

Magnetic fragmentation and fractionalized Goldstone modes in a bilayer quantum spin liquid

Aayush Vijayvargia¹, Emilian Marius Nica^{1,2}, Roderich Moessner³, Yuan-Ming Lu⁴, Omur Erten¹

¹*Department of Physics, Arizona State University, Tempe, AZ 85287, USA*

²*Department of Physics and Astronomy, Rice University, 6100 Main St, Houston 77005 TX, USA*

³*Max-Planck-Institut für Physik komplexer Systeme,
Nöthnitzer Strasse 38, 01187 Dresden, Germany*

⁴*Department of Physics, The Ohio State University, Columbus OH 43210, USA*

We study the phase diagram of a bilayer quantum spin liquid model with Kitaev-type interactions on a square lattice. We show that the low energy limit is described by a π -flux Hubbard model with an enhanced $SO(4)$ symmetry. The antiferromagnetic Mott transition of the Hubbard model signals a magnetic fragmentation transition for the spin and orbital degrees of freedom of the bilayer. The fragmented “Néel order” features a non-local string order parameter for an in-plane Néel component, in addition to an anisotropic local order parameter. The associated quantum order is characterized by an emergent $\mathbb{Z}_2 \times \mathbb{Z}_2$ gauge field when the Néel vector is along the \hat{z} direction, and a \mathbb{Z}_2 gauge field otherwise. We underpin these results with a perturbative calculation, which is consistent with the field theory analysis. We conclude with a discussion on the low energy collective excitations of these phases and show that the Goldstone boson of the $\mathbb{Z}_2 \times \mathbb{Z}_2$ phase is fractionalized and non-local.

Quantum spin liquids (QSLs) are frustrated magnets that do not exhibit long range magnetic order down to zero temperature[1–4]. Quantum fluctuations in these systems give rise to exotic phenomena such as fractionalization and long-range entanglement, which now become the defining properties for QSLs[5–7]. The Kitaev model on the honeycomb lattice[8] is one of the few examples of an exactly solvable model with a QSL ground state (GS). In recent years, remarkable progress in identifying candidate materials with strong Kitaev-type interactions has been achieved, in such instances as the iridates[9, 10] and α - RuCl_3 [11]. Kitaev interactions may also be strong in other van der Waals (vdW) materials[12]. Bilayers and moiré superlattices of vdW materials are new tunable quantum platforms for realizing a multitude of novel phases, with a variety of basic building blocks including graphene[13], semiconductors[14] and superconductors[15].

Motivated by these developments, we study the phase diagram of a bilayer QSL model with Kitaev-type interactions on a square lattice (see Fig 1(a)). First introduced in Ref. 16, the exact ground state of the monolayer model is an algebraic QSL featuring two flavors of Majorana fermions that are delocalized on the π -flux square lattice, and gapped π -flux (vison) excitations[17]. In the bilayer model Eq. (1), we add an Ising-type interlayer spin interaction, which commutes with the intra-layer flux operators and hence allows for controlled calculations. Our main results are summarized as follows: (i) Below the vison gap, we map the low-energy subspace of the bilayer model to a π -flux Hubbard model at half-filling, with an emergent $SO(4)$ symmetry. Monte Carlo studies of this model show an antiferromagnetic (AFM) Mott transition at critical $U_c \sim 6t$ [18]. (ii) The in-plane components of the AFM order parameter, i.e. $n^{x,y}$ of the Néel vector \mathbf{n} , correspond to non-local order parameters whereas the

out-of-plane component (n^z) is a local order parameter in terms of the spin and orbital degrees of freedom (DOF) of the bilayer system. (iii) The system features a $\mathbb{Z}_2 \times \mathbb{Z}_2$ gauge field when the Néel vector points along \hat{z} , and a \mathbb{Z}_2 gauge field otherwise. (iv) To complement the results of the Hubbard model, we perturbatively derive an effective Hamiltonian in the limit of large interlayer interactions. We confirm the magnetic fragmentation and topological degeneracy directly in terms of the original DOF, which are consistent with the Majorana fermion representation of the spin model. (v) We show that the Goldstone modes of the fragmented AFM order is fractionalized in the $\mathbb{Z}_2 \times \mathbb{Z}_2$ phase, in comparison to the normal Goldstone modes in the \mathbb{Z}_2 phase.

Microscopic model. One of the key conditions for the exact solution of the Kitaev model is the anticommutation relations of the Pauli matrices, $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$. Since there are only three Pauli matrices, this method can only be applied to lattices with coordination number $z = 3$ such as honeycomb, hyperhoneycomb and hyperoctagon lattices. However, it is possible to extend Kitaev’s method to Γ matrices that obey the Clifford algebra $\{\Gamma_i, \Gamma_j\} = 2\delta_{ij}$ [19, 20]. For instance, for a four-dimensional representation of the Clifford algebra, there are five Γ^α operators along with ten $\Gamma^{\alpha\beta} = \frac{i}{2}[\Gamma^\alpha, \Gamma^\beta]$ and an identity matrix, which span the local Hilbert space. Therefore, Kitaev’s construction can be extended to lattices with coordination number up to $z = 5$ [19, 20]. We adapt this representation and consider the intra-layer Hamiltonian [16], $H_K = -\sum_{\langle ij \rangle_{\gamma, \nu}} K_\nu (\Gamma_{\nu i}^\gamma \Gamma_{\nu j}^\gamma + \Gamma_{\nu i}^{\gamma 5} \Gamma_{\nu j}^{\gamma 5})$, where $\nu = 1, 2$ is the layer index and γ is the type of the bond as depicted in Fig. 1. We also introduce an inter-layer Ising interaction $H_J = J \sum_i \Gamma_{1i}^5 \Gamma_{2i}^5$. The full Hamiltonian can be expressed in terms of spins (σ) and orbital (τ) Pauli matrices using the relation $\Gamma^\alpha = -\sigma^y \otimes \tau^\alpha$ ($\alpha = x, y, z$), $\Gamma^4 = \sigma^x \otimes \mathbb{I}_2$ and

