

**Topological and magnetic phase transitions in the bilayer Kitaev-Ising model**Aayush Vijayvargia <sup>1</sup>, Urban F. P. Seifert,<sup>2</sup> and Onur Erten<sup>1</sup><sup>1</sup>*Department of Physics, Arizona State University, Tempe, Arizona 85287, USA*<sup>2</sup>*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA*

(Received 16 November 2023; accepted 9 January 2024; published 30 January 2024; corrected 25 April 2024)

We investigate the phase diagram of a bilayer Kitaev honeycomb model with Ising interlayer interactions, deriving effective models via perturbation theory and performing Majorana mean-field theory calculations. We show that a diverse array of magnetic and topological phase transitions occur, depending on the direction of the interlayer Ising interaction and the relative sign of Kitaev interactions. When two layers have the same sign of the Kitaev interaction, a first-order transition from a Kitaev spin liquid to a magnetically ordered state takes place. The magnetic order points along the Ising axis and it is (anti)ferromagnetic for (anti)ferromagnetic Kitaev interactions. However, when two layers have opposite signs of the Kitaev interaction, we observe a notable weakening of magnetic ordering tendencies and the Kitaev spin liquid survives up to a remarkably larger interlayer exchange. Our mean-field analysis suggests the emergence of an intermediate gapped  $\mathbb{Z}_2$  spin-liquid state, which eventually becomes unstable upon vison condensation. The confined phase is described by a highly frustrated  $120^\circ$  compass model. We furthermore use perturbation theory to study the model with the Ising axis pointing along the  $\hat{z}$  axis or lying in the  $xy$  plane. In both cases, our analysis reveals the formation of one-dimensional Ising chains, which remain decoupled in perturbation theory, resulting in a subextensive ground-state degeneracy. Our results highlight the interplay between topological order and magnetic ordering tendencies in bilayer quantum spin liquids.

DOI: [10.1103/PhysRevB.109.024439](https://doi.org/10.1103/PhysRevB.109.024439)**I. INTRODUCTION**

Quantum spin liquids (QSLs) are a unique class of phases in quantum magnets that are not uniquely characterized by local order parameters [1–4], but instead exhibit long-range entanglement, fractionalization, and emergent gauge fields [5–7], which are understood to be stabilized by strong quantum fluctuations. Since the first proposal for a QSL by Anderson [8] in 1973, there has been remarkable progress in the identification of both theoretical models that may exhibit QSL ground states and the discovery of candidate materials that exhibit experimental signatures which might be compatible with QSL behavior. In this regard, the Kitaev model on a honeycomb lattice [9] plays an exceptional role as a spin model for a QSL that *both* can be solved exactly and may be (approximately) realized in materials, most prominently  $\alpha$ - $\text{RuCl}_3$  [10,11].

Further, in recent years, remarkable experimental progress and theoretical analysis has made evident that bilayers and moiré superlattices of two-dimensional (2D) (van der Waals) materials represent new, adjustable quantum platforms for realizing a myriad of novel phases [12–14]. While bilayers of electronic materials have widely been explored, investigations of bilayers of frustrated quantum magnets and magnetic moiré superlattices are still in their early stages [15–23]. Considering bilayers of the Kitaev’s honeycomb spin liquid [9], we note that generic interlayer interactions spoil the integrability of the Kitaev model in each layer [24,25], and the resulting model is no longer exactly solvable. Instead, one can turn to perturbative expansions, starting in solvable limits, perform mean-field treatments [24,26], or use numerical methods

such as exact diagonalization [27,28] to estimate its phase diagram. In contrast,  $\Gamma$ -matrix generalizations of the Kitaev model (with larger local Hilbert spaces) [29–32] allow for interlayer exchange terms that commute with the intralayer fluxes, making controlled calculations feasible. Yet, the lack of candidate materials for these models is a significant challenge. It is worth noting that prior research has predominantly focused on bilayer Kitaev models with Heisenberg interlayer interactions, which stabilize a trivial quantum paramagnet for large interlayer interactions, consisting of interlayer singlets [24,26–28].

Instead, in this paper, we focus on the  $S = \frac{1}{2}$  bilayer Kitaev model with *Ising* interlayer interactions. Unlike a Heisenberg interlayer interaction, this interaction retains a residual degree of freedom in the limit of large interlayer exchange couplings. This opens up the possibility for nontrivial phases in this limit, in particular one may wonder if topologically ordered states or magnetic phases are realized. In particular, the coexistence of topological and magnetic order could give rise to a spontaneously generated chiral spin liquid. In principle, there exists an arbitrariness to fixing the spin-space axis of the Ising interlayer. We note that varying this axis and different choices for the relative sign of the Kitaev couplings adds layers of complexity to our investigation, providing the means for exploration of rich phase diagrams and emergent phenomena.

To construct the phase diagram of the model, we first focus on deriving effective Hamiltonians in the limit of large interlayer exchange interactions. This allows us to determine the ground state in this limit, such as ferromagnetic (FM) or antiferromagnetic (AFM) order. Equipped with these controlled insights, we perform Majorana mean-field theory to determine

the phase at weak and intermediate interlayer exchange, where we use magnetically ordered states as variational *Ansätze*. We emphasize that by construction, the mean-field theory exactly reproduces the  $T = 0$  ground state of the Kitaev model (i.e., in the lowest flux sector) and is thus controlled in *both* limits of vanishing and strong interlayer interactions.

Our main results can be summarized as follows: (i) when the Ising interaction points along  $(n^x, n^y, n^z)$  with all  $n^\alpha \neq 0$ , and both layers have the same Kitaev interaction strength ( $K_1 = K_2$ ), there is a first-order transition from a  $\mathbb{Z}_2 \times \mathbb{Z}_2$  spin-liquid state to a FM or AFM state, depending on the sign of the Kitaev interaction. (ii) For  $K_1 = -K_2$ , the magnetic order is suppressed and the spin-liquid phase is sustained for fairly large interlayer couplings. Beyond a critical  $J/|K|$ , within mean-field theory we find that a gapped bilayer  $\mathbb{Z}_2$  spin liquid emerges, locking the gauge structure of the two layers. This phase then undergoes a confinement-deconfinement transition for larger  $J/|K|$ . Perturbatively, we show that the large interlayer coupling limit of the confined phase is determined by the  $120^\circ$  compass model for the effective degrees of freedom. (iii) In cases when  $\mathbf{n}$  is along a Cartesian axis such as the  $\hat{z}$  direction, or perpendicular to it (i.e.,  $\mathbf{n}$  lies in the  $x$ - $y$  plane), our perturbative analysis shows the existence of Ising chain with a twofold ground-state degeneracy per chain. We find that the splitting of this twofold ground-state degeneracy by interchain couplings is exponentially small in the length of the chains and, therefore, surprisingly, the system possesses a subextensive ground-state degeneracy (given by effectively decoupled chains) in the thermodynamic limit.

The rest of the paper is organized as follows. In Sec. II, we introduce the model and describe our methodology, including the perturbative analysis and Majorana mean-field theory. In Sec. III, we present our results for different parameter regimes. We conclude with a discussion and a summary of our results in Sec. IV.

## II. MODEL AND METHODS

### A. Microscopic model

The Kitaev honeycomb model [9] is a paradigmatic example of a highly frustrated  $S = \frac{1}{2}$  model characterized by bond-dependent interactions. Within this model, interactions are defined along three distinct types of bonds originating from each lattice site within the honeycomb lattice, which we denote using the symbols  $\alpha = x, y, z$  as shown in Figs. 1(a) and 1(b). For the bilayer Kitaev model, we consider an AA stacking configuration, where the A sublattice of the first layer is precisely aligned with the A sublattice of the second layer. The interaction between these two layers is governed by an Ising-type interaction oriented along a specific axis in spin space, characterized by a unit vector  $\mathbf{n}$ , which henceforth will be referred to as the Ising axis. The full Hamiltonian is  $H = H_K + H_J$ ,

$$H_K = \sum_{\nu, (ij)^\alpha} K_\nu S_{\nu i}^\alpha S_{\nu j}^\alpha, \quad (1)$$

$$H_J = -J \sum_i (\mathbf{n} \cdot \mathbf{S}_{1i})(\mathbf{n} \cdot \mathbf{S}_{2i}), \quad (2)$$

