previously, in the nodal B phase, the "chemical potential" $\mu = 2J - J_z$ lies in the band. Thus, when the gap is opened, the system goes into the "weak-coupling" $p + ip$ superconducting phase. As $\vec{q}$ ranges over the Brillouin zone, the vector $(\text{Re}\Delta_q, \text{Im}\Delta_q, \epsilon_q)/(\epsilon^2_q + |\Delta_q|^2)^{1/2}$ wraps around the sphere. The corresponding winding number can not be changed without closing the gap, i.e., without going through a phase transition.

Conversely, when the three-spin interaction is included in the A phase, the chemical potential lies below the band. For sufficiently small J', $(\text{Re}\Delta_q, \text{Im}\Delta_q, \epsilon_q)/(\epsilon^2_q + |\Delta_q|^2)^{1/2}$ remains in the northern hemisphere, and thus has winding number zero. Thus, this is the strong-pairing phase of the chiral p-wave superconductor. In other words, including a weak $\hat{\mathbf{S}}$-breaking perturbation in the A phase leaves the system in the A phase.

Once we have identified the B* phase with the weak-pairing phase of the chiral p-wave superconductor, we are faced with the following riddle: in its usual incarnation, the superconducting coherence length is assumed to be much larger than the lattice scale, so that vortices are well modeled by a continuum theory. In particular, the vortex will have a core that is in the normal state. The argument put forth by Read and Green[25] to show that, in the weak-pairing phase, a zero-energy Majorana fermion is bound to the vortex core, relies on the existence of a domain wall between the vortex core and the superconductor in an essential way. Since phase B* is known to have the same topological order as the chiral p-wave superconductor, in which the existence of Ising anyons is due to the fact that these zero-energy Majorana fermions are bound to the vortex cores, we expect a similar phenomenon. In the lattice model at hand, however, a vortex exists on a single plaquette, and there is no vortex core. How, then, do the Majorana fermions become bound to these vortices?

One answer to this question comes from studying the long-wavelength gauge theory. First, we observe that the key effect of the $\hat{\mathbf{T}}$-breaking three-spin interaction is that it induces a mass term $m_{(\frac{4\pi}{3},0)} = -m_{(\frac{-4\pi}{3},0)} = \frac{3\sqrt{3}}{2}J'$ at the two nodes in the Brillouin zone. As discussed previously, the low-energy effective theory is that of a single species of massive Dirac fermion. If we integrate it out, then as shown explicitly in Appendix D, the one-loop effective action for the gauge fields is precisely what we would expect from a single Dirac cone, except that, since U(1) is broken down to Z_2, a Higgs mass is also generated:

$$\mathcal{L}_g^{(\text{one loop})} = \frac{1}{2}|\Phi|^2 A_\mu A^\mu - \frac{1}{4\pi m}F^{\mu\nu}F_{\mu\nu} + \frac{m}{|m|}\frac{1}{8\pi}\epsilon_{\mu\nu\lambda}A^\mu\partial^\nu A^\lambda. \tag{55}$$

In other words, we obtain the usual Higgs mass term, the field-strength tensor squared, and a Chern-Simons term with level $\frac{1}{2}$ (as usual from a single Dirac cone[12]). The Higgs mass is proportional to the condensate fraction $|\Phi|^2$, and is crucial outside a vortex. However, in a vortex core, the condensate vanishes. We will assume that the Higgs mass can be neglected in the core. Thus, in a vortex core, we have

$$\frac{\delta\mathcal{L}}{\delta A^\mu} = \frac{m}{|m|}\frac{1}{8\pi}\epsilon_{\mu\nu\lambda}\partial^\nu A^\lambda + J_\mu, \tag{56}$$

where J_μ is the fermion current, and we have used $\partial^\nu F_{\mu\nu} = 0$. Taking $\mu = 0, m > 0$, we obtain the constraint

$$\frac{1}{4\pi}\mathbf{B}^z_{\vec{R}} = \rho_{\vec{R}}, \tag{57}$$

where $\rho \equiv J_0$. In the case at hand, we have

$$\rho_q = \sum_{u=1,2}\sum_k [f^\dagger_{k,u\downarrow}f_{k-q,u\downarrow} + f_{-k,u\downarrow}f^\dagger_{-k+q,u\downarrow}]. \tag{58}$$

(Here, k is technically restricted to momenta near the Dirac cone; more generally, we sum over only half the Brillouin zone.) The rather counterintuitive fact that *holes* at the left Dirac cone carry the same charge as particles at the right Dirac cone results from the fact that the two cones have opposite chirality.

The density ρ_q is a sum of the density of particles at the right Dirac cone and holes at the left Dirac cone. The creation operator associated with this density is the Majorana fermion $c_{qu} = i(f^\dagger_{qu\downarrow} - f_{-qu\downarrow})$, which simultaneously creates a particle at q and a hole at $-q$. Hence, Eq. (57) tells us that there is a Majorana fermion c bound to every half-flux quantum. These half-flux quanta are precisely the Z_2 vortices of the superconductor; hence, we conclude that there is a Majorana fermion c bound to each Z_2 vortex.

VI. SPIN-DENSITY-WAVE STATES

As described in the preceding section, a $\hat{\mathbf{T}}$-breaking three-spin term opens up a gap. It is instructive to express the Hamiltonian in terms of the $\tilde{\chi}$ fermions:

$$\tilde{\chi}_p = \begin{pmatrix} \eta_{\vec{Q}/2+\vec{p}} \\ \eta^\dagger_{-\vec{Q}/2-\vec{p}} \end{pmatrix}. \tag{59}$$

At the isotropic point, $J_x = J_y = J_z$, $\vec{Q}/2 = (\frac{4\pi}{3},0)$. The Hamiltonian in the B phase can be written in the form

$$H = \sum_{\vec{p}} \tilde{\chi}^\dagger_p[vp_y\tau_y + vp_x\tau_z]\tilde{\chi}_p, \tag{60}$$

where $v = \frac{\sqrt{3}J}{2}$ at the isotropic point. The Dirac mass term generated by the three-spin interaction is of the form

$$H_{\text{D.M.}} = m\sum_{\vec{p}}\tilde{\chi}^\dagger_p\tau_y\tilde{\chi}_p, \tag{61}$$

where $m = 3J'/2$.

However, this is not the only possible term that can open a gap at the nodes of the B phase. The other possible term is (W is a coupling, which we introduce to parametrize the strength of this term)

$$\begin{aligned} H_{\text{pair}} &= W\sum_{\vec{p}}\tilde{\chi}^T_p i\tau_y\tilde{\chi}_{-p} + \text{H.c.} \\ &= 2W\sum_{\vec{p}}\eta^\dagger_{-\vec{Q}/2+\vec{p}}\eta_{\vec{Q}/2+\vec{p}} + \text{H.c.} \\ &= 4W\sum_{\vec{p}}[c_{\vec{Q}/2-\vec{p},1}c_{\vec{Q}/2+\vec{p},1} + (1\to 2) \\ &\quad + ic_{\vec{Q}/2-\vec{p},1}c_{\vec{Q}/2+\vec{p},2} + (1\leftrightarrow 2)] + \text{H.c.} \\ &= -4W\sum_{\vec{p}}[f^\dagger_{\vec{Q}/2-\vec{p},1}f^\dagger_{\vec{Q}/2+\vec{p},1} - f^\dagger_{\vec{Q}/2-\vec{p},1}f_{-\vec{Q}/2-\vec{p},1} \\ &\quad + \ldots] + \text{H.c.} \end{aligned} \tag{62}$$

Thus, such a mass term breaks translational symmetry. It includes terms that induce superconductivity at nonzero wave vector as well as terms that induce a spin-density wave at wave vector $\vec{Q}$. We can imagine that a spin-spin interaction which is added to the Kitaev model as a perturbation will, upon decoupling, generate such a mass term. However, since the density of states at the nodes is zero, interactions will only generate such a term at $O(1)$ coupling strength (not at infinitesimal coupling, as would be the case for a Fermi surface instability). At $O(1)$ coupling strength, there is no reason to focus on the the nodal regions, so many other instabilities could also occur. It is possible that, in a large-N version of this model, such a translational-symmetry-breaking instability will occur at weak coupling.

Similar but distinct spin-density-wave states have recently been discussed in the context of a hybrid Kitaev-Heisenberg model in Refs. 26 and 27.

VII. DISCUSSION

In describing the spin-liquid ground states of the various phases of Kitaev's honeycomb model using the slave fermion

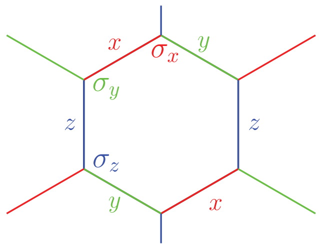

FIG. 4. (Color online) The three-spin interaction that breaks $\hat{\mathbf{T}}$ but commutes with the conserved six-spin product on each plaquette (see Fig. 3) (Ref. 24). The operator is constructed by taking the product of spin operators on three adjacent vertices, where the direction of the central spin is that associated with the external edge at the vertex, while the two external spins match the edges joining their associated vertices to the central vertex. In the figure, the central vertex is at the upper left, and the operator is the product of σ_z at the lower left vertex, σ_y at the upper left vertex, and σ_x at the top vertex.

approach, we may learn several things about the nature of the phases of this model, their potential stability to perturbations away from the solvable point, and their precise relationship to other phases of matter that exhibit similar physics.

First, the fermionic mean-field theory allows us to relate the various phases of the Kitaev model to the ground states of different Bogoliubov–de Gennes Hamiltonians. This can be done in two different ways: (i) in terms of the fermions $f_{\uparrow\downarrow}$ introduced in Eq. (2) and (ii) in terms of the fermions η introduced in Eq. (19). The latter are formed from the propagating part of $f_\downarrow$. Each way has its conceptual and technical advantages, as we have seen.