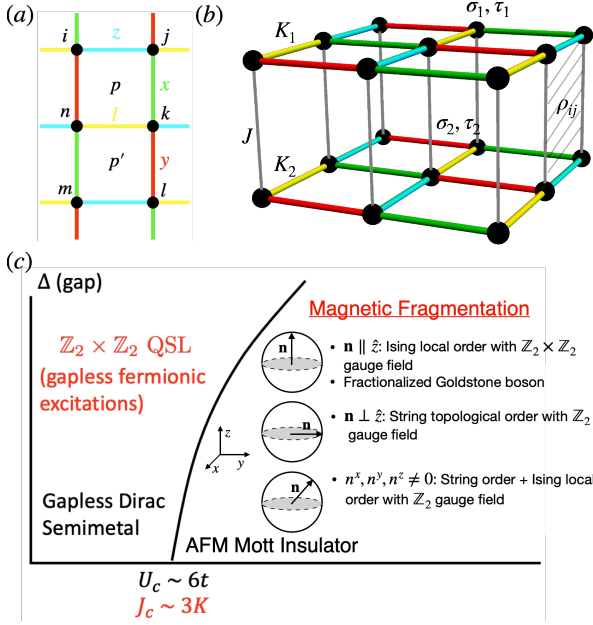


FIG. 1. Schematic of the model and the phase diagram: (a) single layer unit cell. Four different colors depict four types of bonds. There are two inequivalent plaquettes p and p' in a unit cell. (b) Bilayer model with intra-layer Kitaev parameters K_ν and an inter-layer exchange, J . (c) The low-energy description of the model is a π -flux Hubbard model which exhibits a Mott transition at $U/t \sim 6$ (black). In terms of the original degrees of freedom, the Mott transition corresponds to a magnetic fragmentation transition where a local magnetic order coexists with a non-local topological order.

$$\Gamma^5 = -\sigma^z \otimes \mathbb{I}_2,$$

$$H = H_K + H_J - \sum_{\langle ij \rangle_{\gamma, \nu}} K_\nu (\sigma_{\nu i}^x \sigma_{\nu j}^x + \sigma_{\nu i}^y \sigma_{\nu j}^y) (\tau_{\nu i}^\gamma \tau_{\nu j}^\gamma) + J \sum_i \sigma_{1i}^z \sigma_{2i}^z \quad (1)$$

Here $\tau^\gamma = \tau^x, \tau^y, \tau^z, \mathbb{I}$ for $\gamma = 1, 2, 3, 4$ respectively, corresponding to the 4 bonds incident on a vertex of the square lattice as shown in Fig. 1(a) and the sum is over all the γ bonds. Note that the $\gamma = 4$ (yellow) bond, which we refer as the ‘identity’ bond henceforth, has trivial orbital dependence. We consider $K_1 = K_2 = K$, unless specified otherwise. We identify two inequivalent intra-layer flux plaquette operators $W_{\nu p} = \sigma_{\nu k}^z \sigma_{\nu n}^z \tau_{\nu l}^x \tau_{\nu j}^y \tau_{\nu k}^x \tau_{\nu n}^y$ and $W_{\nu p'} = \sigma_{\nu k}^z \sigma_{\nu n}^z \tau_{\nu n}^x \tau_{\nu k}^y \tau_{\nu l}^x \tau_{\nu m}^y$, each with ± 1 eigenvalues. Both types of plaquette operators commute with the Hamiltonian and the Hilbert space is divided into sectors of conserved fluxes. Note that the Ising form of the inter-layer exchange is crucial to preserve $[W_{\nu p/p'}, H] = 0$ [17]. The intra-layer Hamiltonian can be solved by using a Majorana fermion representation of the Γ matrices [16], $H_K = K \sum_{\langle ij \rangle_{\gamma, \nu}} i u_{\nu, ij}^\gamma [c_{\nu i}^x c_{\nu j}^x + c_{\nu i}^y c_{\nu j}^y]$ where $u_{\nu, ij}^\gamma = i b_{\nu i}^\gamma b_{\nu j}^\gamma$ (see supplemental material (SM) for details). This representation is redundant and the

physical states in each layer must be restricted to the eigenstates of $D_{\nu j} = i b_{\nu j}^1 b_{\nu j}^2 b_{\nu j}^3 b_{\nu j}^4 c_{\nu j}^x c_{\nu j}^y$, with eigenvalues ± 1 . As in the Kitaev model, these constraints are imposed by the projection operator $P_\nu = \prod_i (1 + D_{\nu i})/2$. The intra-layer bond operators, $u_{\nu, ij}^\gamma$ commute with H_K and therefore are conserved with eigenvalues ± 1 . A \mathbb{Z}_2 gauge transformation at site i for layer ν involves flipping the signs of the Majorana fermions and bond operators, $c_{\nu i}^\alpha \rightarrow -c_{\nu i}^\alpha$; $u_{\nu, \langle ij \rangle}^\gamma \rightarrow -u_{\nu, \langle ij \rangle}^\gamma$. We combine the Majorana fermions on the two layers to form complex fermions, $f_{\nu i} = (c_{\nu i}^x - i c_{\nu i}^y)/2$ such that $H_K = 2K \sum_{\langle ij \rangle} u_{\nu, ij}^\gamma [i f_{\nu i}^\dagger f_{\nu j} + \text{H.c.}]$.

According to Lieb’s theorem [21], the GS manifold of H_K lies in the π -flux sector and consequently the eigenvalue of $W_{\nu p/p'} = \prod_{p/p'} u_{\nu ij}^\gamma$ is -1 in any GS configuration, for all square plaquettes. The spectrum is given by $E_K = \pm 4K \sqrt{\cos^2 k_x + \sin^2 k_y}$ which includes two inequivalent Dirac points at $(\pm \frac{\pi}{2}, 0)$.

Next, we represent the inter-layer interaction in terms of the Majorana fermions: $H_J = -J \sum_i c_{1i}^x c_{1i}^y c_{2i}^x c_{2i}^y$. H_J commutes with the intra-layer flux operators $W_{p/p'}$. However, the quartic form of the inter-layer exchange precludes the exact solvability of H , which can be expressed as

$$H = 2K \sum_{\langle ij \rangle_{\gamma, \nu}} u_{\nu, ij}^\gamma [i f_{\nu i}^\dagger f_{\nu j} + \text{H.c.}] + 2J \sum_i [n_{1i} + n_{2i} - 1]^2 \quad (2)$$

where $n_{\nu i} = f_{\nu i}^\dagger f_{\nu i}$.

Enhanced emergent symmetry. The Hamiltonian in eq. 2 has a global $U(1)$ symmetry in each layer ($\nu = 1, 2$), $e^{-i\theta \sum_i \sigma_{\nu i}^z} H e^{i\theta \sum_i \sigma_{\nu i}^z} = H$, a \mathbb{Z}_2 layer exchange symmetry \mathcal{X} , a particle-hole symmetry \mathcal{C} , and a time reversal symmetry \mathcal{T} . This results in a full symmetry group $G = O(2)_c \times O(2)_s \times \mathbb{Z}_2^\mathcal{T}$ of model (1), as detailed in SM. After the Majoranization, the two $U(1)$ rotations manifest themselves as:

$$U_c(\theta) f_{\nu i} U_c^{-1}(\theta) = e^{-i\theta} f_{\nu i}, \\ U_s(\theta) f_{\nu i} U_s^{-1}(\theta) = e^{-i\kappa\theta} f_{\nu i} \quad (3)$$

where $\kappa = -1, 1$ for $\nu = 1, 2$ respectively. This can be viewed as ‘charge’: $U_c = e^{i\theta \sum_{\nu i} f_{\nu i}^\dagger f_{\nu i}}$ and ‘pseudo-spin’: $U_s = e^{i\theta \sum_{\nu i} \kappa f_{\nu i}^\dagger f_{\nu i}}$ rotations where $\kappa = +1(-1)$ for $\nu = 1(2)$. The particle-hole symmetry \mathcal{C} and $U(1)$ charge rotations form the $O(2)_c$ subgroup, while the layer exchange \mathcal{X} and $U(1)$ pseudo-spin rotations form the $O(2)_s$ subgroup. Next, we fix the gauge by choosing $u_{\nu, ij}^\gamma = u_{ij}$, for both the layers and pick the π -flux configuration as discussed above. The resulting low-energy Hamiltonian (below the flux/vision gap) is a π -flux Hubbard model at half-filling with a hopping amplitude $t = 2K$ and interaction strength $U = 4J$. It is