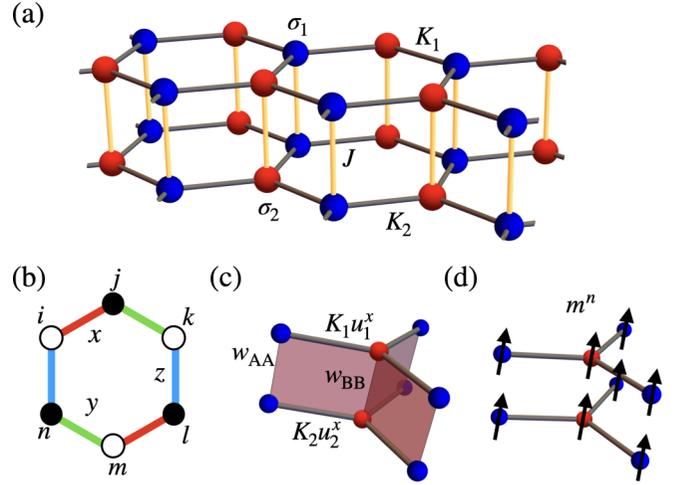


FIG. 1. (a) Illustration of the bilayer Kitaev-Ising model.  $K$  and  $J$  are intralayer Kitaev and interlayer Ising interaction exchange terms, respectively. (b) The bond-dependent interactions of the Kitaev model: red, green, and blue bonds represent the  $x$ ,  $y$ , and  $z$  bonds, respectively. We observe two types of phase transitions: (c) topological phase transitions where the interlayer plaquettes acquire finite expectation value that gap the spectrum, and (d) magnetic order induced by the interlayer interaction or an external magnetic field.

where  $\nu = 1, 2$  is the layer index and  $\mathbf{n}$  is a vector on the unit sphere. Before delving into the bilayer model, we first briefly review the solution of the single-layer Kitaev model following Ref. [9]. The key observation which leads to the exact solvability is based on the plaquette operators  $W_p = \sigma_i^y \sigma_j^z \sigma_k^x \sigma_l^y \sigma_m^z \sigma_n^x$ . These operators commute with the Hamiltonian, and hence the whole Hilbert space can be labeled by the eigenvalues of the plaquette operators. Equation (1) can be solved by representing the spin operators at each site by four Majorana fermions  $2S_i^\alpha = i\chi_i^\alpha \chi_i^0$  where we choose the normalization  $(\chi^\mu)^2 = 1$ . The Majorana representation is overcomplete and the physical Hilbert space can be recovered by projecting states with the operator  $P = \prod_i (1 + D_i)/2$  where  $D_i = \chi_i^0 \chi_i^x \chi_i^y \chi_i^z$ , which enforces that the fermion parity on each site is even,  $D_i \equiv +1$ . Using the Majorana representation, the Kitaev Hamiltonian can then be written as

$$\begin{aligned} H_K &= \frac{K}{4} \sum_{(ij)^\alpha} (i\chi_i^\alpha \chi_i^0)(i\chi_j^\alpha \chi_j^0) \\ &\equiv \frac{K}{4} \sum_{(ij)^\alpha} iu_{ij} \chi_i^0 \chi_j^0, \end{aligned} \quad (3)$$

where in the second line we have introduced  $u_{ij}^\alpha = i\chi_i^\alpha \chi_j^\alpha$ . Notably, both  $\chi_j^\mu$  and  $u_{ij}$  anticommute with the constraint operator  $D_i$ , and thus it becomes clear that the Majorana fermions carry a  $\mathbb{Z}_2$  gauge charge and are coupled to a  $\mathbb{Z}_2$  gauge field given by  $u_{ij}$ , with gauge transformations generated by  $D_i$ .

The plaquette operators can be represented by the product of the bond operators  $W_p = \prod_p u_{ij}$ , corresponding to gauge-invariant Wilson loops in the  $\mathbb{Z}_2$  gauge theory. Given that the  $W_p$  are conserved, the physical Hilbert space decomposes into distinct sectors labeled by the eigenvalues of  $W_p$ . According to

Lieb's theorem [33], the ground state of the Kitaev model lies in the zero-flux sector with all plaquette operators having the eigenvalue  $W_p = 1$ . In this sector, the Majorana fermion dispersion is gapless, and possesses two Majorana-Dirac cones. For a bilayer system with vanishing interlayer couplings  $J = 0$ , there are two copies of gapless  $\mathbb{Z}_2$  QSLs, resulting in a  $\mathbb{Z}_2 \times \mathbb{Z}_2$  phase.

### B. The limit of large interlayer exchange

In the atomic limit with  $K_v = 0$  and  $J \neq 0$ , the effective degrees of freedom are determined by the Ising interlayer interaction. The ground state is a doublet given by  $|\uparrow_1^n \uparrow_2^n\rangle$  and  $|\downarrow_1^n \downarrow_2^n\rangle$  where  $|\uparrow^n\rangle$  and  $|\downarrow^n\rangle$  are eigenstates of  $(\mathbf{n} \cdot \mathbf{S})$ , that is the spin operator aligned to the  $\mathbf{n}$  Ising axis. The excited states also form a doublet,  $|\uparrow_1^n \downarrow_2^n\rangle$  and  $|\downarrow_1^n \uparrow_2^n\rangle$ . For later convenience, we rotate the axis of quantization of the Pauli matrices such that the rotated  $\hat{z}$  axis point along  $\mathbf{n}$ . To achieve this, we choose the axis of rotation and the angle to be  $\mathbf{k} = \mathbf{n} \times \hat{z}$  and  $\theta = \cos^{-1}(n^z)$ . Next, we use the operator  $\exp(i\theta/2\mathbf{k} \cdot \boldsymbol{\sigma})$ , to rotate each spin matrix at every site along the desired axis using the relation  $e^{-i\boldsymbol{\sigma} \cdot \hat{\mathbf{k}}\theta/2} \mathbf{a} \cdot \boldsymbol{\sigma} e^{i\boldsymbol{\sigma} \cdot \hat{\mathbf{k}}\theta/2} = [\hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \mathbf{a}) + \cos(\theta)[\mathbf{a} - \hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \mathbf{a})] + \sin(\theta)\hat{\mathbf{k}} \times \mathbf{a}] \cdot \boldsymbol{\sigma}$ . This rotation maps  $(\mathbf{n} \cdot \mathbf{S}) \rightarrow \tilde{S}^z$ , and then the interlayer interaction can be written as  $H_J = -J \sum_i \tilde{S}_{1i}^z \tilde{S}_{2i}^z$ .

In the following, we will derive effective Hamiltonians within the degenerate ground-state manifold spanned by degenerate doublets on each site. To this end, it will be convenient to introduce pseudospin operators for each interlayer pair of sites. These span a full operator basis for each local ground-state doublet,

$$\begin{aligned} \eta^z &= \frac{1}{2}(\tilde{S}_{1i}^z + \tilde{S}_{2i}^z), \\ \eta^x &= \frac{1}{4}(\tilde{S}_{1i}^x \tilde{S}_{2i}^x - \tilde{S}_{1i}^y \tilde{S}_{2i}^y), \\ \eta^y &= \frac{1}{4}(\tilde{S}_{1i}^x \tilde{S}_{2i}^y + \tilde{S}_{1i}^y \tilde{S}_{2i}^x). \end{aligned} \quad (4)$$

These pseudospin operators satisfy the SU(2) algebra. Note that  $\eta^z$  is a dipolar operator while  $\eta^x$  and  $\eta^y$  are quadrupolar operators [30]. If the sign of  $J$  is flipped from positive to negative, the pseudospin operators need to be redefined as the ground-state sector will then be spanned by  $|\uparrow_1^n \downarrow_2^n\rangle$  and  $|\downarrow_1^n \uparrow_2^n\rangle$ . The effective Hamiltonian acting on this degenerate subspace, obtained via perturbation theory in the large  $J/|K|$  limit, can be expressed using these operators [30]. For instance, the first- and second-order contributions to the effective Hamiltonian are derived as

$$\begin{aligned} H_{\text{eff}}^{(1)} &= P_0 H_K P_0, \\ H_{\text{eff}}^{(2)} &= P_0 H_K S H_K P_0, \end{aligned} \quad (5)$$

where we use projection operator (in the rotated basis)  $P_0 = \prod_i (1 + 4\tilde{S}_{2i}^z)$  onto the low-energy manifold, and  $S = (1 - P_0)/(E_0 - H_J)$ . We stop at the order of perturbation when the effective Hamiltonian exhibits nontrivial magnetic order. If  $H_{\text{eff}}$  has a simple form (i.e., without frustrated interactions), the ground state in the  $J/K \gg 1$  limit can then be readily obtained. We will use the thus-obtained magnetically ordered states as *Ansätze* in our Majorana mean-field theory calculations to explore the weak and intermediate  $J/K$  regions.