The mean-field phase diagram is summarized in Fig. 5, which can be interpreted in terms of the η fermions as follows. The A phase, in which the nodes of the superconductor do not intersect the Fermi surface, is adiabatically connected to an s-wave superconductor. The B phase is a nodal p-wave superconductor. The B* phase is the weak-pairing phase of a chiral p-wave superconductor, with the consequent Ising topological order. The A* phase is the corresponding strong-pairing chiral p-wave superconductor phase. As a result of the strong-pairing nature of this phase, the topological order is, in fact, again that of an s-wave superconductor. The reason for this is that, at the mean-field level (i.e., when treated as a free fermion problem), the A and A* phases can be adiabatically deformed into each other, so the line between them in Fig. 5 is a crossover line. On the other hand, the other transitions in Fig. 5 are genuine phase boundaries, which are essentially the same as the corresponding transitions in the superconductor. One important difference needs to be emphasized. In a two-dimensional superconductor with a three-dimensional electromagnetic field, there is a gapless plasmon. Thus, a thin superconducting film is not fully gapped, even though its fermionic spectrum is fully gapped. However, in the Kitaev honeycomb lattice model, the gauge field is two dimensional. Consequently, the plasmon is gapped and the system is fully gapped.

Although the SU(2) mean-field theory described here is clearly more complicated than that of Kitaev at the soluble point, it has the salient virtue that it is well suited to perturbing

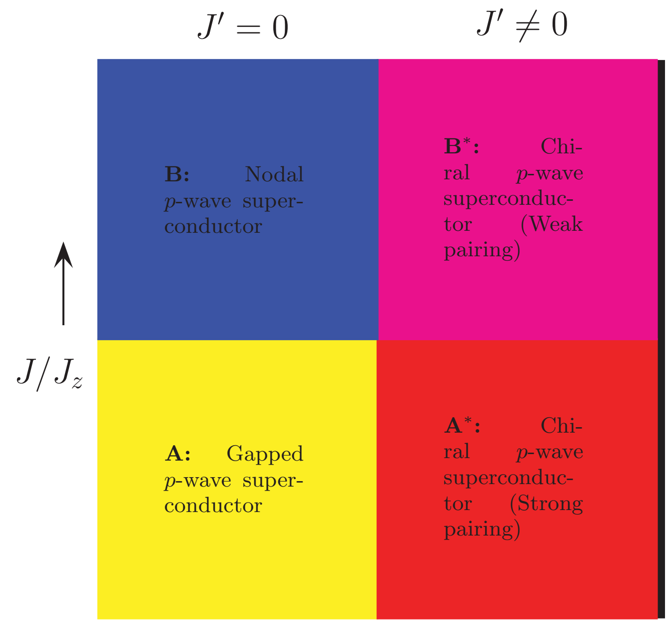

FIG. 5. (Color online) Schematic phase diagram of the Kitaev honeycomb model, and the corresponding superconducting phases. The phase is determined by the ratio J/J_z, and by whether the coefficient J' of the three-spin interaction is non-vanishing.

away from the soluble point, particularly in the gapless B phase.

It is interesting to consider the fate of the phase diagram shown in Fig. 5 when the spin Hamiltonian is deformed away from the exactly solvable point. The fully gapped phases, i.e., the A and B* phases, will be robust against small perturbations by virtue of their energy gaps. As long as the gauge symmetry is broken to Z_2 by the saddle-point solution, the gauge field is gapped, and we do not expect fluctuations to lead to confinement. Therefore, the model still admits effective spinon excitations, and the topological order of the spin liquid will be robust to gauge field fluctuations. Since the spinons are also gapped in these phases, they are stable against adding weak interactions between the fermions. The gapless B phase is a little trickier. Since the gauge field is fully gapped, we believe that the gauge-field action is robust against small perturbations. The fermions, on the other hand, are gapless. However, since they have a single gapless Dirac point (rather than a Fermi surface), weak interactions between the fermions are irrelevant by power counting. This is the reason that an SU(2)-invariant Heisenberg perturbation does not lead to a phase transition until the perturbation is sufficiently strong. Thus, we could say that the stability of the gapless B phase relies on phase-space limitations. However, as we have seen, although the gapless B phase is stable against weak perturbations, some features of the soluble point are not generic to this phase. For instance, a magnetic field will make the spin-spin correlation functions have a power law rather than short-ranged form.

The fact that the bosonic fluctuations are all gapped does not, however, prevent the theory from acquiring a new lowest-energy saddle point if we deform far enough away from the solvable model. For instance, as we have discussed, the gapless B phase can acquire a gap by an alternative method: the development of a spin-density wave, as discussed in

Sec. VI. Various perturbations of the Kitaev model, including a Heisenberg interaction,[26,27] can lead to such an instability. Furthermore, it is well known[7,23] that symmetric spin-liquid states are often prone to dimerization instabilities in which the spins pair with neighbors in a valence-bond crystal, which breaks a lattice symmetry. Away from the solvable limit, therefore, it is likely that the phase diagram will also include some such valence-bond-crystal states. At the symmetric point, the model has a spin-orbit-type three-fold rotation symmetry [entailing a three-fold lattice rotation about a vertex, coupled with a global spin rotation of the form (43)] which makes the saddle point perturbatively stable, although in principle lower-energy symmetry-breaking saddle points might exist. Away from the isotropic point $J_x = J_y = J_z$, such states need not break any symmetries of the Hamiltonian, so that symmetry does not prevent the saddle point from flowing to such a valence-bond crystal upon including fluctuations of the amplitudes of the mean-field hopping and superconducting terms.

The fact that the exact ground state of (1) can be correctly described in the slave fermion mean-field approach used here is also interesting in its own right. As discussed above, since the mean-field state is a Higgs phase of the gauge field, the model is in a regime where the spin-liquid saddle point is most likely to be stable. Even in this case, however, examples of Hamiltonians where the exact ground state can be shown to be a spin liquid are rare. The Kitaev model is thus a potential testing ground for the slave fermion approach since we may begin with a Hamiltonian for which it is demonstrably valid, and consider the fate of the ground state under various perturbations. In particular, on general grounds[28] we expect that, for small perturbations that do not close the gap in the spectrum, the slave fermion mean-field theory will continue to capture the topological order of the gapped phases.

Another interesting prediction of the slave fermion approach is that, near the solvable point, the Kitaev model becomes a superconductor upon doping. Specifically, we imagine starting with a Mott insulator, the effective Hamiltonian at half-filling of which is given by (1). After doping away from half-filling, we must account for the fermion hopping terms, leading to a $t - J$ model, with the spin Hamiltonian given by (1). Following the prescription used to study the cuprates,[29] we may decompose the spin operators as in Eq. (2), and express the electron operator as

$$c^{\dagger}_{i\sigma} = f^{\dagger}_{i\sigma} b_{i\sigma} \tag{63}$$

with the constraint

$$f^{\dagger}_{i\uparrow} f_{i\uparrow} + f^{\dagger}_{i\downarrow} f_{i\downarrow} + b^{\dagger}_i b_i = 1. \tag{64}$$

It follows that, at temperatures below the Bose condensation temperature of the bosons, and at sufficiently low dopings, the mean-field solution described above is a good approximation for the spinons ($f_{i\sigma}$), and the superconducting order parameter is

$$\begin{aligned} \Delta^{\text{phys}}_{k;\sigma,\sigma'} &= \langle c^{\dagger}_{k\sigma} c^{\dagger}_{-k\sigma'} \rangle = \langle f^{\dagger}_{k+q,\sigma} f^{\dagger}_{-k-q,\sigma'} \rangle \langle b_{-q\sigma} b_{q\sigma'} \rangle \\ &= \Delta_{k;\sigma,\sigma'} \rho_s, \end{aligned} \tag{65}$$

where ρ_s is the bosonic superfluid density. Thus, the momentum dependence of the physical superconducting order parameter is set by that of the mean-field superconducting order parameter Δ for the fermionic spinons f. For the Hamiltonian (1), this predicts spin-triplet superconductivity (with equal spin pairing), with a mixed singlet and triplet pseudospin order parameter.

Finally, it is interesting to compare the mean-field ground state of the Kitaev model with existing proposals for generating the B* phase's topological Majorana fermions in physical materials. The mean-field Hamiltonian of the B phase is manifestly equivalent to a $p + ip$ superconducting state of spin-polarized fermions.[25] It also has an interesting relation to the the effective Hamiltonian of Fu and Kane[30] for surface states of a topological insulator in the presence of induced s-wave superconductivity. In the absence of superconductivity, these surface states form a single Dirac fermion. This Dirac fermion is analogous to the Dirac fermion, which we have in the gapless B phase. If a magnetic film is brought into contact with the topological insulator, and the magnetic moment is perpendicular to the interface, then the resulting term in the Hamiltonian is a Dirac mass term, which breaks time-reversal symmetry and opens a gap. This is analogous to the three-spin term in the Kitaev model, which opens a gap and drives the system into the B* phase. Note that this term in the Kitaev model is *not* analogous to the term generated by an s-wave superconducting film on the surface of a topological insulator. Instead, s-wave superconductivity on the surface of a topological insulator is analogous to a term $\tilde{\chi}^T_p i\tau_y \tilde{\chi}_p + \text{H.c.}$, which is a down-spin-density wave at wave vector $(8\pi/3,0)$ at the symmetric point $J_x = J_y = J_z$.

In all cases, the essential ingredients for generating topological Majorana fermions are a two-band model in which the band structure is that of a massive Dirac fermion, and with induced superconductivity. As we described in Sec. V B above, the massive Dirac fermion in all of these models is implicitly coupled to a gauge field since it forms a superconducting state. The fermion mass therefore generates a Chern-Simons term in the effective gauge-field action, which has the effect of binding a half-quantum vortex to each charge since there is only a single Dirac cone. The charge that is bound in the superconducting state is a Bogoliubov–de Gennes quasiparticle, rather than a fermion, which, when the superconducting order parameter has a p-wave component, binds a Majorana fermion to the vortex.