Néel vector	Unbroken subgroup H	G/H	Gauge group
$\mathbf{n} \parallel \hat{z}$	$O(2)_c \times U(1)_s \times \mathbb{Z}_2^{\mathcal{X},\mathcal{T}}$	\mathbb{Z}_2	$\mathbb{Z}_2 \times \mathbb{Z}_2$
$\mathbf{n} \perp \hat{z}$	$O(2)_c \times \mathbb{Z}_2 \times \mathbb{Z}_2^{\mathcal{T}}$	S^1	\mathbb{Z}_2
$n_z \neq 0, n_x n_y \neq 0$	$O(2)_c \times \mathbb{Z}_2^{\mathcal{T}^n}$	$O(2)$	\mathbb{Z}_2

TABLE I. Distinct ground state phases associated with different orientations of the Néel vector \mathbf{n} in (4). In each phase, the full symmetry $G = O(2)_c \times O(2)_s \times \mathbb{Z}_2^{\mathcal{T}}$ of model (1) is spontaneously broken down to a different subgroup $H \leq G$, with an order parameter manifold $\mathcal{M} = G/H$. The gauge group for the associated topological order in each phase is also listed.

well established [22] that the Hubbard model on a bipartite lattice possesses an enhanced $G' = SO(4) \times \mathbb{Z}_2^{\mathcal{T}} = \mathbb{Z}_2^{\mathcal{T}} \times SU(2)_c \times SU(2)_s / \mathbb{Z}_2$ symmetry. The equivalence established above shows that our model also exhibits an enhanced $SO(4)$ symmetry at the low energy sector. In fact, this emergent $SO(4)$ symmetry exists in any subspace with a fixed flux configuration. Emergent symmetries can play a key role to describe the low energy physics of strongly correlated systems including cuprates[23] and iron pnictides[24].

Quantum Monte Carlo studies have shown that the repulsive ($J > 0$) π -flux Hubbard model displays a phase transition from Dirac semimetal to an AFM Mott insulator at $J/K \sim 3$ ($U/t \sim 6$) [25] (see Fig. 1(c)). Moreover, due to $J \rightarrow -J$ mapping in the Hubbard model, the phase diagram is symmetric for ferromagnetic (FM) and AFM inter-layer exchange, for which the Néel order maps to superconducting and charge density wave orders.

The Néel vector of the AFM order

$$\mathbf{n} = \frac{1}{N} \sum_i \mathbf{n}_i, \quad \mathbf{n}_i = (-1)^{r_{ix}+r_{iy}} \langle f_{\mu i}^\dagger \boldsymbol{\sigma}_{\mu\nu} f_{\nu i} \rangle, \quad (4)$$

where N is the number of sites and $r_{ix(y)}$ is the $x(y)$ coordinate of site i , can point along any direction on the Bloch sphere. Goldstone modes always arise since the emergent symmetry $G' = SO(4) \times \mathbb{Z}_2^{\mathcal{T}}$ is spontaneously broken down to $H' = SU(2)_c \times U(1)_s \times \mathbb{Z}_2^{\mathcal{T}}$ in the Néel order (see SM for details). However, as we show below, different orientations of the Néel vector correspond to distinct ground states with different symmetry and topological properties, as summarized in Table I.

(i) The Néel vector points along the z-direction, $\mathbf{n} \parallel \hat{z}$,

$$n_i^z = (-1)^{r_{ix}+r_{iy}} \langle f_{1i}^\dagger f_{1i} - f_{2i}^\dagger f_{2i} \rangle \neq 0 \quad (5)$$

In terms of Majorana fermions, eq. 5 takes the form $n_i^z = (-1)^{r_{ix}+r_{iy}} i (c_{1i}^x c_{1i}^y - c_{2i}^x c_{2i}^y)$. Note that n_i^z is invariant under local \mathbb{Z}_2 gauge transformations ($c_{\nu i}^x, c_{\nu i}^y$) \rightarrow ($-c_{\nu i}^x, -c_{\nu i}^y$), hence corresponding to a physical operator $(-1)^{r_{ix}+r_{iy}} (\sigma_{1i}^z - \sigma_{2i}^z)$. In other words, n^z is a local order parameter of a Landau-type long range order.

The \mathbb{Z}_2 gauge fields for the two layers, $u_{1,ij}$ and $u_{2,ij}$, are decoupled, leading to a $\mathbb{Z}_2 \times \mathbb{Z}_2$ topolog-

ical order described by 4-component Abelian Chern-Simons theory[26] characterized by matrix $\mathbf{K} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$. However, the Goldstone mode of the Néel order

$$n^x + in^y \sim b_{\vec{k}=(\pi,\pi)} \sim \sum_i (-1)^{i_x+i_y} f_{2i}^\dagger f_{1i} \quad (6)$$

is not a gauge-invariant quantity, but instead an anyon obeying mutual semion statistics with the vison in each layer. More precisely, the above Goldstone mode carries the gauge charge for the \mathbb{Z}_2 gauge field from each layer. Incorporating the gapless anyon b in (6) into the low energy description, the effective field theory for this algebraic spin liquid reads

$$\mathcal{L}_{\text{ASL}} = \sum_{I,J} \frac{\epsilon^{\mu\nu\rho}}{4\pi} a_\mu^I \mathbf{K}_{I,J} \partial_\nu a_\rho^J - \sum_{\alpha,I} \frac{\epsilon^{\mu\nu\rho}}{2\pi} A_\mu^\alpha \mathbf{q}_I^\alpha \partial_\nu a_\rho^I + |(-i\partial_\mu - 2A_\mu^s - a_\mu^1 - a_\mu^2 + a_\mu^3 - a_\mu^4)^2 b|^2 + \dots \quad (7)$$

where $A_\mu^{\alpha=c,s}$ label the charge and pseudo-spin external gauge fields, and

$$\mathbf{q}_c = (2, 0, 2, 0)^T, \quad \mathbf{q}_s = (2, 0, -2, 0)^T. \quad (8)$$

are the charge and pseudo-spin vectors[26] for the Chern-Simons theory.