### C. Majorana mean-field theory

In the presence of interlayer interactions, the single-layer Kitaev model as detailed in Sec. II A is no longer solvable, as the plaquette operators are no longer conserved,  $[H_J, W_p] \neq 0$ . To map out phase diagrams, we therefore resort to Majorana mean-field theory (MMFT) for the full bilayer system [34]. In the following, we also incorporate an onsite external magnetic field into the Hamiltonian, which will find utility in specific sections of our analysis.

Within MMFT, we do not enforce the constraint  $D_i = +1$  for each site (which would require significant numerical effort, e.g., using Gutzwiller-projected variational Monte Carlo methods), but instead enforce the constraint on average. To this end, we reformulate  $D_i = 1$  as  $i\chi^\alpha \chi^0 + \frac{i}{2}\epsilon^{\alpha\beta\gamma} \chi^\beta \chi^\delta = 0$  and subsequently enforce it through the introduction of a Lagrange multiplier, as detailed in Refs. [34,35].

To facilitate the analysis, we employ a mean-field approximation to decouple intralayer Majorana fermion interactions as  $i\chi_{i,A}^\alpha \chi_{i,A}^0 i\chi_{j,B}^\alpha \chi_{j,B}^0 \approx m_A^\alpha (i\chi_{j,B}^\alpha \chi_{j,B}^0) + m_B^\alpha (i\chi_{i,A}^\alpha \chi_{i,A}^0) - m_A^\alpha m_B^\alpha - u^\alpha (i\chi_{i,A}^0 \chi_{j,B}^0) - u^0 (i\chi_{i,A}^\alpha \chi_{j,B}^\alpha) + u^\alpha u^0$ , with the mean-field parameters  $u^0 = \langle i\chi_{i,A}^0 \chi_{j,B}^0 \rangle$ ,  $m_A^\alpha = \langle i\chi_{i,A}^\alpha \chi_{i,A}^0 \rangle$ , and  $m_B^\alpha = \langle i\chi_{j,B}^\alpha \chi_{j,B}^0 \rangle$ . The interlayer interaction  $H_J = -\frac{J}{4} \sum_i (i\chi_{1i}^{(\mathbf{n})} \chi_{1i}^0)(i\chi_{2i}^{(\mathbf{n})} \chi_{2i}^0)$ , where  $\chi^{(\mathbf{n})} = \sum_\alpha n^\alpha \chi^\alpha$  is decoupled as  $i\chi_{1i}^{(\mathbf{n})} \chi_{1i}^0 i\chi_{2i}^{(\mathbf{n})} \chi_{2i}^0 \approx w_i^0 (i\chi_{1i}^{(\mathbf{n})} \chi_{2i}^{(\mathbf{n})}) + w_i^{(\mathbf{n})} (i\chi_{1i}^0 \chi_{2i}^0) - w_i^{(\mathbf{n})} w_i^0 - m_{1i}^{(\mathbf{n})} (i\chi_{2i}^{(\mathbf{n})} \chi_{2i}^0) - m_{2i}^{(\mathbf{n})} (i\chi_{1i}^{(\mathbf{n})} \chi_{1i}^0) + m_{1i}^{(\mathbf{n})} m_{2i}^{(\mathbf{n})}$  with  $w_i^0 = \langle i\chi_{1i}^0 \chi_{2i}^0 \rangle$ ,  $w_i^{(\mathbf{n})} = \langle i\chi_{1i}^{(\mathbf{n})} \chi_{2i}^{(\mathbf{n})} \rangle$  denoting mean fields in the Hartree channel, while  $m_{1i}^{(\mathbf{n})} = \langle i\chi_{1i}^{(\mathbf{n})} \chi_{1i}^0 \rangle$  and  $m_{2i}^{(\mathbf{n})} = \langle i\chi_{2i}^{(\mathbf{n})} \chi_{2i}^0 \rangle$  is the decoupling in the magnetic channel. Note that  $m^{(\mathbf{n})}$  is the magnetization along the direction of the  $\mathbf{n}$  axis and  $m^\alpha$  is the magnetization along  $x$ ,  $y$ ,  $z$  axes. Note that here we also considered the mean-field channels of the form  $\langle i\chi_{i,A}^\alpha \chi_{j,B}^0 \rangle$  and  $\langle i\chi_{1i}^\alpha \chi_{2i}^0 \rangle$  which correspond to spin-polarized intralayer and interlayer Majorana hopping. However, these channels never attained any finite expectation values in our calculations. Thus, for brevity, we drop them from our mean-field Hamiltonian. Incorporating all the other mean-field parameters, we write the full mean-field Hamiltonian as

$$\begin{aligned} H &= \sum_{v,i} \sum_{\alpha} \frac{1}{2} \left( \frac{K_v}{2} m_B - h^\alpha - \lambda^\alpha \right) i\chi_{vi,A}^\alpha \chi_{vi,A}^0 + \frac{1}{2} \left( \frac{K_v}{2} m_A - h^\alpha - \lambda^\alpha \right) i\chi_{v,j,B}^\alpha \chi_{v,j,B}^0 - \frac{K_v u^\alpha}{4} (i\chi_{vi,A}^0 \chi_{vj,B}^0) \\ &\quad - \frac{K_v u^0}{4} (i\chi_{vi,A}^\alpha \chi_{vj,B}^\alpha) - \lambda^\alpha \frac{\epsilon^{\alpha\beta\gamma}}{4} (i\chi_{vi,A}^\beta \chi_{vi,A}^\gamma + i\chi_{vj,B}^\beta \chi_{vj,B}^\gamma) \\ &\quad - \frac{J}{4} \sum_i w_i^0 (i\chi_{1i}^{(\mathbf{n})} \chi_{2i}^{(\mathbf{n})}) + w_i^{(\mathbf{n})} (i\chi_{1i}^0 \chi_{2i}^0) - m_{1i}^{(\mathbf{n})} (i\chi_{2i}^{(\mathbf{n})} \chi_{2i}^0) - m_{2i}^{(\mathbf{n})} (i\chi_{1i}^{(\mathbf{n})} \chi_{1i}^0) + E_{\text{const}}[m, u, w], \end{aligned} \quad (6)$$

where in total eight independent mean-field parameters ( $u^0, u^\alpha, w^0, w^{(\mathbf{n})}, m_A, m_B$ ) with  $\alpha = x, y, z$  and three Lagrange multipliers  $\lambda^\alpha$  are to be determined self-consistently.  $E_{\text{const}}[m, u, w]$  is a constant term that depends on the mean-field parameters. We use an iterative procedure to solve the mean-field self-consistency equations and determine the Lagrange multipliers, where we diagonalize Eq. (6) on momentum-space grids of  $4 \times 10^4$  points.

As discussed in previous works, the mean-field decoupling of the single-layer Kitaev interaction in Eq. (6) can be seen to exactly reproduce static spin-spin correlations and the spectrum of the itinerant Majorana fermions in the 0-flux ground-state sector [36,37], where intuitively the mean-field parameter  $u^\alpha$  can be identified with a (gauge-fixed) configuration of the gauge field  $u_{ij}$  in Eq. (3).

Next, we present the results obtained using the methods above for various possibilities of  $\mathbf{n}$  and the relative sign of the Kitaev interactions in the two layers.

### III. RESULTS

#### A. Arbitrary Ising axis with same Kitaev interaction

$$(n^\alpha \neq 0, K_1 = K_2)$$

We first consider the case where the Ising interaction has components along all Cartesian coordinates  $\mathbf{n} = (n^x, n^y, n^z)$ , with  $n^\alpha \neq 0$ . We proceed according to the method described in the previous section, and first derive an effective Hamiltonian for  $J/K \gg 1$  via perturbative expansion. First-order perturbation theory leads to

$$H_{\text{eff}}^{(1)} = \sum_{(ij)^\alpha} U^\alpha \eta_i^z \eta_j^z, \quad (7)$$

where  $U^\alpha = (K_1 + K_2)(n^\alpha)^2/2$ . For the isotropic direction,  $\mathbf{n} = (1, 1, 1)/\sqrt{3}$ , and  $K_1 = K_2 = K$  we obtain  $U^\alpha = K/3$  for all bonds. Equation (7) suggests the ground state exhibits FM or AFM long-range order depending on the sign of  $K$ . It is noteworthy that highly frustrated Kitaev interactions lead to a simple, nonfrustrated effective model in this limit with a straightforward AFM/FM ground state aligned along the Ising axis.