ACKNOWLEDGMENTS

We gratefully acknowledge the hospitality of the Aspen Center for Physics, where part of this work was completed. We thank J. Chalker, J. Cardy, and S. Trebst for discussions. C.N. has been supported in part by the DARPA QuEST program.

APPENDIX A: MAPPING BETWEEN SU(2) AND MAJORANA FERMIONIZATIONS

Here, we explain in more detail the correspondence between the fermionization (2) and the Majorana fermionization employed by Kitaev.[1] We begin with the mean-field correspondence

$$\begin{aligned} b^x_{\mathbf{q}u} &= i(f^{\dagger}_{\mathbf{q}u\uparrow} - f_{-\mathbf{q}u\uparrow}), \quad b^y_{\mathbf{q}u} = f^{\dagger}_{\mathbf{q}u\uparrow} + f_{-\mathbf{q}u\uparrow}, \\ b^z_{\mathbf{q}u} &= f^{\dagger}_{\mathbf{q}u\downarrow} + f_{-\mathbf{q}u\downarrow}, \quad c_{\mathbf{q}u} = i(f^{\dagger}_{\mathbf{q}u\downarrow} - f_{-\mathbf{q}u\downarrow}), \end{aligned} \tag{A1}$$

which gives a mapping between unprojected spinful fermions and unprojected Majorana fermions. This mapping is not unique, as each Majorana fermion can be represented by any linear combination:

$$c_{\mathbf{q}u} = f^\dagger_{\mathbf{q}u\sigma} e^{i\phi} + \text{H.c.} \tag{A2}$$

and any choice of four such combinations, which mutually anticommute, could be associated with $\{b^x,b^y,b^z,c\}$. However, this difference is not physical as all such mappings are equivalent under SU(2) gauge transformations.

The mapping (A1) does not preserve the form of the unprojected spin operators, however. Specifically, the fermionization (2) gives

$$S^x_i = (f^\dagger_{i\uparrow} f_{i\downarrow} + f^\dagger_{i\downarrow} f_{i\uparrow}), \quad S^y_i = -i(f^\dagger_{i\uparrow} f_{i\downarrow} + f^\dagger_{i\downarrow} f_{i\uparrow}), \\ S^z_i = (f^\dagger_{i\uparrow} f_{i\uparrow} - f^\dagger_{i\downarrow} f_{i\downarrow}), \tag{A3}$$

while Kitaev's Majorana fermionization stipulates

$$\begin{aligned} \tilde{S}^x_i &= i b^x_i c_i = -i(f^\dagger_{i\uparrow} - f_{i\uparrow})(f^\dagger_{i\downarrow} - f_{i\downarrow}), \\ \tilde{S}^y_i &= i b^y_i c_i = -(f^\dagger_{i\uparrow} + f_{i\uparrow})(f^\dagger_{i\downarrow} - f_{i\downarrow}), \\ \tilde{S}^z_i &= i b^z_i c_i = -(f^\dagger_{i\downarrow} + f_{i\downarrow})(f^\dagger_{i\downarrow} - f_{i\downarrow}). \end{aligned} \tag{A4}$$

This gives

$$\tilde{S}^x_i = -S^y_i - i(f^\dagger_{i\uparrow} f^\dagger_{i\downarrow} + f_{i\uparrow} f_{i\downarrow}), \quad \tilde{S}^y_i = S^x_i - (f^\dagger_{i\uparrow} f^\dagger_{i\downarrow} - f_{i\uparrow} f_{i\downarrow}), \\ \tilde{S}^z_i = -S^z_i + (n_{i\uparrow} + n_{i\downarrow} - 1), \tag{A5}$$

which, after a gauge transformation to rotate the spins and eliminate the extra phases, differs from the spin operators (A4) by terms that vanish under projection onto the physical Hilbert space. It is these extra terms that lead to the fact that the mean-field Hamiltonian (6) does not conserve $b^x_i b^x_j$ on x links (and similarly for y and z) so that it is not obvious that the mean-field theory captures the essentials of the spin-spin correlations, as it is in the Majorana description.

However, one way to view the equivalence of the two descriptions is via the wave functions that they produce *after* projection. The Majorana projector is

$$\begin{aligned} D_i &\equiv b^x_i b^y_i b^z_i c_i = \hat{\mathbf{1}} \\ &= -(f^\dagger_{i\uparrow} + f_{i\uparrow})(f^\dagger_{i\uparrow} - f_{i\uparrow})(f^\dagger_{i\downarrow} + f_{i\downarrow})(f^\dagger_{i\downarrow} - f_{i\downarrow}). \end{aligned} \tag{A6}$$

By expanding the constraint in terms of Dirac fermion operators, we obtain

$$\begin{aligned} D_i &= -(2n_{i\uparrow} - 1)(2n_{i\downarrow} - 1) \\ &= -2(n_{i\uparrow} + n_{i\downarrow} - 1)^2 + 1. \end{aligned} \tag{A7}$$

Hence, imposing the diagonal SU(2) constraint

$$n_{i\uparrow} + n_{i\downarrow} - 1 = 0 \tag{A8}$$

automatically imposes the Majorana constraint $D_i = \hat{\mathbf{1}}$.

Therefore, if we begin with a mean-field wave function expressed in terms of the spinful fermions, and project onto the physical Hilbert space of singly occupied states, this is equivalent to studying the same mean-field wave function expressed in terms of Majorana fermions, and applying the projector (A6) at each site. This gives an alternative perspective on why the mean-field theory is exact.

APPENDIX B: MEAN-FIELD THEORY OF THE QUADRATIC SPIN MODEL

Here, we will review the detailed derivation of the mean-field Hamiltonian (7). We will first show how to derive the full effective action, and then present the self-consistent mean-field solution.

1. Hubbard-Stratonovich decoupling of the Kitaev model

In the Dirac fermion basis, the three different types of terms in the Hamiltonian (1) are

$$\begin{aligned} \hat{S}^x_i \hat{S}^x_j &= -\tfrac{1}{4}[f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} f_{i\downarrow} f_{j\downarrow} + f^\dagger_{i\downarrow} f^\dagger_{j\downarrow} f_{i\uparrow} f_{j\uparrow} + f^\dagger_{i\uparrow} f_{j\uparrow} f^\dagger_{j\downarrow} f_{i\downarrow} + f^\dagger_{i\downarrow} f_{j\downarrow} f^\dagger_{j\uparrow} f_{i\uparrow}], \\ \hat{S}^y_i \hat{S}^y_j &= -\tfrac{1}{4}[-f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} f_{i\downarrow} f_{j\downarrow} - f^\dagger_{i\downarrow} f^\dagger_{j\downarrow} f_{i\uparrow} f_{j\uparrow} + f^\dagger_{i\uparrow} f_{j\uparrow} f^\dagger_{j\downarrow} f_{i\downarrow} + f^\dagger_{i\downarrow} f_{j\downarrow} f^\dagger_{j\uparrow} f_{i\uparrow}], \\ \hat{S}^z_i \hat{S}^z_j &= -\tfrac{1}{4}[f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} f_{j\uparrow} f_{i\uparrow} + f^\dagger_{i\downarrow} f^\dagger_{j\downarrow} f_{j\downarrow} f_{i\downarrow} + f^\dagger_{i\uparrow} f_{j\uparrow} f^\dagger_{j\uparrow} f_{i\uparrow} + f^\dagger_{i\downarrow} f_{j\downarrow} f^\dagger_{j\downarrow} f_{i\downarrow}], \end{aligned} \tag{B1}$$

where we have used $n_{i\uparrow} = 1 - n_{i\downarrow}$ in the last expression.

To decouple the four-fermion interactions using Hubbard-Stratonovich fields, we take the Lagrangian

$$\begin{aligned} \mathcal{L}_x &= -\frac{8(|\Phi_1|^2 + |\Phi_2|^2)}{J_x} + \Phi_1(f^\dagger_{i\uparrow} f_{j\uparrow} + f^\dagger_{i\downarrow} f_{j\downarrow}) + i\Phi_2(f^\dagger_{i\uparrow} f_{j\uparrow} - f^\dagger_{i\downarrow} f_{j\downarrow}) + \tilde{\text{H.c.}} \\ &\quad - \frac{8(|\Theta_1|^2 + |\Theta_2|^2)}{J_x} + \Theta_1(f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} + f^\dagger_{i\downarrow} f^\dagger_{j\downarrow}) + i\Theta_2(f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} - f^\dagger_{i\downarrow} f^\dagger_{j\downarrow}) + \tilde{\text{H.c.}}, \\ \mathcal{L}_y &= -\frac{8(|\Phi_1|^2 + |\Phi_2|^2)}{J_y} + \Phi_1(f^\dagger_{i\uparrow} f_{j\uparrow} + f^\dagger_{i\downarrow} f_{j\downarrow}) + i\Phi_2(f^\dagger_{i\uparrow} f_{j\uparrow} - f^\dagger_{i\downarrow} f_{j\downarrow}) + \tilde{\text{H.c.}} \\ &\quad - \frac{8(|\Theta_1|^2 + |\Theta_2|^2)}{J_y} + i\Theta_1(f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} + f^\dagger_{i\downarrow} f^\dagger_{j\downarrow}) + \Theta_2(f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} - f^\dagger_{i\downarrow} f^\dagger_{j\downarrow}) - \tilde{\text{H.c.}}, \\ \mathcal{L}_z &= -\frac{4(|\Phi_1|^2 + |\Phi_2|^2)}{J_z} + \Phi_1 f^\dagger_{i\uparrow} f_{j\uparrow} + \Phi_2 f^\dagger_{i\downarrow} f_{j\downarrow} + \tilde{\text{H.c.}} - \frac{4(|\Theta_1|^2 + |\Theta_2|^2)}{J_z} + \Theta_1 f^\dagger_{i\uparrow} f^\dagger_{j\uparrow} + \Theta_2 f^\dagger_{i\downarrow} f^\dagger_{j\downarrow} + \tilde{\text{H.c.}}, \end{aligned} \tag{B2}$$