(ii) The Néel vector lies in-plane, e.g. $\mathbf{n} \perp \hat{z}$ with

$$n^+ \equiv n_i^x + in_i^y = (-1)^{r_{ix}+r_{iy}} \langle f_{1i}^\dagger f_{2i} \rangle \neq 0 \quad (9)$$

Unlike n^z , the in-plane components, n^x and n^y are not gauge invariant as the local gauge transformation maps $n^{x(y)} \rightarrow -n^{x(y)}$. However, a non-local gauge invariant correlator can be defined [27].

$$\mathcal{C}^{x(y)}(r, r') = \langle n^{x(y)}(r) B(r, r') n^{x(y)}(r') \rangle, \quad (10)$$

where the gauge string for fermions, $B(r, r') = \prod_{ij \in (r, r')} u_{1ij} u_{2ij}$, connects operators at the end sites (r, r') . The value of $\mathcal{C}^{x(y)}(r, r')$ is the same in all gauge choices. Therefore, the ground state, symmetrized over all gauge configurations through the projection procedure, also has the same value of $\mathcal{C}^{x(y)}(r, r')$, signifying a string order parameter. Physically, the long-range string order corresponds to the condensation of anyon b in the field theory (7), hence breaking the gauge group down to \mathbb{Z}_2 via the Higgs mechanism[28].

An alternative way to understand the gauge structure is to notice the following local order parameter for the in-plane Néel order

$$S_{\langle i,j \rangle}^+ \equiv \langle n_i^+ B(i, j) n_j^+ \rangle \quad (11)$$

for a pair of nearest neighbor sites $\langle i, j \rangle$. Due to the mutual braiding phase of $e^{i\pi}$ between a vison and a fermion in each layer, a vison from layer 1 (or 2) is nothing but

a vortex for the above local order parameter, since $S_{\langle i,j \rangle}^+$ acquires a $e^{\pm 2\pi i}$ phase as it travels around a vison from layer 1 (2). The logarithmic confinement of vortices in the in-plane Neel phase suggest that the vison from layer 1 (or 2) is confined, therefore reducing the $Z_2 \times Z_2$ gauge group down to Z_2 . A similar conclusion can be drawn if the Néel vector has both in-plane and \hat{z} components.

In general, the ground state can have both non-zero out-of-plane (n^z) and in-plane (n^x, n^y) components. The term, ‘magnetic fragmentation’ is coined for phases that display a coexistence of a local Landau-type order parameter and a non-local topological order [29–34]. Magnetic fragmentation is theoretically predicted[29] and experimentally observed[30] in spin ice materials such as $\text{Nd}_2\text{Zr}_2\text{O}_7$ where a local AFM order coexists with a spin liquid with FM correlations. The conclusions we draw from the Hubbard model rely on mean-field order parameters. Next, we underpin these results by a perturbative analysis.

Perturbative analysis in the limit of large inter-layer exchange. We corroborate the results of the Hubbard model (Eq. 2) by considering the bilayer in the large- J limit, without reference to the Majorana representation (Eq. 1), on a torus. We introduce effective pseudo-spin and orbital DOF appropriate to this limit. We next derive effective models on the large- J GS manifold, to fourth order in the intra-layer coupling K . By analogy to the Hubbard model, we distinguish between cases with a) Z_2 and b) $Z_2 \times Z_2$ topological order. For a), we show that the GS manifold is a state of uniform π flux, with a finite $\mathcal{C}^{x(y)}$ correlator (Eq. 10). We also demonstrate that the GS manifold has four-fold topological degeneracy and that the visons are confined. For b), we also obtain a GS with uniform π flux, which has sixteen-fold topological degeneracy and deconfined vison excitations. These results naturally lead to the conclusion that the two phases are separated by a topological phase transition.

Effective degrees of freedom. We first introduce the effective pseudo-spin and orbital DOF. For $K = 0$ and finite FM inter-layer interactions ($J < 0$), the spins on overlapping sites form GS doublets $|\uparrow_1\uparrow_2\rangle, |\downarrow_1\downarrow_2\rangle$. These can be represented by a bilayer pseudo-spin

$$\begin{aligned}\eta_i^z &= \frac{1}{4}(\sigma_{1i}^z + \sigma_{2i}^z) \\ \eta_i^\pm &= \frac{1}{4}\sigma_{1i}^\pm\sigma_{2i}^\pm,\end{aligned}\quad (12)$$

obeying an $\text{SU}(2)$ algebra. In addition, the orbital DOF for each pair of overlapping sites form a four dimensional Hilbert space, corresponding to one singlet and three triplet configurations. To represent these states, we introduce the inter-layer orbital operators $q_i^\gamma = \tau_{1i}^\gamma\tau_{2i}^\gamma$, $\gamma = x, y, z$, which mutually commute as $[q_i^\alpha, q_i^\beta] = 0$. The four orbital states can be labeled by the three eigenvalues $q_i^\gamma = \pm 1$, constrained to obey $\prod_\gamma q_i^\gamma = -1$. The Hilbert space thus includes all states of the form

$$|\{\eta^z\}, \{q^\gamma\}\rangle = |\{\eta_i^z\}\rangle \otimes |\{(q_i^x, q_i^y, q_i^z)\}\rangle, \quad (13)$$

with the implicit local constraint. Pairs of nearest-neighbor $q_{i/j}^\gamma$ define the bond variables

$$\rho_{ij}^\gamma = q_i^\gamma q_j^\gamma \quad (14)$$

which take on values of ± 1 for $\gamma \in \{x, y, z\}$, while they are trivially equal to 1 for additional identity bonds, labeled by $\rho^{\gamma=I} \equiv 1$. To any $\{q_i^\gamma\}$ configuration, we can associate a unique $\{\rho_{ij}^\gamma\}$ bond configuration, while the converse is not true. Orbital states like $|\phi\rangle_0 = |\forall q_i^\gamma = -1\rangle$, which have uniform $\rho_{ij}^\gamma = 1$, play an important role in all subsequent discussions.

Defects in $|\phi\rangle_0$ take the form of strings of negative bonds as shown in Fig. 2 (b). Defects in both pseudo-spin and ρ_{ij}^γ bonds are introduced by operating with the flux operators $W_{\nu p/p'}$. Each of these flips the pseudo-spin components along x/y as $\eta_{n,k}^{x/y} \rightarrow -\eta_{n,k}^{x/y}$ in the corresponding unit cell (2 (a)). Each also changes the signs of all six ρ^γ bonds connected with sites n, k . Note that any string defect cannot be eliminated by application of $W_{\nu p/p'}$ operators.