Next, we perform Majorana mean-field theory calculations to explore the intermediate- $J$  region. We begin with solving the mean-field Hamiltonian in Eq. (6), with no external field. We find a transition from the  $\mathbb{Z}_2 \times \mathbb{Z}_2$  gapless spin liquid, which is characterized by a vanishing magnetization and no interlayer Hartree channel ( $w^{(\mathbf{n})}$  and  $w^0$ ), to a fully polarized state with a uniform magnetization  $m^{(\mathbf{n})} = 1$ .

This holds for both FM or AFM Kitaev interactions. We find that in this parameter regime, the results of our iterative numerical procedure for finding self-consistent solutions to the mean-field equations heavily depend on the initial conditions for the iteration, in particular on an *Ansatz* for systems' magnetization. This suggests that the transition is first-order phase transition, between two (locally stable) competing saddle points. Thus, to pinpoint the exact value of  $J_c/K$ , we compare the energies of the  $\mathbb{Z}_2 \times \mathbb{Z}_2$  gapless spin liquid and the fully polarized state and find that the energies intersect at  $J_c/K = 0.55$  as shown in Fig. 2. This demonstrates that, based

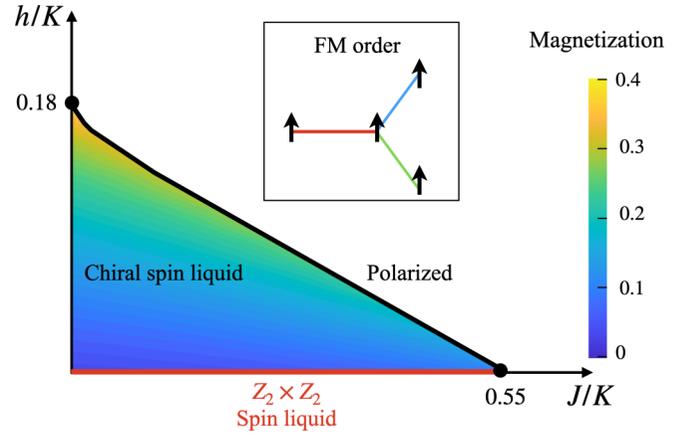


FIG. 2. Phase diagram for  $\mathbf{n} = (1, 1, 1)/\sqrt{3}$  and  $K_1 = K_2$ . For both FM and AFM Kitaev interactions, the  $\mathbb{Z}_2 \times \mathbb{Z}_2$  gapless spin liquid (red line) undergoes a first-order transition to a polarized phase as  $J_c/K = 0.55$ . For FM Kitaev interaction, external magnetic field  $h$  along the [111] direction lowers the critical  $J$ . External magnetic field induces a finite magnetization, which is shown with the color coding. This phase is a gapped chiral spin liquid.

on our mean-field analysis, we do not expect a phase that simultaneously exhibits local magnetic order and topological order.

Focussing on the case of FM Kitaev interactions, we consider the impact of a magnetic field in the [111] direction. In the absence of interlayer interactions ( $J = 0$ ), we obtain a chiral spin liquid up to  $h_c/K = 0.18$ , in agreement with Ref. [35]. With the inclusion of interlayer couplings,  $h_c$  diminishes, as expected since the FM interlayer exchange functions similar to magnetic field at mean-field level, leading to a higher effective magnetic field experienced by each layer. We also observe that if the magnetic channel is artificially turned off, the interlayer Hartree channel acquires a finite expectation value at  $J/K = 0.9$ . Given that this value surpasses the critical exchange needed for the fully polarized phase, we can infer that magnetic ordering is preferred compared to the interlayer Hartree channel.

It is important to note that Majorana mean-field calculations on the Kitaev model tend to overestimate the critical values for the destruction of the Kitaev QSL phase since they ignore the quantum fluctuations due to dynamical vortons as excitations of the  $\mathbb{Z}_2$  gauge field [28,38,39]. An appropriate treatment is an interesting direction for future research. Nevertheless, the phase diagrams of mean-field calculations and numerical approaches can be expected to be similar, with renormalized values for the critical coupling constants.

#### B. Suppressed magnetic ordering for Kitaev interaction with opposite sign ( $K_1 = -K_2$ )

Equation (7) implies that the first-order correction in the effective Hamiltonian vanishes when  $K_1 = -K_2$ . Motivated by this observation, we investigate the phase diagram for  $K_1 = -K_2 = K$  and  $\mathbf{n} = (1, 1, 1)/\sqrt{3}$ . Then, second-order perturbation theory leads to the following effective spin Hamiltonian

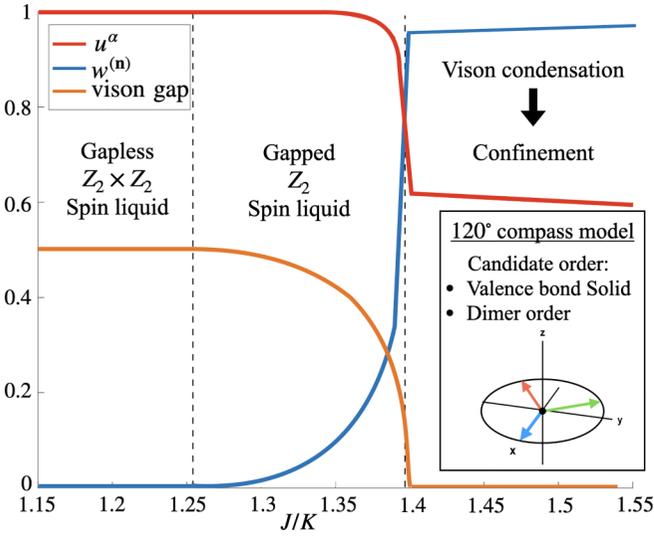


FIG. 3. Phase diagram  $K_1 = -K_2$  and  $\mathbf{n} = (1, 1, 1)/\sqrt{3}$ . At  $J/K = 1.25$ , the Hartree order parameter  $w^\mu$  acquires a finite expectation value which gaps the spectrum and locks the gauge fields on each layer. This is followed by a confinement-deconfinement transition via the condensation of visons, which occurs (in a treatment beyond mean-field theory) when the vison gap closes. Here, we take the energy gap of the  $\chi^\alpha$  bands as a proxy for the energy cost of a single vison excitation in the full interacting  $\mathbb{Z}_2$  gauge theory. Using perturbation theory, we predict that this phase, at large values of  $J/K$ , is described by the  $120^\circ$  compass model.

in the large- $J$  limit,

$$H_{\text{eff}}^{(2)} = \frac{2K^2}{|J|} \left[ \sum_{(ij)^z} \eta_i^x \eta_j^x + \sum_{(ij)^x} R_{120}^z(\eta_i^x) R_{120}^z(\eta_j^x) + \sum_{(ij)^y} R_{-120}^z(\eta_i^x) R_{-120}^z(\eta_j^x) \right], \quad (8)$$

where  $R_\theta^z(\eta^x) = \exp(i\frac{\theta}{2}\eta^z)\eta^x \exp(-i\frac{\theta}{2}\eta^z)$  is the rotation operation on the pseudospin operators about the  $\hat{z}$  axis by  $\theta = \pm 120^\circ$ . Notably, Eq. (8) is the  $120^\circ$  compass model for the  $\eta$  degrees of freedom. It is a highly frustrated model and its ground state has still not been unambiguously identified. Candidate orders include valence bond solid, long-range dimer order [40,41].

Since the ground state of the  $120^\circ$  compass model is not well established, a major reason being that the energy differences between the candidate magnetic orders are quite small, we instead use for simplicity FM and AFM (Néel order) mean-field *Ansätze* for our mean-field theory calculations:  $m_{v,A}^{(\mathbf{n})} = m_{v,B}^{(\mathbf{n})}$ , for FM and  $m_{v,A}^{(\mathbf{n})} = -m_{v,B}^{(\mathbf{n})}$  for AFM case. We find that these magnetically polarized phases exhibit higher energies compared to the  $K_1 = K_2$  case since the energy gain from the Kitaev term on each layer cancels each other due to the opposite sign. This allows for the Hartree channel order parameter  $w^\mu$  to attain a finite expectation value prior to magnetic order. Consequently, the interlayer plaquette operator, as shown in Fig. 1(c), attains a nonzero value, leading to gapped QSL with a topologically trivial band structure (i.e., Chern number  $\nu = 0$ ) at  $J_c/K = 1.25$  as shown in Fig. 3.