where the fields Φ_i,Θ_i are to be understood as being evaluated on the link in question, and the $\tilde{\text{H}}$.c. is the Hermitian conjugate with all spin directions reversed. We can check that this decoupling gives back the original action by integrating out the bosonic fields. For example, completing the square for the first line of $\mathcal{L}_x$ gives

$$\begin{aligned}\mathcal{L}_x = &-\frac{8}{J_x}\left[\Phi_1 - \frac{J_x}{8}(f^\dagger_{j\uparrow}f_{i\uparrow} + f^\dagger_{j\downarrow}f_{i\downarrow})\right]\left[\Phi^\dagger_1 - \frac{J_x}{8}(f^\dagger_{i\uparrow}f_{j\uparrow} + f^\dagger_{i\downarrow}f_{j\downarrow})\right] + \frac{J_x}{8}(f^\dagger_{i\uparrow}f_{j\uparrow} + f^\dagger_{i\downarrow}f_{j\downarrow})(f^\dagger_{j\uparrow}f_{i\uparrow} + f^\dagger_{j\downarrow}f_{i\downarrow})\\ &-\frac{8}{J_x}\left[\Phi_2 - i\frac{J_x}{8}(f^\dagger_{j\uparrow}f_{i\uparrow} - f^\dagger_{j\downarrow}f_{i\downarrow})\right]\left[\Phi^\dagger_2 - i\frac{J_x}{8}(f^\dagger_{i\uparrow}f_{j\uparrow} - f^\dagger_{i\downarrow}f_{j\downarrow})\right] - \frac{J_x}{8}(f^\dagger_{i\uparrow}f_{j\uparrow} - f^\dagger_{i\downarrow}f_{j\downarrow})(f^\dagger_{j\uparrow}f_{i\uparrow} - f^\dagger_{j\downarrow}f_{i\downarrow}).\end{aligned} \quad \text{(B3)}$$

Integrating out the factors involving Φ_1 and Φ_2 gives a constant; the sum of the remaining pieces gives

$$\frac{J_x}{4}(f^\dagger_{i\uparrow}f_{j\uparrow}f^\dagger_{j\downarrow}f_{i\downarrow} + f^\dagger_{i\downarrow}f_{j\downarrow}f^\dagger_{j\uparrow}f_{i\uparrow}) \quad \text{(B4)}$$

as expected.

Now we proceed in the usual way for mean-field theories, namely, the fields Θ and Φ have gapped amplitude fluctuations as well as phase fluctuations. We will thus begin with a mean-field solution $\Theta_{\sigma,ij}(t) \equiv \Delta_{\sigma,ij}, \Phi_{\sigma,ij}(t) \equiv t_{\sigma,ij}$, which reproduces the quadratic fermionic spectrum of the exact solution. We then consider the fate of the fluctuations of both gapped amplitude modes and gapless phase modes about mean field.

2. Mean-field solution

At mean-field level, the relevant information contained in Eq. (B2) is that on each link there are potentially four bosonic fields: $t_\uparrow$ associated with hopping of up spins, $t_\downarrow$ with hopping of down spins (which formally transforms in the opposite way under time reversal), and separate superconducting order parameters $\Delta_\uparrow,\Delta_\downarrow$ for the spin-up and spin-down sectors. Formally, in terms of the fields of the previous section, we take

$$\begin{gathered}t_\uparrow = \langle\Phi_1 + i\Phi_2\rangle \text{ on } x \text{ and } y \text{ links}, \quad t_\uparrow = \langle\Phi_1\rangle \text{ on } z \text{ links},\\ t_\downarrow = \langle\Phi_1 - i\Phi_2\rangle \text{ on } x \text{ and } y \text{ links}, \quad t_\downarrow = \langle\Phi_2\rangle \text{ on } z \text{ links},\\ \Delta_\uparrow = \langle i\Theta_1 + \Theta_2\rangle \text{ on } y \text{ links}, \quad \Delta_\downarrow = \langle i\Theta_1 - \Theta_2\rangle \text{ on } y \text{ links},\\ \Delta_\uparrow = \langle\Theta_1 + i\Theta_2\rangle \text{ on } x \text{ links}, \quad \Delta_\uparrow = \langle\Theta_1\rangle \text{ on } z \text{ links},\\ \Delta_\downarrow = \langle\Theta_1 - i\Theta_2\rangle \text{ on } x \text{ links}, \quad \Delta_\downarrow = \langle\Theta_2\rangle \text{ on } z \text{ links}.\end{gathered} \quad \text{(B5)}$$

From the Lagrangian (B2), the saddle-point equations are

$$\begin{aligned}t^{(x,y)}_\uparrow &= \frac{J_{x,y}}{4}\langle f^\dagger_{j\downarrow}f_{i\downarrow}\rangle, \quad t^{(x,y)}_\downarrow = \frac{J_{x,y,}}{4}\langle f^\dagger_{j\uparrow}f_{i\uparrow}\rangle,\\ t^{(z)}_\uparrow &= \frac{J_z}{4}\langle f^\dagger_{j\uparrow}f_{i\uparrow}\rangle, \quad t^{(z)}_\downarrow = \frac{J_{z,}}{4}\langle f^\dagger_{j\downarrow}f_{i\downarrow}\rangle,\\ \Delta^{(x)}_\uparrow &= \frac{J_x}{4}\langle f_{j\downarrow}f_{i\downarrow}\rangle, \quad \Delta^{(x)}_\downarrow = \frac{J_x}{4}\langle f_{j\uparrow}f_{i\uparrow}\rangle,\\ \Delta^{(y)}_\uparrow &= \frac{-J_y}{4}\langle f_{j\downarrow}f_{i\downarrow}\rangle, \quad \Delta^{(y)}_\downarrow = \frac{-J_y}{4}\langle f_{j\uparrow}f_{i\uparrow}\rangle,\\ \Delta^{(z)}_\uparrow &= \frac{J_z}{4}\langle f_{j\uparrow}f_{i\uparrow}\rangle, \quad \Delta^{(z)}_\downarrow = \frac{J_z}{4}\langle f_{j\downarrow}f_{i\downarrow}\rangle.\end{aligned} \quad \text{(B6)}$$

To satisfy the mean-field conditions (B6), we take

$$\begin{aligned}t_{ij,\downarrow} &= -\Delta_{ij,\downarrow} = \frac{iJ_x}{16} \quad \text{on } x \text{ links},\\ t_{ij,\downarrow} &= -\Delta_{ij,\downarrow} = \frac{iJ_y}{16} \quad \text{on } y \text{ links},\\ t_{ij,\uparrow} &= \Delta_{ij,\uparrow} = 0 \quad \text{on } z \text{ links},\\ t_{ij,\uparrow} &= -\Delta_{ij,\uparrow} = \frac{iJ_x}{16} \quad \text{on } x \text{ links},\\ t_{ij,\uparrow} &= \Delta_{ij,\uparrow} = \frac{iJ_y}{16} \quad \text{on } y \text{ links},\\ t_{ij,\downarrow} &= i\frac{J_z}{8} \qquad \Delta_{ij,\downarrow} = 0 \quad \text{on } z \text{ links},\end{aligned} \quad \text{(B7)}$$

which gives the mean-field Hamiltonian (7).

3. Theory of fluctuations about mean field

We now turn to the fluctuations about the mean-field solutions. Since symmetry dictates that these can not change the fermionic band structure, our focus will be to describe the bosonic degrees of freedom in this theory, and demonstrate that the gauge field is in a Higgsed phase with a residual Z_2 symmetry group.

The Hubbard-Stratonovich decoupling introduces four bosonic fields: $\Phi_{1,2}$, whose saddle-point expectation values are associated with fermion hopping terms; and $\Theta_{1,2}$, associated with the spin-triplet superconductivity. We parametrize their fluctuations according to

$$\begin{aligned}\Phi_{1ij} &= \mp\left(\frac{i}{16}(J_x\delta_{ij,x} + J_y\delta_{ij,y} + 2J_z\delta_{ij,z})e^{ia_{ij}} + i\phi_{ij}\right),\\ \Phi_{2ij} &= \pm i\left(\frac{J_z}{8}\delta_{ij,z}e^{i\tilde{\theta}_{ij}} + \tilde{\rho}_{ij}\right),\\ \Theta_{1ij} &= \pm i\left(\frac{J_y}{16}\delta_{ij,y}e^{i\theta_{ij}} + \rho_{ij}\right),\\ \Theta_{2ij} &= \mp i\left(\frac{J_x}{16}\delta_{ij,x}e^{i\theta_{ij}} + \rho_{ij}\right),\end{aligned} \quad \text{(B8)}$$

where the functions $\delta_{ij,x,y,z}$ have support on x,, y, and z links, respectively, and the top (bottom) sign is taken for edges oriented from sublattice 1(2) to sublattice 2(1).

The physical interpretation of these fields is as follows. Φ_1 is associated with the spin-rotation-invariant hopping terms familiar from spinon decompositions of the Heisenberg model.[7,8] The phase variables a_{ij} are the spatial components of the gauge fields associated with the constraints (3); fluctuations in the amplitude of this hopping term are parametrized by the scalar ϕ.

The remaining terms parametrize fluctuations of a condensed superfluid, which breaks the SU(2) gauge group down to Z_2. We combine the fields associated with Θ_1 and Θ_2, each of which is nonvanishing at mean field either on x or y links, respectively, into a single pair of scalar fields ρ,θ defined on all links in the lattice. Since, at mean field, Θ's expectation value generates a spinful superconducting pairing, θ is the phase of a charged superfluid and, hence, in the condensed phase becomes the longitudinal component of the corresponding gauge field. ρ parametrizes the (gapped) fluctuations in this superfluid density.

That Φ_2, the hopping antisymmetric in spin, is associated with a charged superfluid is less obvious. We will show shortly, however, that $\langle\Phi_2\rangle$ breaks the off-diagonal generators of SU(2). As these are not the same as the generator broken by the superconducting terms, we use a new field $\tilde{\theta}$ to denote the phase fluctuations.