Effective Hamiltonian. The effective Hamiltonian $H_{\eta-\rho}$, projected onto the $K = 0$ GS manifold, reads

$$H_{\eta-\rho} = H_{g_2} + H_{g_4}, \quad (15)$$

where

$$H_{g_2} = g_2 \left[\sum_{\langle ij \rangle} \eta_i^z \eta_j^z + \sum_{\langle ij \rangle_\gamma} \frac{1}{2} (\eta_i^+ \eta_j^- + \eta_i^- \eta_j^+) \rho_{ij}^\gamma \right] + \sum_i (-1)^{i_x + i_y} (h_x \eta_i^x + h_z \eta_i^z), \quad (16)$$

is obtained at second order in K ($g_2 = K^2/4J$). Note the

distinction between NN, in-plane pseudo-spins connected

via variable and trivial identity ρ_{ij}^γ bonds, respectively. We introduce small perturbations $h_{x/z} > 0$ to explicitly break the continuous symmetry, enforcing the staggered pseudo-spin configurations along x/z , respectively. As it turns out, these respectively correspond to \mathbb{Z}_2 and $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order. The fourth-order contribution is

$$H_{g_4} = g_4 \left[\sum_{\nu,p} \eta_p^\square W_{\nu p} + \sum_{\nu,p'} \eta_{p'}^\square W_{\nu p'} \right], \quad (17)$$

where $g_4 = K^4/J^3$. η_p^\square and $\eta_{p'}^\square$ include linear combinations of products of η and q^γ operators around p/p' plaquettes. H_{g_2} commutes with all flux operators for $h_x \rightarrow 0$, while this always holds for H_{g_4} . Technical details and derivations related to the effective Hamiltonian and the following sections are relegated to SM.

\mathbb{Z}_2 topological order. We consider $H_{\eta-\rho}$ with $h_x > 0$ and $h_z = 0$, on a torus. Exact diagonalization calculations indicate that the GS manifold of H_{g_2} includes $|\eta_{xx}, \phi_0\rangle$, with η_{xx} denoting a finite staggered pseudo-spin along x (see SM). Importantly, any configuration with string defects in the ρ^γ bonds are gapped, with an energy cost which scales as the string length.

We next consider the evolution of the GS manifold at H_{g_4} level for $h_x \ll g_4 \ll g_2$. In SM, we show that exact diagonalization calculations indicate H_{g_4} projects $|\eta_{xx}, \phi_0\rangle$ onto a state of uniform π flux per plaquette:

$$|\Psi_{\text{GS}}\rangle = \prod_{\nu,p,p'} \frac{(1 - W_{\nu p})(1 - W_{\nu p'})}{4} |\tilde{\eta}_{xx}; \phi_0\rangle + O\left(\frac{h_x}{g_4}\right). \quad (18)$$

Note that $\tilde{\eta}_{xx}$ is a state of staggered pseudo-spins including corrections at both H_{g_2} and H_{g_4} levels. This is a state of definite π flux since $W_{\nu p}(1 - W_{\nu p}) = -(1 - W_{\nu p})$. Moreover, all $W_{\nu p/p'}$ commute with the operator

$$\mathcal{D}_{ij}^x = \eta_i^x \left(\prod_{i'j' \in C_{ij}} \rho_{i'j'}^\gamma \right) \eta_j^x. \quad (19)$$

Consequently, the latter has a finite expectation value in $|\Psi_{\text{GS}}\rangle$ for any pair of i, j , reflecting a locking of pseudo-spin and ρ_{ij}^γ bond configuration. Moreover, \mathcal{D}_{ij}^x is equivalent to a gauge-invariant correlator of the Hubbard model corresponding to an in-plane Néel vector (Eq. 10).

Note that any state with an open string defect, obtained by first including strings in ϕ_0 , involves one or more visons on the plaquettes at each end (2 (b)). As the energy of this excitation depends on the string length and diverges for infinite vison separation, it follows that the latter are confined in an infinite system.

$\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order. In this case, we consider $H_{\eta-\rho}$ with $h_z > 0$ and $h_x = 0$. By analogy with the

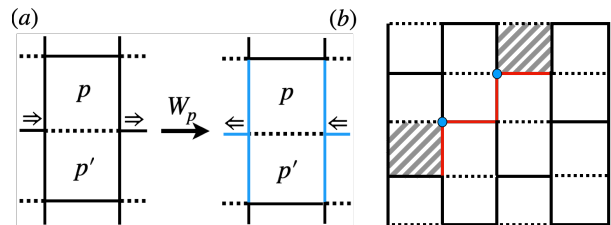


FIG. 2. Illustration of the pseudo-spin and bond ($\eta - \rho$) configurations. (a) Effect of flux operator $W_{\nu p}$ on a state with fixed ρ_{ij}^γ bonds. $W_{\nu p}$ changes the signs of the bonds marked in blue. It also changes the signs of $\eta^{x/y}$ pseudo-spins on the identity (dashed) bond. (b) String defect, with red lines indicating bonds with signs opposite to the background bonds, marked in black. The string shown here is created by operating with $\tau_{\nu i}^x$ on any orbital state along the sites marked with blue dots. When operating on the GS in Eq. 18, the τ 's change the fluxes on the hashed plaquettes, since they anti-commute with $W_{\nu p/p'}$. Consequently, this open string terminates with a pair of visons.

case with \mathbb{Z}_2 topological order, the GS manifold of H_{g_2} now includes $|\eta_{zz}, \phi_0\rangle$, where η_{zz} indicates finite staggered pseudo-spins along z . Similarly, any open string defects are gapped. However, in contrast to the case for \mathbb{Z}_2 topological order, the gap for these excitations remains finite for arbitrary string length, in the infinite-system size. In the same limit, states with strings forming non-contractible loops become degenerate with $|\phi_0\rangle$ for $h_z \gg g_{2/4}$.

The effect of H_{g_4} is analogous to the case with \mathbb{Z}_2 topological order. Consequently, the GS has the form in Eq. 18, with the replacement $|\tilde{\eta}_{xx}; \phi_0\rangle \rightarrow |\tilde{\eta}_{zz}; \phi_0\rangle$, indicating a surviving pseudo-spin staggering. While the two-point correlator for the η^z pseudo-spins is always finite, \mathcal{D}_{ij}^x vanishes for infinite separation as in the Hubbard model with $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order.

Since the energy cost of an open string remains finite even in an infinite-size system, the visons are deconfined. Similarly, as states with non-contractible loops of negative ρ^γ bonds become degenerate with the GS, in the limit of infinite system-size, the latter acquires additional topological degeneracy. Consequently, this entails a sixteen-fold topological degeneracy on the torus, consistent with $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order.

Conclusion and outlook. We have studied a bilayer adaptation of a QSL model on a square lattice with Kitaev-type interactions. We have shown that the low energy model exhibits an AFM Mott transition which corresponds to magnetic fragmentation in terms of the original DOF. We have corroborated these results by a perturbative calculation for the topological degeneracy, which is consistent with field theory analysis. The analysis we have presented here may be of particular value as a largely tractable yet highly non-trivial instance of

magnetic fragmentation.

Interesting future directions include examining the role of fluctuations on the emergent symmetry which may reduce the ground state manifold via order by disorder[35–37]. Another direction is to generalize our mechanism for fractionalized Goldstone modes in bilayer systems to multilayer systems with larger emergent symmetries, such as $SU(N)$.