We now comment on the interpretation of our results beyond the mean-field treatment of the model. The mean-field Hamiltonian in Eq. (6) can be understood to constitute a particular gauge-fixed configuration of some (nonintegrable) gauge theory. Equivalence classes of such mean-field *Ansätze* which are equivalent (up to gauge transformations) can be classified with respect to their projective symmetry group (PSG) [36]. We refrain from such a full classification for the bilayer system here. However, importantly, we note that a finite  $w^\mu$  implies that independent gauge transformations on each layer no longer leave the Hamiltonian invariant, only conjoint gauge transformations do. This reduces the gauge group from  $\mathbb{Z}_2 \times \mathbb{Z}_2$  to  $\mathbb{Z}_2$  [30]. Moreover, we stress that the operators  $W_i^\mu = i\chi_{1i}^\mu \chi_{2i}^\mu$  are in general not gauge invariant, and thus the fields  $w_i^\mu$  can not be used to construct a local Landau-Ginzburg analysis for the transition out of the  $\mathbb{Z}_2 \times \mathbb{Z}_2$  spin liquid to the bilayer system with a residual  $\mathbb{Z}_2$  gauge group. Explicitly, gauge transformations induced by the operators  $D_{vi}$  change the sign of the associated  $w_i^\mu$ , in addition to the three bond operators  $u_{ij}^\alpha$  emanating from that site. Consequently,  $w_i^\mu$  vanishes for the physical wave function which is symmetrized over all gauge configurations [9]. However, it is possible to introduce a gauge-invariant correlator [29,30]

$$\langle C_{ij}^\mu \rangle = \langle W_i^\mu B_{ij} W_j^\mu \rangle, \quad (9)$$

where  $B_{ij} = \prod_{(i'j')} \text{sgn}(u_{i'j'}^\mu) \text{sgn}(u_{2i'j'}^\mu)$  is the product of the signs of the  $u_{ij}^\mu$  operators that connect the two  $W_{i'j'}$  operators. The value of  $\langle C_{ij}^\mu \rangle$  is the same in all gauge choices. Therefore, it is also finite for the physical wave function. Finite  $w_{ij}^\mu \neq 0$  implies  $\langle C_{ij}^\mu \rangle \neq 0$ , signaling a nonlocal string order parameter.

For larger values of interlayer exchange, we observe that the energy gap of  $\chi^\alpha$  bands vanishes as shown in Fig. 3. These bands are associated with the Majorana fermions of flavor  $\alpha$  that are localized on the  $\alpha$  bonds in the pure Kitaev limit, which in the exact solution give rise to the  $\mathbb{Z}_2$  gauge field [compare also Eq. (3)]. While the vison in Kitaev's exact solution is a *nonlocal* excitation of the  $\mathbb{Z}_2$  gauge field, the delocalization of the  $\alpha$  Majoranas (i.e., dispersive bands) can be taken as a proxy for the dynamics of the visons that is induced by breaking integrability, and we therefore (loosely) associate the gap of the  $\chi^\alpha$ -Majorana fermion dispersion with the gap of dispersing visons in the full (nonintegrable)  $\mathbb{Z}_2$  gauge theory. Equipped with this understanding, we suggest that the  $\chi^\alpha$  Majoranas becoming gapless can be interpreted as the single-vison gap closing, which allows for the condensation of visons, tantamount to a confinement-deconfinement transition [42,43]. From our mean-field computations, we find a critical coupling of approximately  $J/K \simeq 1.4$ . The resulting state will be accurately described by the  $120^\circ$  compass model, as presented in Eq. (8), for which previous studies have identified nonfractionalized states with magnetic and valence bond solid (VBS) ordering as possible ground states.

### C. Special cases for the Ising axis

The first-order correction to the effective Hamiltonian in Eq. (7) also becomes suppressed if the Ising axis is oriented such that  $n^\alpha$  ( $\alpha = x, y, z$ ) vanishes for certain bonds. Unlike the ( $K_1 = -K_2$ ) case in Sec. III B, where  $H_{\text{eff}}^{(1)}$  vanishes entirely, orienting the  $\mathbf{n}$  such that  $n^\alpha = 0$  for particular Cartesian

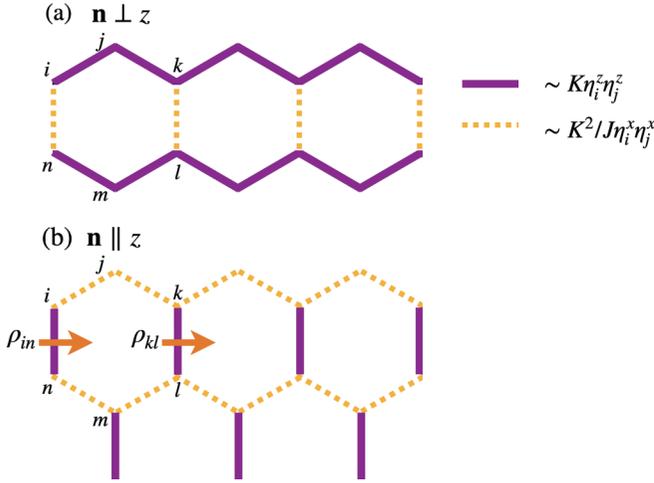


FIG. 4. Depiction of the  $H_{\text{eff}}^{(2)}$  for special cases of the Ising axis. Purple and dashed orange lines represent first- and second-order terms in the effective Hamiltonian. (a) For the  $\mathbf{n} = [1, 1, 0]/\sqrt{2}$ , the first-order correction forms Ising chains along  $x$  and  $y$ . These chains remain decoupled within perturbation theory. (b) For  $\mathbf{n} = [0, 0, 1]$ , the first-order correction leads to formation of Ising dimers. These dimers couple to form chains in fifth order in perturbation theory. Once again, the chains remain decoupled perturbatively, leading to subextensive degeneracy in both cases.

axes only suppresses the bonds along the  $\alpha$  directions. To investigate the consequences of these interactions, we consider two cases, where  $n^\alpha = 0$  for one and two Cartesian axes, respectively, below.

### 1. Effective chain geometry for $\mathbf{n} = (1, 1, 0)/\sqrt{2}$

We first consider the case when a single  $n^\alpha$  vanishes. We pick  $\mathbf{n} = (1, 1, 0)/\sqrt{2}$ , which preserves the symmetry between the  $x$  and  $y$  bonds, but the first-order correction the energy along the  $z$  bond vanishes. We obtain the following effective Hamiltonian up to second order in perturbation expansion:

$$H_{\text{eff}}^{(1)} = K \sum_{(ij)^{x/y}} \eta_i^z \eta_j^z,$$

$$H_{\text{eff}}^{(2)} = \frac{2K^2}{|J|} \sum_{(ij)^{x/y}} \eta_i^x \eta_j^x \quad (x/y \text{ bonds}), \quad (10)$$

$$H_{\text{eff}}^{(2)} = \frac{2K^2}{|J|} \sum_{(ij)^z} \eta_i^x \eta_j^x \quad (z \text{ bonds}). \quad (11)$$

Equation (10) leads to the formation of chains along  $x/y$  bonds, coupled along the Ising axis [as depicted in Fig. 4(a)]. This is the largest interaction in the perturbation theory  $O(K)$ , and at this order, each chain exhibits two degenerate ground states. Meanwhile, at each lattice site, the spins along a chain interact with spins on adjacent chains in the transverse direction in spin space, with a notably diminished interaction strength on the order of  $O(K^2/J)$ . Considering the two adjacent Ising chains, a single  $H_{\text{eff}}^{(2)}$  bond flips two spins and therefore takes the state outside the ground-state manifold of Eq. (10). Consequently, the interchain interactions in  $H_{\text{eff}}^{(2)}$  do

not split the degeneracy between different chains in leading order  $K/|J|$ .