To find the residual symmetry group, we must evaluate the SU(2) flux through each lattice plaquette at mean field.[8] It is enlightening to express the fermionic degrees of freedom in terms of the usual BCS spinors

$$\chi_q = \begin{pmatrix} f_{\uparrow,q} \\ f^{\dagger}_{\downarrow,-q} \end{pmatrix}, \tag{B9}$$

which transform under gauge transformations by $e^{i\vec{\alpha}\cdot\vec{\sigma}}$ as

$$\chi_q \rightarrow e^{i\vec{\alpha}\cdot\vec{\sigma}}\chi_q. \tag{B10}$$

In this basis, the spin-symmetric and spin-antisymmetric hopping terms can be expressed as

$$\begin{aligned} &it_{\uparrow+\downarrow}(ij)(f^{\dagger}_{i\uparrow}f_{j\uparrow} + f^{\dagger}_{i\downarrow}f_{j\downarrow} - f^{\dagger}_{j\uparrow}f_{i\uparrow} - f^{\dagger}_{j\downarrow}f_{i\downarrow}) \\ &\quad = it_{\uparrow+\downarrow}(ij)(\chi_i^{\dagger}\chi_j - \chi_j^{\dagger}\chi_i), \\ &it_{\uparrow-\downarrow}(ij)(f^{\dagger}_{i\uparrow}f_{j\uparrow} - f^{\dagger}_{i\downarrow}f_{j\downarrow} - f^{\dagger}_{j\uparrow}f_{i\uparrow} + f^{\dagger}_{j\downarrow}f_{i\downarrow}) \\ &\quad = it_{\uparrow-\downarrow}(ij)(\chi_i^{\dagger}\sigma_z\chi_j - \chi_j^{\dagger}\sigma_z\chi_i). \end{aligned} \tag{B11}$$

As promised, the first term is gauge invariant under all generators. The effect of a gauge transformation on the second term is to conjugate the matrix σ_z by $e^{i\vec{\alpha}\cdot\vec{\sigma}}$. Hence, this term is invariant under the U(1) subgroup comprised of rotations about the z axis, but not under rotations by the two generators σ_x and σ_y. Fluctuations in $\tilde{\theta}$ are therefore associated with the longitudinal modes of the broken generators $a_{ij}^{(x,y)}$.

The remaining U(1) symmetry is broken by the superconducting terms. As the pairing occurs here in the spin-triplet channel, these can not naturally be expressed in the BCS basis; however, they are clearly charged under the residual U(1) symmetry $f_{i\sigma} \Rightarrow e^{i\alpha_i}f_{i\sigma}$. Hence, the U(1) symmetry is broken to the Z_2 subgroup $f_{i\sigma} \Rightarrow \pm f_{i\sigma}$, which is the residual gauge symmetry of the Hamiltonian. [Indeed, the spin-triplet superconducting terms are certainly not gauge equivalent to the terms associated with $t_{\uparrow-\downarrow}$, guaranteeing that the SU(2) gauge symmetry is fully broken to Z_2, rather than to a residual U(1) as might otherwise be the case.] As usual, the phase fluctuations θ can be absorbed by means of a gauge transformation into the longitudinal modes of the broken U(1) generator.

As an aside: Equation (B8) reveals that the longitudinal modes of the broken generators are confined to x-y chains and z links in the lattice, respectively. Since the corresponding gauge fluctuations are no longer purely transverse in the condensed phase, this means that only the residual Z_2 gauge field and the amplitude fluctuations are free to propagate in both dimensions of the lattice. This explains, to a large degree, why the effect of including these bosons in the theory is so innocuous.

In summary, the fluctuations about mean field are described by the real scalars ρ, $\tilde{\rho}$, and ϕ, describing fluctuations in the amplitudes of the various condensed bosonic fields, and the SU(2) gauge field that is Higgsed in a bi-adjoint representation to a residual symmetry group Z_2, which we may consider to have absorbed the remaining phase fluctuations as two Goldstone bosons.

APPENDIX C: MEAN-FIELD THEORY OF THE GAPPED B PHASE

Here, we describe the mean-field theory in the presence of the three-spin interaction, which leads to the gapped topological B phase. We will show that the band structure discussed in Sec. V is, up to irrelevant operators, a saddle point of an appropriate action and thus constitutes at least a self-consistent mean-field solution to the fermion problem, if not a global minimum of the action.

We begin by rewriting the three-spin interaction as a sum of products of six-fermion interaction terms

$$\begin{aligned} S_i^x S_j^y S_k^z = \frac{i}{8}(&f^{\dagger}_{i\uparrow}f^{\dagger}_{j\uparrow}f_{j\downarrow}f_{i\downarrow} - f_{i\uparrow}f_{j\uparrow}f^{\dagger}_{j\downarrow}f^{\dagger}_{i\downarrow} \\ &+ f^{\dagger}_{i\uparrow}f_{j\uparrow}f^{\dagger}_{j\downarrow}f_{i\downarrow} - f^{\dagger}_{j\uparrow}f_{i\uparrow}f^{\dagger}_{i\downarrow}f_{j\downarrow})(2f^{\dagger}_{k\downarrow}f_{k\downarrow} - 1), \end{aligned} \tag{C1}$$

where we have used $n_{i\uparrow} = 1 - n_{i\downarrow}$ to express S_i^z in terms of down spins only. Of the possible fermion bilinears, only $(f^{\dagger}_{i\uparrow}f_{j\uparrow})$,$(f^{\dagger}_{i\downarrow}f_{j\downarrow})$, and $(f^{\dagger}_{j\downarrow}f_{k\downarrow})$ (together with their analogs in the particle-particle and hole-hole channels) have nonvanishing expectation values at mean field $(\langle f^{\dagger}_{k\downarrow}f_{k\downarrow} - f^{\dagger}_{k\uparrow}f_{k\uparrow}\rangle = 0)$. This gives us two possible ways to replace two of the three fermion bilinears by their mean-field values. First, we may take

$$\begin{aligned} \frac{i}{8}(&\langle f^{\dagger}_{i\uparrow}f^{\dagger}_{j\uparrow}\rangle\langle f_{j\downarrow}f_{i\downarrow}\rangle - \langle f_{i\uparrow}f_{j\uparrow}\rangle\langle f^{\dagger}_{j\downarrow}f^{\dagger}_{i\downarrow}\rangle + \langle f^{\dagger}_{i\uparrow}f_{j\uparrow}\rangle\langle f^{\dagger}_{j\downarrow}f_{i\downarrow}\rangle \\ &- \langle f^{\dagger}_{j\uparrow}f_{i\uparrow}\rangle\langle f^{\dagger}_{i\downarrow}f_{j\downarrow}\rangle)(2f^{\dagger}_{k\downarrow}f_{k\downarrow} - 1), \end{aligned} \tag{C2}$$

which vanishes in the mean-field solution relevant to the Kitaev model as the fermion bilinears are purely imaginary in position space. The only remaining possibility is

$$\begin{aligned} \frac{i}{8}[&\langle f^{\dagger}_{i\uparrow}f^{\dagger}_{j\uparrow}\rangle\langle f^{\dagger}_{k\downarrow}f_{j\downarrow}\rangle f_{i\downarrow}f_{k\downarrow} - \langle f^{\dagger}_{i\uparrow}f^{\dagger}_{j\uparrow}\rangle\langle f_{j\downarrow}f_{k\downarrow}\rangle f^{\dagger}_{k\downarrow}f_{i\downarrow} \\ &- \langle f_{i\uparrow}f_{j\uparrow}\rangle\langle f^{\dagger}_{k\downarrow}f^{\dagger}_{j\downarrow}\rangle f^{\dagger}_{i\downarrow}f_{k\downarrow} + \langle f_{i\uparrow}f_{j\uparrow}\rangle\langle f^{\dagger}_{j\downarrow}f_{k\downarrow}\rangle f^{\dagger}_{k\downarrow}f^{\dagger}_{i\downarrow} \\ &+ \langle f^{\dagger}_{i\uparrow}f_{j\uparrow}\rangle\langle f^{\dagger}_{k\downarrow}f^{\dagger}_{j\downarrow}\rangle f_{i\downarrow}f_{k\downarrow} - \langle f^{\dagger}_{i\uparrow}f_{j\uparrow}\rangle\langle f^{\dagger}_{j\downarrow}f_{k\downarrow}\rangle f^{\dagger}_{k\downarrow}f_{i\downarrow} \\ &+ \langle f^{\dagger}_{j\uparrow}f_{i\uparrow}\rangle\langle f^{\dagger}_{k\downarrow}f_{j\downarrow}\rangle f^{\dagger}_{i\downarrow}f_{k\downarrow} - \langle f^{\dagger}_{j\uparrow}f_{i\uparrow}\rangle\langle f_{j\downarrow}f_{k\downarrow}\rangle f^{\dagger}_{k\downarrow}f^{\dagger}_{i\downarrow}]. \end{aligned} \tag{C3}$$

Taking $\langle ij\rangle$ to be an x link and $\langle jk\rangle$ to be a z link, and substituting in the mean-field values given in Eq. (B7), this becomes

$$\begin{aligned} &\frac{i}{2^7}[f_{i\downarrow}f_{k\downarrow} - f^{\dagger}_{k\downarrow}f^{\dagger}_{i\downarrow} + f^{\dagger}_{k\downarrow}f_{i\downarrow} - f^{\dagger}_{i\downarrow}f_{k\downarrow}] \\ &\quad = \frac{i}{2^7}(f^{\dagger}_{i\downarrow} - f_{i\downarrow})(f^{\dagger}_{k\downarrow} - f_{k\downarrow}). \end{aligned} \tag{C4}$$

In light of the correspondence (10) between our Dirac fermions and the Majorana basis originally used to diagonalize the problem, this is exactly the term originally proposed by Ref. 1 to break $\hat{\mathbf{T}}$ and open a gap in the B phase.