The study of moiré superlattices of QSLs is another intriguing direction[27, 38]. Our work suggests that manifold new phenomena arising from a combination of emergent symmetry and strong interactions are awaiting discovery here, providing a new vista on strongly correlated magnetism.

Acknowledgements. We thank Nandini Trivedi and Natalia Perkins for fruitful discussions. OE acknowledge support from NSF Award No. DMR 2234352. YML is supported by NSF under Award No. DMR 2011876. EMN acknowledges support by NSF under Grant No. DMR-2220603. This work in part supported by the Deutsche Forschungsgemeinschaft (DFG) via SFB 1143 (project-id 247310070) and cluster of excellence ct.qmat (EXC 2147, project-id 390858490).

-
- [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* **464**, 199–208 (2010).
- [3] L. Savary and L. Balents, “Quantum spin liquids: a review,” *Reports on Progress in Physics* **80**, 016502 (2016).
- [4] R. Moessner and J. E. Moore, *Topological Phases of Matter* (Cambridge University Press, 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,” *Annual Review of Condensed Matter Physics* **10**, 451–472 (2019).
- [7] X. G. Wen, “Colloquium: Zoo of quantum-topological phases of matter,” *Rev. Mod. Phys.* **89**, 041004 (2017).
- [8] A. Kitaev, “Anyons in an exactly solved model and beyond,” *Annals of Physics* **321**, 2–111 (2006).
- [9] S. Hwan Chun, J. W. Kim, J. Kim, H. Zheng, C. C. Stoumpos, C. D. Malliakas, J. F. Mitchell, Kavita Mehlawat, Yogesh Singh, Y. Choi, T. Gog, A. Al-Zein, M. Moretti Sala, M. Krisch, J. Chaloupka, G. Jackeli, G. Khaliullin, and B. J. Kim, “Direct evidence for dominant bond-directional interactions in a honeycomb lattice iridate Na_2IrO_3 ,” *Nature Physics* **11**, 462–466 (2015).
- [10] K. Kitagawa, T. Takayama, Y. Matsumoto, A. Kato, R. Takano, Y. Kishimoto, S. Bette, R. Dinnebier, G. Jackeli, and H. Takagi, “A spin-orbital-entangled quantum liquid on a honeycomb lattice,” *Nature* **554**, 341–345 (2018).
- [11] H. Takagi, T. Takayama, G. Jackeli, G. Khaliullin, and S. E. Nagler, “Concept and realization of Kitaev quantum spin liquids,” *Nature Reviews Physics* **1**, 264–280 (2019).
- [12] I. Lee, F. G. Utermohlen, D. Weber, K. Hwang, C. Zhang, J. van Tol, J. E. Goldberger, N. Trivedi, and P. C. Hammel, “Fundamental spin interactions underlying the magnetic anisotropy in the Kitaev ferromagnet CrI_3 ,” *Phys. Rev. Lett.* **124**, 017201 (2020).
- [13] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, “Unconventional superconductivity in magic-angle graphene superlattices,” *Nature* **556**, 43–50 (2018).
- [14] T. Devakul, V. Crépel, Y. Zhang, and L. Fu, “Magic in twisted transition metal dichalcogenide bilayers,” *Nature Communications* **12**, 6730 (2021).
- [15] S. Y. F. Zhao, N. Poccia, X. Cui, P. A. Volkov, H. Yoo, R. Engelke, Y. Ronen, R. Zhong, G. Gu, S. Plugge, T. Tummuru, M. Franz, J. H. Pixley, and P. Kim, “Emergent interfacial superconductivity between twisted cuprate superconductors,” *arXiv e-prints* (2021), 10.48550/ARXIV.2108.13455.
- [16] R. Nakai, S. Ryu, and A. Furusaki, “Time-reversal symmetric Kitaev model and topological superconductor in two dimensions,” *Phys. Rev. B* **85**, 155119 (2012).
- [17] U. F. P. Seifert, X. Y. Dong, S. Chulliparambil, M. Vojta, H. H. Tu, and L. Janssen, “Fractionalized fermionic quantum criticality in spin-orbital Mott insulators,” *Phys. Rev. Lett.* **125**, 257202 (2020).
- [18] Y. Otsuka, S. Yunoki, and S. Sorella, “Mott transition in the 2d Hubbard model with π -flux,” *Proceedings of the International Conference on Strongly Correlated Electron Systems (SCES2013)*, **3**, 013021 (2014).
- [19] C. Wu, D. Arovas, and H. H. Hung, “T-matrix generalization of the Kitaev model,” *Phys. Rev. B* **79**, 134427 (2009).
- [20] H. Yao, S. C. Zhang, and S. A. Kivelson, “Algebraic spin liquid in an exactly solvable spin model,” *Phys. Rev. Lett.* **102**, 217202 (2009).
- [21] E. H. Lieb, “Flux phase of the half-filled band,” *Phys. Rev. Lett.* **73**, 2158–2161 (1994).
- [22] C. N. Yang and S.C. Zhang, “ $SO(4)$ symmetry in a Hubbard model,” *Modern Physics Letters B* **04**, 759–766 (1990).
- [23] E. Demler, W. Hanke, and S. C. Zhang, “ $SO(5)$ theory of antiferromagnetism and superconductivity,” *Rev. Mod. Phys.* **76**, 909–974 (2004).
- [24] D. Podolsky, H.-Y. Kee, and Y. B. Kim, “Collective modes and emergent symmetry of superconductivity and magnetism in the iron pnictides,” *Europhysics Letters* **88**, 17004 (2009).
- [25] C. C. Chang and R. T. Scalettar, “Quantum disordered phase near the Mott transition in the staggered-flux Hubbard model on a square lattice,” *Phys. Rev. Lett.* **109**, 026404 (2012).
- [26] X. G. Wen and A. Zee, “Classification of abelian quantum Hall states and matrix formulation of topological fluids,” *Phys. Rev. B* **46**, 2290–2301 (1992).
- [27] 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).
- [28] T.H. Hansson, V. Oganesyan, and S.L. Sondhi, “Superconductors are topologically ordered,” *Annals of Physics* **313**, 497–538 (2004).
- [29] M. E. Brooks-Bartlett, S. T. Banks, L. D. C. Jaubert, A. Harman-Clarke, and P. C. W. Holdsworth, “Magnetic-moment fragmentation and monopole crystallization,” *Phys. Rev. X* **4**, 011007 (2014).