In order to determine if there are higher-order contributions to  $H_{\text{eff}}$  which lift the degeneracy, we perform exact diagonalization on a 12-site system, which is a single hexagon on both layers. We extract the following effective Hamiltonian:

$$H_{\text{eff}}^{\text{ED}} = \sum [c_1 K^3 / J^2 (\eta_i^x \eta_j^x \eta_k^x + \eta_i^x \eta_m^x \eta_n^x) + c_2 K^6 / |J|^5 (\eta_i^x \eta_j^x \eta_k^x \eta_l^x \eta_m^x \eta_n^x)], \quad (12)$$

where  $c_1 \approx 10^{-2}$  and  $c_2 \approx 10^{-5}$ . The details of this calculation are given in Appendix. While the second term involves interactions between spins on different chains, it flips three bonds on each chain, and therefore takes the chains outside their ground-state manifold determined by Eq. (10), similar to  $H_{\text{eff}}^{(2)}$ .

Next, we argue that the degeneracy between distinct chains, determined by  $H_{\text{eff}}^{(1)}$ , remains at arbitrarily high order when including the effects of interchain interactions in  $H_{\text{eff}}^{(2)}$  perturbatively in  $K/J \ll 1$ . To this end, we denote the two degenerate Ising ground states of a chain according to  $H_{\text{eff}}^{(1)}$  as  $|\uparrow\rangle$  ( $|\downarrow\rangle$ ) =  $\prod_i |\uparrow\rangle$  ( $|\downarrow\rangle$ ). Considering two chains, labeled  $t$  and  $b$ , interactions lift the fourfold ground-state degeneracy if there exists some nontrivial Hamiltonian  $\tilde{H}_{\text{eff}}$  acting on  $|\uparrow_t \uparrow_b\rangle, \dots, |\downarrow_t \downarrow_b\rangle$ . We first note that symmetry strongly constrains the form of  $\tilde{H}_{\text{eff}}$ : Performing a  $\pi$  rotation about the  $x$  axis of the spins along a given chain  $U = \exp(-i\frac{\pi}{2} \sum_{vi} \sigma_{vi}^x)$  flips the spins from  $|\uparrow\rangle \rightarrow |\downarrow\rangle$  and vice versa, but commutes both with Eq. (11) and any effective Hamiltonian  $H_{\text{eff}}^{(n)} = P_0 V S V S \dots S V P_0$  obtained at arbitrarily high order in perturbation theory. This implies that all diagonal matrix elements of  $\tilde{H}_{\text{eff}}$  must be identical to any order in perturbation theory,  $\langle \uparrow_t \uparrow_b | \tilde{H}_{\text{eff}} | \uparrow_t \uparrow_b \rangle = \langle \uparrow_t \downarrow_b | \tilde{H}_{\text{eff}} | \uparrow_t \downarrow_b \rangle = \dots$ , and similarly all off-diagonal matrix elements must be identical (and real),  $\langle \uparrow_t \uparrow_b | \tilde{H}_{\text{eff}} | \downarrow_t \downarrow_b \rangle = \langle \uparrow_t \downarrow_b | \tilde{H}_{\text{eff}} | \downarrow_t \uparrow_b \rangle$ . Crucially, this implies that  $\tilde{H}_{\text{eff}}$  becomes trivial if these off-diagonal matrix elements vanish. These off-diagonal elements only emerge at order approximately  $L = \sqrt{N}$  (length of a chain) in perturbation theory in  $K/J$  since tunneling  $|\uparrow\rangle \rightarrow |\downarrow\rangle$  requires flipping all spins of a given chain, and  $H_{\text{eff}}^{(n)}$  consists of *local* interactions. This implies that  $\langle \uparrow_t \uparrow_b | \tilde{H}_{\text{eff}} | \downarrow_t \downarrow_b \rangle \sim (K/|J|)^L \Delta^{-L}$ , where  $\Delta > 0$  is characteristic dimensionless energy difference between the ground state and excited states. Importantly, this implies that such off-diagonal matrix elements are exponentially suppressed with the length of the chains, and in the thermodynamic limit  $L \rightarrow \infty$ , these chains are effectively uncoupled. We therefore conclude that the ground state has a subextensive degeneracy  $O(\sqrt{N})$ , consisting of  $\sim 2^{\sqrt{N}}$  states corresponding to a twofold degree of freedom per chain. Note that our arguments are only valid in the perturbative limit and will eventually break down for  $K/J \not\ll 1$ . Similar states are also obtained in bilayer Kitaev model with Heisenberg interaction for different stacking orders and can be referred to as ‘‘classical’’ spin liquids [24], formed by Ising ‘‘macrospins’’ corresponding to the twofold-degenerate chains.

## 2. Coupled dimers for $\mathbf{n} = \hat{z}$

For  $\mathbf{n} = [0, 0, 1]$ , the first-order contribution for both  $x$  and  $y$  bonds vanishes. We obtain the following effective Hamiltonian:

$$H_{\text{eff}}^{(1)} = K \sum_{\langle ij \rangle^z} \eta_i^z \eta_j^z \quad (z \text{ bonds}), \quad (13)$$

$$H_{\text{eff}}^{(2)} = \frac{2K^2}{J} \sum_{\langle ij \rangle^{x/y}} \eta_i^x \eta_j^x \quad (x/y \text{ bonds}). \quad (14)$$

Note that there are no second- or higher-order contribution on the  $z$  bonds in this case since  $[H_K, P_0] = 0$ , which implies that the higher-order contributions in the perturbation theory vanish as  $(1 - P_0)H_K P_0$  type terms are identically zero. The  $H_{\text{eff}}^{(1)}$  forms Ising dimers [see Fig. 4(b)] such that the spins along  $z$  bonds are ‘‘locked’’ along the  $\hat{z}$  axis, which forms a doublet. A single  $H_{\text{eff}}^{(2)}$  bond acting on these dimers flips two spins, thereby breaking the Ising dimers. The doublet operators can be expressed as a pseudospin in terms of the  $\eta$  degrees of freedom,

$$\begin{aligned} \rho_{in}^z &= \frac{1}{2}(\eta_i^z + \eta_n^z), \\ \rho_{in}^\pm &= \eta_i^\pm \eta_n^\pm, \end{aligned} \quad (15)$$

where  $\rho^z$  is a dipolar and  $\rho^x$  and  $\rho^y$  are octupolar operators. In terms of the new degrees of freedom, the ground state of Eq. (13) is given by the eigenstates of  $\rho_{in}^z$ . In order to determine if the dimers are coupled via higher-order processes, we treat  $H_{\text{eff}}^{(2)}$  on the  $x/y$  bonds as a perturbation on the ground state. We obtain a nonzero contribution involving all four  $x/y$  bonds which can be expressed as a ring-exchange term

$$\begin{aligned} H_{\text{eff}}^{\text{ring}} &= P_0 H_{g2} S H_{g2} S H_{g2} S H_{g2} P_0 \\ &= \frac{2K^5}{J^4} \sum_{\square} P_0 (\eta_i^x \eta_k^x \eta_l^x \eta_n^x) P_0, \end{aligned} \quad (16)$$

where  $P_0 = \prod_{\langle ij \rangle^z} (1 + \eta_i^z \eta_j^z)/2$  and the sum over all the hexagons. In terms of the new pseudospin degrees of freedom, Eq. (16) can be expressed as

$$H_{\text{eff}}^{\text{ring}} = \frac{2K^5}{J^4} \sum_{\square} [\rho_{in}^x \rho_{kl}^x], \quad (17)$$

where  $\langle in \rangle$  and  $\langle kl \rangle$  are the two  $z$  bonds belonging to the ring. The ring-exchange term couples the dimer degrees of freedom along the  $x$  direction and once again forms chains for the octupolar degrees of freedom  $\rho^x$ . We also conducted an exact diagonalization study on a 16-site lattice, which included a central hexagonal region, along with two additional  $z$ -bond connections (see Fig. 4) which agrees with the splitting due to Eq. (17) and indicates no further splitting.

Similar to the previous subsection, here we argue that the chains remain decoupled within the perturbation theory. Considering two adjacent dimer chains, a  $\pi$  rotation about the  $x$  axis,  $U' = \exp[-i\frac{\pi}{2} \sum_{v, \langle in \rangle} (\sigma_{vi}^x + \sigma_{vn}^x)]$ , the dimers along that chain flip from  $|\uparrow\rangle \rightarrow |\downarrow\rangle$  and vice versa. Via this rotation, it is possible to map all diagonal matrix elements. The off-diagonal matrix elements require flipping all the spins on the dimer chains, leading to a vanishingly small matrix element in the thermodynamic limit.