Before analyzing the resulting band structure, let us understand why we may simply replace the fermion bilinears by their mean-field values, as we have done above. In fact, we can modify the Lagrangian (B2) to produce just such a term at mean-field level. To see why this is so, we consider the action

$$\mathcal{L}_{\mathcal{F}} = \chi_1^\dagger \chi_1 + \chi_2^\dagger \chi_2 + i J' \chi_1^\dagger \chi_2^\dagger \chi_3^\dagger + \text{H.c.} \tag{C5}$$

We will show that $\mathcal{L}_{\mathcal{F}}$ is well approximated by the Hubbard-Stratonovich–type action

$$\begin{aligned}\mathcal{L} = -|\Phi_1|^2 - |\Phi_2|^2 + \chi_1 \Phi_1 + \chi_2 \Phi_2 \\ - i J'(\Phi_1 \Phi_2 - \chi_2^\dagger \Phi_1 - \chi_1^\dagger \Phi_2)\chi_3^\dagger + \text{H.c.},\end{aligned} \tag{C6}$$

where $\chi_{1,2,3}$ are fermion bilinears. [The Lagrangian (B2) is of the general form of the quadratic terms in Eq. (C6), albeit with more different scalar fields. This multiplicity of indices will not affect our qualitative result.] The saddle-point equations are

$$\begin{aligned}\Phi_1 &= \chi_1^\dagger - i J' \chi_3 (\Phi_2^\dagger - \chi_2), \\ \Phi_2 &= \chi_2^\dagger - i J' \chi_3 (\Phi_1^\dagger - \chi_1).\end{aligned} \tag{C7}$$

For $J' = 0$, the saddle-point equations specify that $\Phi_i = \chi_i^\dagger$. This is also the unique solution of the saddle-point equations for $J' \neq 0$ (although in this case one might worry about instabilities that tend to drive $\Phi_{1,2}$ toward ∞ if $\langle \chi_3 \rangle \neq 0$). Hence, the extra term does not modify the structure of the mean-field equations, except inasmuch as $\langle \chi_{1,2} \rangle$ might be modified by the new interaction.

As in the standard Hubbard-Stratonovich decoupling, we would like to integrate out $\Phi_{1,2}$ to obtain $\mathcal{L}_{\mathcal{F}}$. As the Lagrangian (C6) is no longer quadratic in the variables Φ_i, χ_i, we will not be able to perform the integral exactly; rather, we will obtain $\mathcal{L}_{\mathcal{F}}$ as the lowest-order term in an expansion in J'. To see this, it is helpful to reexpress $\mathcal{L}$ as

$$\mathcal{L} = -|\tilde{\Phi}_1|^2 - |\tilde{\Phi}_2|^2 - i J' \tilde{\Phi}_1 \tilde{\Phi}_2 \chi_3^\dagger + \text{H.c.} + \mathcal{L}_{\mathcal{F}}, \tag{C8}$$

where $\tilde{\Phi}_i \equiv \Phi_i - \chi_i^\dagger$. In the standard Hubbard-Stratonovich transformation, there would be at this point no cross terms coupling fermions to the scalar fields. We could therefore integrate out the latter exactly, and this proves that (C6) is exactly equivalent to $\mathcal{L}_{\mathcal{F}}$. Here, we are unable to eliminate the cross term $\Phi_1 \Phi_2 \chi_3^\dagger$ by further shifting the scalar fields, so that integrating out the $\tilde{\Phi}$ fields will not reproduce $\mathcal{L}_{\mathcal{F}}$ exactly. If we take J' small, however, we may consider the effect of the cross term perturbatively, and ask what the undesired additions to the fermionic action will be. The exact correction is given by evaluating the series

$$\begin{aligned}\delta \mathcal{L}_{\mathcal{F}} = \log \left\{ \int [D\tilde{\Phi}_1][D\tilde{\Phi}_2] e^{i \int |\tilde{\Phi}_1|^2 + |\tilde{\Phi}_2|^2} \right. \\ \left. \times \sum_{n=0}^{\infty} \frac{(i J')^n}{n!} (\tilde{\Phi}_1 \tilde{\Phi}_2 \chi_3^\dagger + \text{H.c.})^n \right\}.\end{aligned} \tag{C9}$$

Terms with n odd integrate to 0 since the action contains only even powers of $\tilde{\Phi}_i$. Hence, the leading correction is of order J'^2; to linear order in J', then, we have recovered exactly the fermionic action we wanted. Since the scalar-scalar-fermion bilinear interaction is decidedly irrelevant (all scalars here are massive), we may conclude that the difference between the action (C6) and the true fermionic action $\mathcal{L}_{\mathcal{F}}$ is unimportant, at least for the low-energy physics.

The general form of this correction is simple to understand. The leading-order correction in the series (C9) is proportional to $\frac{(J')^2}{2} \chi_3^\dagger \chi_3$. If we take χ_3 to have the form $f_{i\downarrow} f_{k\downarrow}$, then we have $\chi_3^\dagger \chi_3 = (\chi_3^\dagger \chi_3)^r = \hat{n}_{i\downarrow} \hat{n}_{k\downarrow}$ for all r, and all terms in the series induce the same type of extraneous interaction, which is to induce a second-neighbor Coulomb repulsion term.

We conclude that at least the low-energy structure of the phase we are interested in can be obtained by studying the Lagrangian (C6). We may now proceed as in Sec. B 2, obtaining a mean-field solution, which satisfies

$$\langle \Phi_i \rangle = \langle \chi_i^\dagger \rangle. \tag{C10}$$

As noted above, the mean-field consistency conditions are identical to those at $J' = 0$; the only new feature of this saddle point is that it now includes quadratic terms coupling fermions on the same sublattice, such as

$$J \, \langle \Phi_1 \rangle \langle \Phi_2 \rangle f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger. \tag{C11}$$

This means that, to lowest order in J', the effect of the three-spin interaction is, exactly as originally postulated by Ref. 1, to modify the band structure by adding next-nearest-neighbor quadratic couplings. (We now also have to contend with the four-fermion interactions; however, when the quadratic problem has no Fermi surface, we do not expect these to be associated with instabilities of the free fermion problem and, hence, we can safely drop them without altering the qualitative nature of the physics.)

1. Form of the mean-field Hamiltonian with three-spin interactions

Here we will derive the expression (50) for the terms induced by the set of all three-spin interactions at mean-field. There are three distinct three-spin interactions that we must consider:

$$\begin{aligned}& S_i^x S_j^y S_k^z \quad \text{if } r_{ik} = \hat{l}_1, \\ & S_i^y S_j^x S_k^z \quad \text{if } r_{ik} = \hat{l}_2, \\ & S_i^x S_j^z S_k^y \quad \text{if } r_{ik} = \hat{x}.\end{aligned} \tag{C12}$$

The contributions to mean field involve decoupling the resulting six-fermion interactions into combinations of a pair of two-point functions multiplying a fermion bilinear.

First, we show that only contributions multiplying bilinears of the form $f_{i\sigma} f_{k\sigma}, f_{i\sigma}^\dagger f_{k\sigma}$, etc., are nonvanishing. The mean-field eigenfunctions imply that $\langle S_i^\alpha \rangle = 0$ on each site. To show that $\langle S_i^\alpha S_j^{\alpha'} \rangle = 0$ if $\alpha \neq \alpha'$, we first note that if $\alpha = x, y$ and $\alpha' = z$, any grouping of the resulting four-fermion interaction into pairs involves one term in each pair, which contains both a spin-up and spin-down fermion. Since the two-point functions of all terms involving spin flips are strictly 0, these terms consequently all vanish. If $\alpha = x, \alpha' = y$, then we have

$$-i(f_{i\uparrow}^\dagger f_{i\downarrow} + f_{i\downarrow}^\dagger f_{i\uparrow})(f_{j\uparrow}^\dagger f_{j\downarrow} - f_{j\downarrow}^\dagger f_{j\uparrow})$$
$$= i(\langle f_{i\uparrow}^\dagger f_{j\uparrow}^\dagger\rangle\langle f_{i\downarrow} f_{j\downarrow}\rangle + \langle f_{i\downarrow}^\dagger f_{j\downarrow}\rangle\langle f_{j\uparrow}^\dagger f_{i\uparrow}\rangle - \text{H.c.}), \quad \text{(C13)}$$

which vanishes since the two-point function on every link is purely imaginary, so that the products shown are purely real.

The only remaining possibility is terms in which the two-point functions whose mean-field expectation we take involve fermion operators from all three sites. Since all two-point functions between sites i and k vanish at mean field (this is guaranteed by the discrete symmetries $\hat{\mathbf{C}}$ and $\hat{\mathbf{T}}$), the only possibility is terms that multiply fermion bilinears, which couple the sites i and k.

Our next task is to understand the precise form of these terms. For $r_{ij} = \hat{l}_{1,2}$, it is convenient to write $S_i^z = f_{i\downarrow} f_{i\downarrow}^\dagger - f_{i\downarrow}^\dagger f_{i\downarrow}$; for $r_{ij} = \hat{x}$, we write $S_i^z = f_{i\uparrow}^\dagger f_{i\uparrow} - f_{i\uparrow} f_{i\uparrow}^\dagger$. The resulting expressions contain couplings only between the spin-down fermions on sites i and k. Thus, the three-spin interaction does not modify the band structure of the spin-up fermions, which remain localized, at least at the mean-field level.