- [30] S. Petit, E. Lhotel, B. Canals, M. Ciomaga Hatnean, J. Ollivier, H. Mutka, E. Ressouche, A. R. Wildes, M. R. Lees, and G. Balakrishnan, “Observation of magnetic fragmentation in spin ice,” *Nature Physics* **12**, 746–750 (2016).
- [31] E. Lefrançois, V. Cathelin, E. Lhotel, J. Robert, P. Lejay, C. V. Colin, B. Canals, F. Damay, J. Ollivier, B. Fåk, L. C. Chapon, R. Ballou, and V. Simonet, “Fragmentation in spin ice from magnetic charge injection,” *Nature Communications* **8**, 209 (2017).
- [32] A. Zorko, M. Pregelj, M. Klanjšek, M. Gomilšek, Z. Jagličić, J. S. Lord, J. A. T. Verezhak, T. Shang, W. Sun, and J.-X. Mi, “Co-existence of magnetic order and persistent spin dynamics in a quantum kagome antiferromagnet with no intersite mixing,” *Phys. Rev. B* **99**, 214441 (2019).
- [33] C. Mauws, A. M. Hallas, G. Sala, A. A. Aczel, P. M. Sarte, J. Gaudet, D. Ziat, J. A. Quilliam, J. A. Lussier, M. Bieringer, H. D. Zhou, A. Wildes, M. B. Stone, D. Abernathy, G. M. Luke, B. D. Gaulin, and C. R. Wiebe, “Dipolar-octupolar Ising antiferromagnetism in $\text{Sm}_2\text{Ti}_2\text{O}_7$: A moment fragmentation candidate,” *Phys. Rev. B* **98**, 100401 (2018).
- [34] Y. Q. Wang, C. Liu, and Y. M. Lu, “Theory of topological defects and textures in two-dimensional quantum orders with spontaneous symmetry breaking,” arXiv e-prints (2022).
- [35] E. F. Shender, “Antiferromagnetic garnets with fluctuationally interacting sublattices,” *Soviet Journal of Experimental and Theoretical Physics* **56**, 178 (1982).
- [36] C. L. Henley, “Ordering due to disorder in a frustrated vector antiferromagnet,” *Phys. Rev. Lett.* **62**, 2056–2059 (1989).
- [37] A. G. Green, G. Conduit, and F. Krüger, “Quantum order-by-disorder in strongly correlated metals,” *Annual Review of Condensed Matter Physics* **9**, 59–77 (2018).
- [38] Z. X. Luo, U. F. P. Seifert, and L. Balents, “Twisted bilayer $U(1)$ Dirac spin liquids,” *Phys. Rev. B* **106**, 144437 (2022).

Supplementary material for “Emergent symmetry and fractionalized Goldstone modes in a bilayer quantum spin liquid”

Aayush Vijayvargia¹, Emilian Marius Nica^{1,2}, Roderich Moessner³, Yuan-Ming Lu⁴, Onur Erten¹

¹*Department of Physics, Arizona State University, Tempe, AZ 85287, USA*

²*Department of Physics and Astronomy, Rice University, 6100 Main St, Houston 77005 TX, USA*

³*Max-Planck-Institut für Physik komplexer Systeme,
Nöthnitzer Strasse 38, 01187 Dresden, Germany*

⁴*Department of Physics, The Ohio State University, Columbus OH 43210, USA*

I. MAJORANA FERMION REPRESENTATION

The Hamiltonian of the bilayer is

$$H = H_K + H_J \quad (1)$$

where H_K and H_J are the intra- and inter-layer terms, respectively. H_K has the form [?],

$$H_K = -K \sum_{\langle ij \rangle_{\gamma, \nu}} (\Gamma_{\nu i}^{\gamma} \Gamma_{\nu j}^{\gamma} + \Gamma_{\nu i}^{\gamma^5} \Gamma_{\nu j}^{\gamma^5}) \quad (2)$$

where Γ^{γ} matrices obey the Clifford algebra, $\{\Gamma^{\alpha}, \Gamma^{\beta}\} = \delta_{\alpha, \beta}$ and $\Gamma^{\alpha\beta} = [\Gamma^{\alpha}, \Gamma^{\beta}]$. Using $\Gamma^{\alpha} = -\sigma^y \otimes \tau^{\alpha}$ ($\alpha = x, y, z$), $\Gamma^4 = \sigma^x \otimes \mathbb{I}_2$ and $\Gamma^5 = -\sigma^z \otimes \mathbb{I}_2$, H_K can be re-cast as

$$H_K = - \sum_{\langle ij \rangle_{\gamma, \nu}} K_{\nu} (\sigma_{\nu i}^x \sigma_{\nu j}^x + \sigma_{\nu i}^y \sigma_{\nu j}^y) (\tau_{\nu i}^{\gamma} \tau_{\nu j}^{\gamma}) \quad (3)$$

where $\sigma(\tau)$ Pauli matrices act on spin(orbital) degrees of freedom (DOF). Here $\tau^{\gamma} = \tau^x, \tau^y, \tau^z, \mathbb{I}$ and $\gamma = 1, 2, 3, 4$ correspond to the 4 inequivalent bonds at each site. We introduce a Majorana fermion representation via

$$\Gamma_j^{\alpha} = i b_j^{\alpha} c_j, \quad \Gamma_i^{\alpha\beta} = i b_i^{\alpha} b_i^{\beta} \quad (4)$$

and relabel $b_i^5 \rightarrow c_i^x$ and $c_i \rightarrow c_i^y$ for convenience to obtain

$$H_K = K \sum_{\langle ij \rangle_{\gamma}} i u_{\nu, ij} [c_{\nu i}^x c_{\nu j}^x + c_{\nu i}^y c_{\nu j}^y] \quad (5)$$

where $u_{\nu, ij} = i b_{\nu i}^{\gamma} b_{\nu j}^{\gamma}$. The inter-layer interaction Hamiltonian, $H_J = J \sum_i \Gamma_{1i}^5 \Gamma_{2i}^5$ can be similarly expressed as

$$\begin{aligned} H_J &= J \sum_i \sigma_{1i}^z \sigma_{2i}^z \\ &= -J \sum_i c_{1,i}^x c_{1,i}^y c_{2,i}^x c_{2,i}^y \end{aligned} \quad (6)$$

Note that the fermion Hilbert space per site is $2^3 = 8$ -dimensional, twice the size of the 4-dimensional physical Hilbert space expanded by Pauli matrices $\vec{\sigma}$ and $\vec{\tau}$. To faithfully represent the physical system with the Majorana fermion degrees of freedom, we have to enforce the following onsite constraint

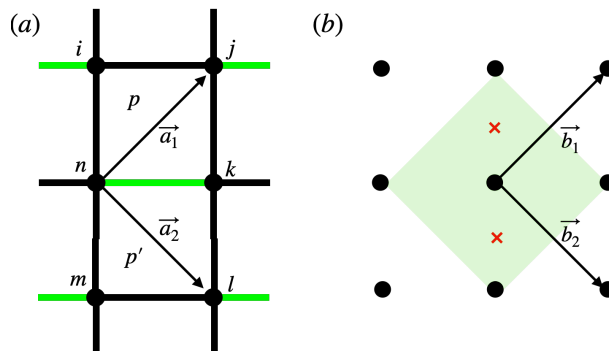
$$D_{\nu j} = i b_{\nu j}^1 b_{\nu j}^2 b_{\nu j}^3 b_{\nu j}^4 c_{\nu j}^x c_{\nu j}^y = 1, \quad \forall j, \nu. \quad (7)$$