## IV. CONCLUSIONS

In conclusion, the investigation of the phase diagram of a bilayer Kitaev honeycomb model with Ising interlayer interactions using both perturbative arguments as well as Majorana mean-field theory has yielded valuable insights into the complex interplay between topological order and magnetic tendencies in quantum spin liquids.

When the Kitaev interaction is of the same sign in both layers, we observe a first-order transition from the Kitaev spin-liquid state to a magnetically ordered state. However, when the layers have opposite signs of the Kitaev interaction, our study uncovered a higher stability of the Kitaev spin liquid. We also find that on a mean-field level, an additional intermediate gapped  $\mathbb{Z}_2$  spin-liquid state emerges, which ultimately becomes unstable for larger  $J/|K|$ , when visons are expected to condense and topological order is destroyed. The stability and nature (in particular, topological order) of this intermediate spin liquid beyond mean-field theory is an

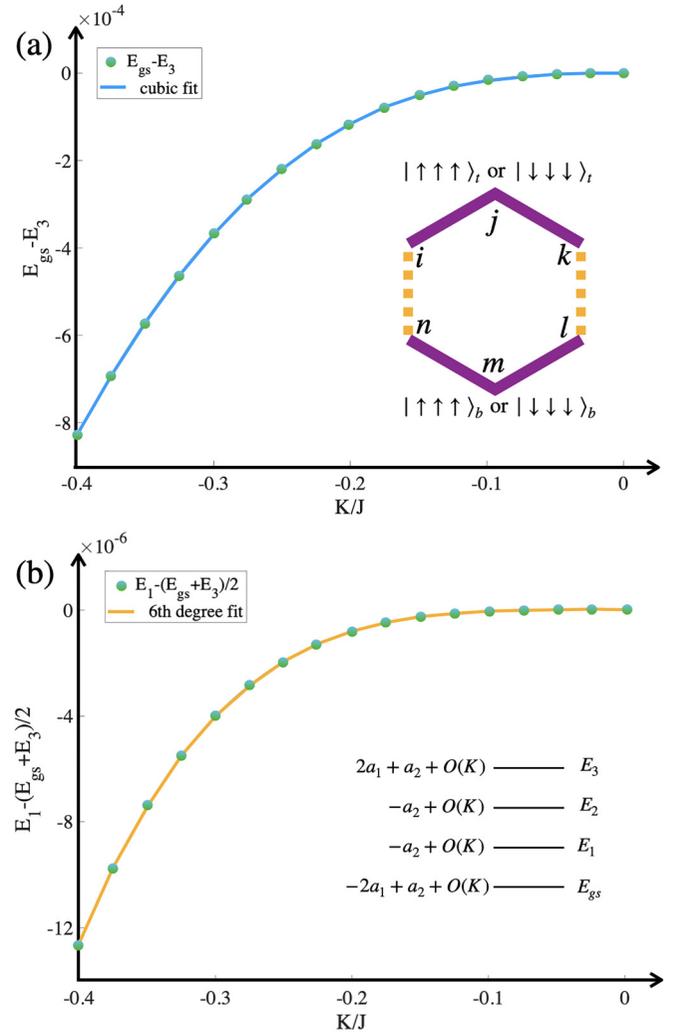


FIG. 5. (a), (b) Fitting and relevant energy differences are plotted as a function of  $K/J$ . Inset (a): Primary hexagon is denoted with the possible unperturbed ‘‘chain’’ states for top and bottom two  $x/y$  bonds. (b) Lowest 4 eigenvalues, from which coefficient data are extracted.

interesting direction for further study, e.g., using advanced numerical methods. The confined phase at large  $J/|K| \gg 1$  is aptly described by a highly frustrated  $120^\circ$  compass model.

Furthermore, we have performed perturbative analyses for the cases where the Ising axis lies along the  $\hat{z}$  axis or in the  $xy$  plane. Remarkably, in both instances, we find that 1D Ising chains that intriguingly remain decoupled within perturbation theory, and can be viewed as “macrospin” degrees of freedom. Interesting directions for future studies include exploring different stacking orders, and twisting the two layers, likely leading to a rich interplay of various orders preferred by spatially modulating stacking patterns.

### ACKNOWLEDGMENTS

We thank J. Knolle and E. Nica for fruitful discussions. A.V. and O.E. acknowledge support from NSF Award No. DMR-2234352. U.F.P.S. was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through a Walter Benjamin fellowship, Project No. 449890867, and the DOE office of BES, through Award No. DE-SC0020305. This research was supported by the National Science Foundation under Grant No. NSF PHY-1748958.

### APPENDIX: EXACT DIAGONALIZATION

#### FOR $\mathbf{n} = (1, 1, 0)/\sqrt{2}$

We describe here the exact diagonalization calculation of the effective Hamiltonian when  $\mathbf{n} = (1, 1, 0)/\sqrt{2}$ . Considering a hexagon (12 sites), there are four  $x/y$  bonds. The effective Hamiltonian (10) fixes the

spins along these bonds to be either  $|\uparrow\rangle$  or  $|\downarrow\rangle$  state (along the  $z$  axis). The ground-state manifold spans  $|\uparrow\uparrow\uparrow\rangle_t |\uparrow\uparrow\uparrow\rangle_b, |\uparrow\uparrow\uparrow\rangle_t |\downarrow\downarrow\downarrow\rangle_b, |\downarrow\downarrow\downarrow\rangle_t |\uparrow\uparrow\uparrow\rangle_b, |\downarrow\downarrow\downarrow\rangle_t |\downarrow\downarrow\downarrow\rangle_b$ , where  $t$  and  $b$  represent the “top” and “bottom” three spins (see also Fig. 5).

To find the coupling between these two segments of the Ising chains, we perform an exact diagonalization on the full Hamiltonians (1) and (2) for  $\mathbf{n} = (1, 1, 0)/\sqrt{2}$ . The four lowest eigenvalues, and the corresponding eigenvectors are extracted. In this four-dimensional subspace, we perform a rotation of basis to the ground-space basis of Eq. (10), mentioned above. This four-dimensional Hamiltonian can be written in terms of spin matrices (up to additional constants)  $\Sigma_t^\alpha$  and  $\Sigma_b^\alpha$ , where  $\alpha = x, y, z$ :

$$H_{\square}^{\text{ED}} = a_1(\Sigma_t^x + \Sigma_b^x) + a_2 \Sigma_t^x \Sigma_b^x, \quad (\text{A1})$$

where  $a_1$  and  $a_2$  are coefficients that we determine in the following steps. First, the eigenvalues of the above Hamiltonian can be written as  $E_{gs} = -2a_1 + a_2$ ,  $E_{e1} = -a_2$ ,  $E_{e2} = -a_2$ ,  $E_{e3} = 2a_1 + a_2$ . In addition, there is an  $O(K)$  term in all of these eigenvalues, from the unperturbed Hamiltonian. To extract coefficient  $a_1$ , eigenvalues  $E_{gs}$  and  $E_3$  are subtracted, and plotted as a function of  $K/J$  [Fig. 5(a)]. A cubic fit suggests that  $a_1 \approx 0.01 \frac{K^3}{J^2}$ . Similarly, for  $a_2$ , the combination  $E_1 - (E_{gs} + E_3)/2$  gets rid of the  $O(K)$  term and retains  $a_2$ . Plotting this as a function of  $K/J$  and fitting suggests a sixth-order fit with  $a_2 \approx 10^{-5} \frac{K^6}{J^5}$ .

These operators with their coefficients can be rewritten in terms of the  $\eta$  spins, as  $\Sigma_t^x = \eta_i^x \eta_j^x \eta_k^x$  and  $\Sigma_b^x = \eta_l^x \eta_m^x \eta_n^x$  to obtain Eq. (12).