The quadratic couplings between the down spins induced by the three-spin interactions can be expressed as

$$S_i^x S_j^y S_k^z = \frac{i}{8}\big[\big(T^{(1)}_{ijk;\downarrow} + T^{(3)}_{ijk;\downarrow}\big) f_{k\downarrow}^\dagger f_{i\downarrow} + \big(T^{(2)}_{ijk;\downarrow} + T^{(4)}_{ijk;\downarrow}\big) f_{k\downarrow} f_{i\downarrow} + \big(T^{(6)}_{ijk;\downarrow} + T^{(8)}_{ijk;\downarrow}\big) f_{i\downarrow}^\dagger f_{k\downarrow} + \big(T^{(5)}_{ijk;\downarrow} + T^{(7)}_{ijk;\downarrow}\big) f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger\big],$$

$$S_i^y S_j^x S_k^z = \frac{i}{8}\big[\big(T^{(1)}_{ijk;\downarrow} - T^{(3)}_{ijk;\downarrow}\big) f_{k\downarrow}^\dagger f_{i\downarrow} + \big(T^{(2)}_{ijk;\downarrow} - T^{(4)}_{ijk;\downarrow}\big) f_{k\downarrow} f_{i\downarrow} - \big(T^{(6)}_{ijk;\downarrow} - T^{(8)}_{ijk;\downarrow}\big) f_{i\downarrow}^\dagger f_{k\downarrow} - \big(T^{(5)}_{ijk;\downarrow} - T^{(7)}_{ijk;\downarrow}\big) f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger\big], \quad \text{(C14)}$$

$$S_i^x S_j^z S_k^y = \frac{i}{8}\big[\big(T^{(1)}_{ijk;\uparrow} - T^{(3)}_{ijk;\uparrow}\big) f_{k\downarrow}^\dagger f_{i\downarrow} + \big(T^{(2)}_{ijk;\uparrow} - T^{(4)}_{ijk;\uparrow}\big) f_{k\downarrow} f_{i\downarrow} - \big(T^{(6)}_{ijk;\uparrow} - T^{(8)}_{ijk;\uparrow}\big) f_{i\downarrow}^\dagger f_{k\downarrow} - \big(T^{(5)}_{ijk;\uparrow} - T^{(7)}_{ijk;\uparrow}\big) f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger\big]$$

with

$$T^{(1)}_{ijk;\sigma} = -\langle f_{i\uparrow}^\dagger f_{j\uparrow}^\dagger\rangle\langle f_{j\sigma} f_{k\sigma}\rangle = -\frac{16}{J_{ij}J_{jk}}(\Delta_\uparrow^{(ij)})^*\Delta_\sigma^{(kj)},$$
$$T^{(2)}_{ijk;\sigma} = -\langle f_{i\uparrow}^\dagger f_{j\uparrow}^\dagger\rangle\langle f_{k\sigma}^\dagger f_{j\sigma}\rangle = -\frac{16}{J_{ij}J_{jk}}(\Delta_\uparrow^{(ij)})^*\left(t_\sigma^{(kj)}\right)^*,$$
$$T^{(3)}_{ijk;\sigma} = -\langle f_{i\uparrow}^\dagger f_{j\uparrow}\rangle\langle f_{j\sigma}^\dagger f_{k\sigma}\rangle = -\frac{16}{J_{ij}J_{jk}}(t_\uparrow^{(ij)})^* t_\sigma^{(kj)},$$
$$T^{(4)}_{ijk;\sigma} = \langle f_{i\uparrow}^\dagger f_{j\uparrow}\rangle\langle f_{k\sigma}^\dagger f_{j\sigma}^\dagger\rangle = \frac{16}{J_{ij}J_{jk}}(t_\uparrow^{(ij)})^*\left(\Delta_\sigma^{(kj)}\right)^*,$$
$$T^{(5)}_{ijk;\sigma} = \langle f_{j\uparrow}^\dagger f_{i\uparrow}\rangle\langle f_{j\sigma} f_{k\sigma}\rangle = -\frac{16}{J_{ij}J_{jk}} t_\uparrow^{(ij)}\Delta_\sigma^{(kj)}, \quad \text{(C15)}$$
$$T^{(6)}_{ijk;\sigma} = \langle f_{j\uparrow}^\dagger f_{i\uparrow}\rangle\langle f_{k\sigma}^\dagger f_{j\sigma}\rangle = \frac{16}{J_{ij}J_{jk}} t_\uparrow^{(ij)}\left(t_\sigma^{(kj)}\right)^*,$$
$$T^{(7)}_{ijk;\sigma} = \langle f_{j\uparrow} f_{i\uparrow}\rangle\langle f_{j\sigma}^\dagger f_{k\sigma}\rangle = \frac{16}{J_{ij}J_{jk}}\Delta_\uparrow^{(ij)} t_\sigma^{(kj)},$$
$$T^{(8)}_{ijk;\sigma} = \langle f_{j\uparrow} f_{i\uparrow}\rangle\langle f_{k\sigma}^\dagger f_{j\sigma}^\dagger\rangle = \frac{16}{J_{ij}J_{jk}}\Delta_\uparrow^{(ij)}\left(\Delta_\sigma^{(kj)}\right)^*,$$

where we have used $t^{(jk)*} = t^{(kj)}$, $\Delta^{(jk)*} = \Delta^{(kj)}$. [Here we have defined $\Delta^{(ab)} = \Delta^{(x,z)}$ on x and z links, and $-\Delta^{(y)}$ on y links, in accordance with Eq. (B7).]

We next substitute in the mean-field values given in Eq. (B7) for t, Δ on each link. We take t to be the hopping from sublattice 1 to sublattice 2 ($t_\sigma^{(ij)} = \langle f_{\vec{R}1\sigma}^\dagger f_{\vec{R}'2\sigma}\rangle$), and similarly for Δ. Here we write the induced quadratic couplings between two sites on sublattice 1; the couplings between sites on sublattice 2 are the same, but with $r_{ij} \to -r_{ij}$.

For $r_{ij} = \hat{l}_1$, the interaction is of the form $J' S_i^x S_j^y S_k^z$, with ij an x link and jk a z link. We thus have $\Delta_\downarrow^{(jk)} = 0$, giving an interaction of

$$2iJ'[-t_\uparrow^{(x)*} t_\downarrow^{(z)} f_{k\downarrow}^\dagger f_{i\downarrow} - \Delta_\uparrow^{(x)*} t_\downarrow^{(z)*} f_{k\downarrow} f_{i\downarrow}$$
$$+ t_\uparrow^{(x)} t_\downarrow^{(z)*} f_{i\downarrow}^\dagger f_{k\downarrow} + \Delta_\uparrow^{(x)} t_\downarrow^{(z)} f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger] \quad \text{(C16)}$$

with

$$\Delta_\uparrow^{(x)} = -i\frac{J_x}{16}, \quad t_\uparrow^{(x)} = -i\frac{J_x}{16}, \quad t_\downarrow^{(z)} = -i\frac{J_z}{8}. \quad \text{(C17)}$$

Similarly, for $r_{ij} = \hat{l}_2$, we have $J' S_i^y S_j^x S_k^z$, with ij a y link and jk a z link. Hence, again $\Delta_\downarrow^{(jk)} = 0$, and the interaction is

$$2iJ'[t_\uparrow^{(y)*} t_\downarrow^{(z)} f_{k\downarrow}^\dagger f_{i\downarrow} + \Delta_\uparrow^{(y)*} t_\downarrow^{(z)*} f_{k\downarrow} f_{i\downarrow}$$
$$- t_\uparrow^{(y)} t_\downarrow^{(z)*} f_{i\downarrow}^\dagger f_{k\downarrow} - \Delta_\uparrow^{(y)} t_\downarrow^{(z)} f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger] \quad \text{(C18)}$$

with

$$\Delta_\uparrow^{(y)} = i\frac{J_y}{16}, \quad t_\uparrow^{(y)} = -i\frac{J_y}{16}, \quad t_\downarrow^{(z)} = -i\frac{J_z}{8}. \quad \text{(C19)}$$

For $r_{ij} = \hat{x}$, we have $J' S_i^x S_j^z S_k^y$, with ij an x link and jk a y link. This gives the interaction

$$iJ'[(\Delta_\uparrow^{(x)*}\Delta_\uparrow^{(y)*} - t_\uparrow^{(x)*} t_\uparrow^{(y)}) f_{k\downarrow}^\dagger f_{i\downarrow} + (\Delta_\uparrow^{(x)*} t_\uparrow^{(y)*}$$
$$+ t_\uparrow^{(x)*}\Delta_\uparrow^{(y)*}) f_{k\downarrow} f_{i\downarrow} + (-t_\uparrow^{(x)} t_\uparrow^{(y)*} - \Delta_\uparrow^{(x)}\Delta_\uparrow^{(y)}) f_{i\downarrow}^\dagger f_{k\downarrow}$$
$$+ (t_\uparrow^{(x)}\Delta_\uparrow^{(y)*} + \Delta_\uparrow^{(x)} t_\uparrow^{(y)}) f_{i\downarrow}^\dagger f_{k\downarrow}^\dagger] \quad \text{(C20)}$$

with

$$\Delta_{\uparrow}^{(x)} = -i\frac{J_x}{16}, \quad \Delta_{\uparrow}^{(y)} = i\frac{J_y}{16}, \\ t_{\uparrow}^{(x)} = -i\frac{J_x}{16}, \quad t_{\uparrow}^{(y)} = -i\frac{J_y}{16}. \tag{C21}$$

In all three cases, we obtain the mean-field interaction

$$\pm 2iJ'[f_{k\downarrow}^{\dagger}f_{i\downarrow} - f_{k\downarrow}f_{i\downarrow} + f_{i\downarrow}^{\dagger}f_{k\downarrow} - f_{i\downarrow}^{\dagger}f_{k\downarrow}^{\dagger}] \\ = \pm 2iJ'(f_{k\downarrow}^{\dagger} - f_{k\downarrow})(f_{i\downarrow}^{\dagger} - f_{i\downarrow}). \tag{C22}$$

We see that this induces a coupling only between Majorana modes in the dispersing band, leaving the band structure of the Majoranas localized on the z links unaltered.

Hence, the net effect of adding the three-spin interaction, at mean-field level, is exactly to add the next-nearest-neighbor couplings to the dynamical Majorana modes, while leaving the localized modes unchanged.

APPENDIX D: INDUCING CHERN-SIMONS TERMS BY INTEGRATING OUT FERMIONS IN THE GAPPED B PHASE

Here, we will consider the one-loop perturbative correction to the effective U(1) gauge-field propagator due to the low-energy fermions in the gapped phase. We demonstrate that although the Dirac point is intrinsically a property of the band structure of the superconductor—such that the electron bubble has both particle-particle and particle-hole contributions—the matrix structure about the Dirac point is such that integrating out the low-energy fermions produces exactly the same Chern-Simons correction to the effective action as doing so for a normal Dirac cone.

Since the Dirac cone is in only one of the four fermion bands, and we are interested only in the long-wavelength theory, we will isolate the effect of the propagator of the dispersing Majorana band. The general form of the spin-down propagator in the gapped B phase is

$$G_{\downarrow\downarrow q} = \frac{1}{2}\left\{\frac{1}{4m_q^2+\omega^2+|\Delta_q-t_q|^2}\begin{pmatrix} -2m_q-i\omega & -i(\Delta_q-t_q) & 2m_q+i\omega & i(\Delta_q-t_q) \\ i(\Delta_q^*-t_q^*) & 2m_q-i\omega & -i(\Delta_q^*-t_q^*) & i\omega-2m_q \\ 2m_q+i\omega & i(\Delta_q-t_q) & -2m_q-i\omega & -i(\Delta_q-t_q) \\ -i(\Delta_q^*-t_q^*) & i\omega-2m_q & i(\Delta_q^*-t_q^*) & 2m_q-i\omega \end{pmatrix}\right. \\ \left. + \frac{1}{\omega^2+|\Delta_q+t_q|^2}\begin{pmatrix} -i\omega & i(\Delta_q+t_q) & -i\omega & i(\Delta_q+t_q) \\ -i(\Delta_q^*+t_q^*) & -i\omega & -i(\Delta_q^*+t_q^*) & -i\omega \\ -i\omega & i(\Delta_q+t_q) & -i\omega & i(\Delta_q+t_q) \\ -i(\Delta_q^*+t_q^*) & -i\omega & -i(\Delta_q^*+t_q^*) & -i\omega \end{pmatrix}\right\}, \tag{D1}$$

where we use the basis $\psi = (c_{q1}, c_{q2}, c_{-q1}^{\dagger}, c_{-q2}^{\dagger})^T$.

Here, we choose $t_q = -2J_z - J_x e^{i\vec{q}\cdot\hat{l}_1} - J_y e^{i\vec{q}\cdot\hat{l}_2}, \Delta_q = J_x e^{i\vec{q}\cdot\hat{l}_1} + J_y e^{i\vec{q}\cdot\hat{l}_2}$. In this case, the bottom line is the propagator of the flat band (energies given by $\pm|t+\Delta| = \pm 2J_z$); the top line is the propagator of the dispersing band, which captures all of the low-energy physics near the Dirac cones. It is easy to check that cross terms between the two spin-down bands vanish at one-loop order in the fermion correction, so that we will drop contributions of the flat gapped band entirely.

In the vicinity of the Dirac cone $\vec{q} = (\frac{4\pi}{3}, 0)$, at the isotropic point $J_x = J_y = J_z$, we have

$$\Delta_q - t_q \approx \sqrt{3}J, \quad m_q \approx \tfrac{3}{2}\sqrt{3}J'. \tag{D2}$$

Near this point in the Brillouin zone, then, the part of the propagator that we are interested in can be expressed as

$$G_{c;q,\omega} = \frac{1}{2}\left(G_{c;q,\omega}^{(0)} + G_{c;q,\omega}^{(sc)}\right), \\ G_{c;q,\omega}^{(0)} = \frac{1}{4m_q^2+\omega^2+|\Delta_q-t_q|^2}(p^{\mu}\sigma_{\mu} + 2m\sigma_z)\otimes\mathbf{1}, \\ G_{c;q,\omega}^{(sc)} = \frac{1}{4m_q^2+\omega^2+|\Delta_q-t_q|^2}(p^{\mu}\sigma_{\mu} + 2m\sigma_z)\otimes\sigma_x, \tag{D3}$$

with $\sigma^{\mu} = (\mathbf{1}, \sigma_y, \sigma_x)$. In addition to the usual term ($G_{c;q,\omega}^{(0)}$), the fermion propagator contains an anomalous term ($G_{c;q,\omega}^{(sc)}$) due to the presence of superconductivity. The 2×2 matrix structure of both of these terms is, however, the same.

In this long-wavelength limit, the interaction between fermions and the gauge field is

$$A_q^{\mu}\sum_k \psi_k^{\dagger}\gamma_{\mu}\psi_{k-q} - 2\delta_{\mu 0}\delta_{q0}, \tag{D4}$$

where $\gamma_{\mu} = \sigma_{\mu}\otimes\mathbf{1}$, and the last term occurs due to normal ordering. (Here it should be understood that the sum encompasses only half the Brillouin zone.) The one-loop correction to the gauge-field effective action induced by the fermion terms is therefore

$$\mathcal{L}_{\mu\nu}^{(G)}(\vec{p},\Omega) = \int \frac{d^3p}{(2\pi)^3}\mathrm{Tr}[\gamma_{\mu}G_{c;q,\omega}\gamma_{\nu}G_{c;q+p,\omega+\Omega}]. \tag{D5}$$

By using the expression (D3), we find that traces of the cross terms between $G_{c;q,\omega}^{(0)}$ and $G_{c;q,\omega}^{(sc)}$ vanish, leaving

$$\mathcal{L}_{\mu\nu}^{(G)}(\vec{p},\Omega) \\ = \frac{1}{4}\left\{2\mathcal{L}_{\mu\nu}^{(1)} + \int\frac{d^3p}{(2\pi)^3}\mathrm{Tr}[\gamma_{\mu}G_{c;q,\omega}^{(sc)}\gamma_{\nu}G_{c;q+p,\omega+\Omega}^{(sc)}]\right\}, \tag{D6}$$

where $\mathcal{L}_{\mu\nu}^{(1)}$ is the effective action induced by the usual (2 + 1)-dimensional Dirac cone (appearing here with a multiplicative factor of 2 since we have counted both terms of the form $f_{qi}^{\dagger}f_{qi}$ and $f_{-q,i}^{\dagger}f_{-qi}$, effectively counting the contribution of both Dirac cones). The second contribution, due to the superconducting terms, also has precisely the same form as the first, since $G^{(sc)}$ has the same 2×2 structure as $G^{(0)}$. The factor of $\frac{1}{4}$ (due to the $\frac{1}{2}$ in $G^{(0)}$ relative to its usual value) is exactly canceled by the factor of 4 from these contributions. This gives exactly the one-loop correction expected from a single Dirac cone in QED, albeit with a mass of $2m$ rather than m.

---

[1] A. Kitaev, Ann. Phys. (NY) **321**, 2 (2006).
[2] A. Y. Kitaev, Ann. Phys. (NY) **303**, 2 (2003).
[3] P. W. Anderson, Mater. Res. Bull. **8**, 153 (1973).
[4] P. W. Anderson, Science **235**, 1196 (1987).
[5] I. Affleck, Z. Zou, T. Hsu, and P. W. Anderson, Phys. Rev. B **38**, 745 (1988).
[6] I. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774 (1988).
[7] J. B. Marston and I. Affleck, Phys. Rev. B **39**, 11538 (1989).
[8] X.-G. Wen, Phys. Rev. B **65**, 165113 (2002).
[9] V. Kalmeyer and R. B. Laughlin, Phys. Rev. Lett. **59**, 2095 (1987).
[10] R. B. Laughlin and Z. Zou, Phys. Rev. B **41**, 664 (1990).
[11] Z. Zou and R. B. Laughlin, Phys. Rev. B **42**, 4073 (1990).
[12] X.-G. Wen, F. Wilczek, and A. Zee, Phys. Rev. B **39**, 11413 (1989).
[13] P. Coleman and N. Andrei, J. Magn. Magn. Mater. **76-77**, 504 (1988).
[14] N. Andrei and P. Coleman, Phys. Rev. Lett. **62**, 595 (1989).
[15] Usually, this is done by introducing a fermion flavor index N, and treating $\frac{1}{N}$ as a small parameter so that the perturbative expansion is controlled, as described in Ref. 7.
[16] K. P. Schmidt, S. Dusuel, and J. Vidal, Phys. Rev. Lett. **100**, 057208 (2008).
[17] S. Dusuel, K. P. Schmidt, and J. Vidal, Phys. Rev. Lett. **100**, 177204 (2008).
[18] G. Kells, A. T. Bolukbasi, V. Lahtinen, J. K. Slingerland, J. K. Pachos, and J. Vala, Phys. Rev. Lett. **101**, 240404 (2008).
[19] J. Vidal, K. P. Schmidt, and S. Dusuel, Phys. Rev. B **78**, 245121 (2008).
[20] T. H. Hansson, V. Oganesyan, and S. L. Sondhi, Ann. Phys. (NY) **313**, 497 (2004).
[21] M. de Wild Propitius, Ph.D. thesis, University of Amsterdam, 1995.
[22] K. S. Tikhonov, M. V. Feigel'man, and A. Y. Kitaev, e-print arXiv:1008.4106.
[23] D. S. Rokhsar, Phys. Rev. B **42**, 2526 (1990).
[24] G. Kells, J. K. Slingerland, and J. Vala, Phys. Rev. B **80**, 125415 (2009).
[25] N. Read and D. Green, Phys. Rev. B **61**, 10267 (2000).
[26] J. Chaloupka, G. Jackeli, and G. Khaliullin, Phys. Rev. Lett. **105**, 027204 (2010).
[27] H. Jiang, Z. Gu, X. Qi, and S. Trebst, Phys. Rev. B **83**, 245104 (2011).
[28] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).
[29] P. A. Lee and N. Nagaosa, Phys. Rev. B **46**, 5621 (1992).
[30] L. Fu and C. L. Kane, Phys. Rev. Lett. **100**, 096407 (2008).