- 
- [1] C. Broholm, R. J. Cava, S. A. Kivelson, D. G. Nocera, M. R. Norman, and T. Senthil, Quantum spin liquids, *Science* **367**, eaay0668 (2020).
  - [2] L. Balents, Spin liquids in frustrated magnets, *Nature (London)* **464**, 199 (2010).
  - [3] L. Savary and L. Balents, Quantum spin liquids: a review, *Rep. Prog. Phys.* **80**, 016502 (2017).
  - [4] R. Moessner and J. E. Moore, *Topological Phases of Matter* (Cambridge University Press, Cambridge, 2021).
  - [5] Y. Zhou, K. Kanoda, and T.-K. Ng, Quantum spin liquid states, *Rev. Mod. Phys.* **89**, 025003 (2017).
  - [6] J. Knolle and R. Moessner, A field guide to spin liquids, *Annu. Rev. Condens. Matter Phys.* **10**, 451 (2019).
  - [7] X.-G. Wen, *Colloquium: Zoo of quantum-topological phases of matter*, *Rev. Mod. Phys.* **89**, 041004 (2017).
  - [8] P. Anderson, Resonating valence bonds: A new kind of insulator? *Mater. Res. Bull.* **8**, 153 (1973).
  - [9] A. Kitaev, Anyons in an exactly solved model and beyond, *Ann. Phys.* **321**, 2 (2006).
  - [10] S. Trebst and C. Hickey, Kitaev materials, *Phys. Rep.* **950**, 1 (2022).
  - [11] G. Jackeli and G. Khaliullin, Mott insulators in the strong spin-orbit coupling limit: From heisenberg to a quantum compass and Kitaev models, *Phys. Rev. Lett.* **102**, 017205 (2009).
  - [12] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Unconventional superconductivity in magic-angle graphene superlattices, *Nature (London)* **556**, 43 (2018).
  - [13] T. Devakul, V. Crépel, Y. Zhang, and L. Fu, Magic in twisted transition metal dichalcogenide bilayers, *Nat. Commun.* **12**, 6730 (2021).
  - [14] G.-D. Zhao, X. Liu, T. Hu, F. Jia, Y. Cui, W. Wu, M.-H. Whangbo, and W. Ren, Difference in magnetic anisotropy of the ferromagnetic monolayers  $\text{VI}_3$  and  $\text{CrI}_3$ , *Phys. Rev. B* **103**, 014438 (2021).
  - [15] K. Hejazi, Z.-X. Luo, and L. Balents, Noncollinear phases in moiré magnets, *Proc. Natl. Acad. Sci. USA* **117**, 10721 (2020).
  - [16] K. Hejazi, Z.-X. Luo, and L. Balents, Heterobilayer moiré magnets: Moiré skyrmions and commensurate-incommensurate transitions, *Phys. Rev. B* **104**, L100406 (2021).
  - [17] M. Akram and O. Erten, Skyrmions in twisted van der Waals magnets, *Phys. Rev. B* **103**, L140406 (2021).
  - [18] M. Akram, H. LaBollita, D. Dey, J. Kapeghian, O. Erten, and A. S. Botana, Moiré skyrmions and chiral magnetic phases in twisted  $\text{CrX}_3$  ( $X = \text{I, Br, and Cl}$ ) bilayers, *Nano Lett.* **21**, 6633 (2021).
  - [19] J. Das, M. Akram, and O. Erten, Revival of antibiskyrmionic magnetic phases in bilayer  $\text{NiI}_2$ , *arXiv:2308.01484*.

- [20] Y. Xu, A. Ray, Y.-T. Shao, S. Jiang, K. Lee, D. Weber, J. E. Goldberger, K. Watanabe, T. Taniguchi, D. A. Muller, K. F. Mak, and J. Shan, Coexisting ferromagnetic–antiferromagnetic state in twisted bilayer  $\text{CrI}_3$ , *Nat. Nanotechnol.* **17**, 143 (2022).
- [21] T. Song, Q.-C. Sun, E. Anderson, C. Wang, J. Qian, T. Taniguchi, K. Watanabe, M. A. McGuire, R. Stöhr, D. Xiao, T. Cao, J. Wrachtrup, and X. Xu, Direct visualization of magnetic domains and moiré; magnetism in twisted 2D magnets, *Science* **374**, 1140 (2021).
- [22] H. Xie, X. Luo, Z. Ye, Z. Sun, G. Ye, S. H. Sung, H. Ge, S. Yan, Y. Fu, S. Tian, H. Lei, K. Sun, R. Hovden, R. He, and L. Zhao, Evidence of non-collinear spin texture in magnetic moiré superlattices, *Nat. Phys.* **19**, 1150 (2023).
- [23] M. Akram, J. Kapeghian, J. Das, R. Valenti, A. S. Botana, and O. Erten, Theory of moiré magnetism in twisted bilayer  $\alpha\text{-RuCl}_3$ , [arXiv:2310.12211](https://arxiv.org/abs/2310.12211).
- [24] U. F. P. Seifert, J. Gritsch, E. Wagner, D. G. Joshi, W. Brenig, M. Vojta, and K. P. Schmidt, Bilayer Kitaev models: Phase diagrams and novel phases, *Phys. Rev. B* **98**, 155101 (2018).
- [25] J. May-Mann and T. L. Hughes, Twisted Kitaev bilayers and the moiré Ising model, *Phys. Rev. B* **101**, 245126 (2020).
- [26] S. Haskell and A. Principi, Emergent hypermagic manifold in twisted Kitaev bilayers, *Phys. Rev. B* **106**, L161404 (2022).
- [27] H. Tomishige, J. Nasu, and A. Koga, Interlayer coupling effect on a bilayer Kitaev model, *Phys. Rev. B* **97**, 094403 (2018).
- [28] H. Tomishige, J. Nasu, and A. Koga, Low-temperature properties in the bilayer Kitaev model, *Phys. Rev. B* **99**, 174424 (2019).
- [29] E. Nica, M. Akram, A. Vijayvargia, R. Moessner, and O. Erten, Kitaev spin-orbital bilayers and their moiré superlattices, *npj Quantum Mater.* **8**, 9 (2023).
- [30] A. Vijayvargia, E. M. Nica, R. Moessner, Y.-M. Lu, and O. Erten, Magnetic fragmentation and fractionalized goldstone modes in a bilayer quantum spin liquid, *Phys. Rev. Res.* **5**, L022062 (2023).
- [31] M. A. Keskiner, O. Erten, and M. O. Oktel, Kitaev-type spin liquid on a quasicrystal, *Phys. Rev. B* **108**, 104208 (2023).
- [32] M. Akram, E. M. Nica, Y.-M. Lu, and O. Erten, Vison crystals, chiral and crystalline phases in the Yao-Lee model, *Phys. Rev. B* **108**, 224427 (2023).
- [33] E. H. Lieb, Flux phase of the half-filled band, *Phys. Rev. Lett.* **73**, 2158 (1994).
- [34] U. F. P. Seifert, T. Meng, and M. Vojta, Fractionalized fermi liquids and exotic superconductivity in the Kitaev-Kondo lattice, *Phys. Rev. B* **97**, 085118 (2018).
- [35] F. Yilmaz, A. P. Kampf, and S. K. Yip, Phase diagrams of Kitaev models for arbitrary magnetic field orientations, *Phys. Rev. Res.* **4**, 043024 (2022).
- [36] Y.-Z. You, I. Kimchi, and A. Vishwanath, Doping a spin-orbit mott insulator: Topological superconductivity from the Kitaev-Heisenberg model and possible application to  $(\text{Na}_2/\text{Li}_2)\text{IrO}_3$ , *Phys. Rev. B* **86**, 085145 (2012).
- [37] W. Choi, P. W. Klein, A. Rosch, and Y. B. Kim, Topological superconductivity in the Kondo-Kitaev model, *Phys. Rev. B* **98**, 155123 (2018).
- [38] C. Hickey and S. Trebst, Emergence of a field-driven  $U(1)$  spin liquid in the Kitaev honeycomb model, *Nat. Commun.* **10**, 530 (2019).
- [39] M. Gohlke, R. Moessner, and F. Pollmann, Dynamical and topological properties of the kitaev model in a  $[111]$  magnetic field, *Phys. Rev. B* **98**, 014418 (2018).
- [40] H. Zou, B. Liu, E. Zhao, and W. V. Liu, A continuum of compass spin models on the honeycomb lattice, *New J. Phys.* **18**, 053040 (2016).
- [41] J. Lou, L. Liang, Y. Yu, and Y. Chen, Global phase diagram of the extended Kitaev-Heisenberg model on honeycomb lattice, [arXiv:1501.06990](https://arxiv.org/abs/1501.06990).
- [42] Y. Huh, M. Punk, and S. Sachdev, Vison states and confinement transitions of  $\mathbb{Z}_2$  spin liquids on the kagome lattice, *Phys. Rev. B* **84**, 094419 (2011).
- [43] T. Senthil and M. P. A. Fisher,  $\mathbb{Z}_2$  gauge theory of electron fractionalization in strongly correlated systems, *Phys. Rev. B* **62**, 7850 (2000).

*Correction:* A typographical error in the NSF award number in the Acknowledgment section has been fixed.