II. SOLUTION OF THE INTRA-LAYER HAMILTONIAN

According to Lieb's theorem[?], the ground state of each layer of H_K lies in the π -flux sector. Such states can be obtained by choosing gauge configurations where every x-bond (see Fig. 1(a) in the main text) has $u_{ij} = -1$, while all remaining bonds have $u_{ij} = 1$. The unit cell includes two inequivalent square plaquettes with primitive lattice vectors $a_{1,2} = (1, \pm 1)$ and corresponding reciprocal vectors $b_{1,2} = \pi(1, \pm 1)$ (in units of $a = 1$) as shown in Supplementary figure 1. Performing a Fourier transform $f_{(A,B)i} = \sum_k e^{-ikr_i} f_{(A,B)k}$, we obtain

$$H_K = \sum_{k \in 1^{\text{st}} \text{ BZ}} \begin{pmatrix} f_{Ak}^{\dagger} & f_{Bk}^{\dagger} \end{pmatrix} \begin{pmatrix} 0 & 4K(\cos k_y + i \sin k_x) \\ 4K(\cos k_y - i \sin k_x) & 0 \end{pmatrix} \begin{pmatrix} f_{Ak} \\ f_{Bk} \end{pmatrix} \quad (8)$$

with a dispersion $E_K = \pm 4K \sqrt{\sin^2 k_x + \cos^2 k_y}$. E_K has two inequivalent Dirac points at $(0, \pm \frac{\pi}{2})$ as shown in Supplementary figure 1.



Supplementary figure 1. (a) Lattice vectors $\vec{a}_{1,2} = (1, \pm 1)$. The unit cell includes two inequivalent plaquettes labeled p and p' . Note the specific gauge choice adopted here, where green bonds and black bonds correspond to $u_{ij} = -1$ and $u_{ij} = 1$, respectively. (b) Brillouin zone (BZ) (green square) with reciprocal vectors $\vec{b}_{1,2} = \pi(1, \pm 1)$. Red crosses indicate the two inequivalent Dirac points at $(0, \pm \frac{\pi}{2})$.

III. SYMMETRY AND TOPOLOGICAL PROPERTIES

A. Symmetry of the full Hamiltonian

We first consider the symmetry of the single layer Hamiltonian (2). The (spinless) time reversal symmetry operator $\mathcal{T} = \mathcal{K}$ is simply implemented by the complex conjugation \mathcal{K} . In addition, we have a unitary $U(1)_\nu \times \mathbb{Z}_2^{\mathcal{C}}$ symmetry, generated by $U(1)$ “charge” rotations in each layer ν

$$U(1)_\nu = \{e^{i\theta Q_\nu} | Q_\nu = \sum_j \frac{ic_{\nu j}^x c_{\nu j}^y + 1}{2}, 0 \leq \theta < 2\pi\}, \quad (9)$$

and the particle-hole symmetry in layer ν

$$\mathcal{C}_\nu = \prod_j (ic_{\nu j}^y b_{\nu j}^\gamma), \quad (10)$$

where we can choose an arbitrary (but fixed) bond direction $1 \leq \gamma \leq 4$. Note that all these symmetries commute with the onsite constraint (7).

Next we take the inter layer interaction term (6) into account to analyze the symmetry of the full Hamiltonian (1). Note that the particle-hole symmetry (10) in each layer does not commute with the onsite constraint (7), and hence ceases to be a symmetry of the full Hamiltonian (1). Instead, the following particle-hole symmetry for both layers

$$\mathcal{C} \equiv \mathcal{C}_1 \mathcal{C}_2 = \prod_{j,\nu} (ic_{\nu j}^y b_{\nu j}^\gamma) \quad (11)$$

remain as a symmetry of the full Hamiltonian. Moreover, there is a layer exchange symmetry generated by

$$\mathcal{X} = \prod_i \left[\frac{1 + c_{1,i}^x c_{2,i}^x}{\sqrt{2}} \cdot \frac{1 + c_{1,i}^y c_{2,i}^y}{\sqrt{2}} \cdot \prod_\gamma \frac{1 + b_{1,i}^\gamma b_{2,i}^\gamma}{\sqrt{2}} \right] \quad (12)$$

As a result, the symmetry of the full Hamiltonian is given by

$$G = [U(1) \times U(1) \times \mathbb{Z}_2^{\mathcal{T}}] \times [\mathbb{Z}_2^{\mathcal{X}} \times \mathbb{Z}_2^{\mathcal{C}}] \quad (13)$$

The symmetry of the full Hamiltonian can be conveniently understood in the basis of complex fermions $f_{\nu,j}$ defined as

$$f_{\nu,j} \equiv \frac{c_{\nu,j}^x - ic_{\nu,j}^y}{2} \implies \{f_{\nu,i}, f_{\mu,j}^\dagger\} = \delta_{i,j} \delta_{\mu,\nu} \quad (14)$$

In this basis, the $U(1)$ symmetries for each layer can be recombined into “charge” $U(1)_c$ and pseudo-spin (i.e. layer) $U(1)_s$ rotations:

$$U(1)_c : f_{\nu,j} \rightarrow e^{i\theta} f_{\nu,j}; \quad \mathcal{C} f_{\nu,j} \mathcal{C}^{-1} = f_{\nu,j}^\dagger; \quad (15)$$

$$U(1)_s : f_{\nu,j} \rightarrow e^{i\kappa\theta} f_{\nu,j}; \quad \mathcal{X} f_{\nu,j} \mathcal{X}^{-1} = \kappa f_{\nu',j}, \quad (16)$$

where $\kappa = -1, 1$ for $\nu = 1, 2$ respectively and $\nu' = \nu + 1 \pmod{2}$. If we represent the pseudo-spin-1/2 complex fermions using the following matrix

$$F_j \equiv \begin{pmatrix} f_{1,j} & f_{2,j}^\dagger \\ f_{2,j} & -f_{1,j}^\dagger \end{pmatrix} \quad (17)$$

the pseudo-spin and charge rotations correspond to left and right rotations on the matrix F_j :

$$R(U_\theta^s) = e^{i\theta Z_L}; \quad R(\mathcal{X}) = iY_L; \quad (18)$$

$$R(U_\theta^c) = e^{i\theta Z_R}; \quad R(\mathcal{C}) = X_R. \quad (19)$$

where $\vec{\sigma}_{L/R} \equiv (X_{L/R}, Y_{L/R}, Z_{L/R})$ stands for Pauli matrices multiplied to the left/right of matrix F_j in (17). Now it is clear that the full symmetry group can be written as

$$G = O(2)_c \times O(2)_s \times \mathbb{Z}_2^T \quad (20)$$

where $O(2) = U(1) \times \mathbb{Z}_2$ is the rotation group for either the charge or pseudo-spin degrees of freedom.

B. Emergent symmetry for a fixed flux configuration

One type of gapped excitations in model (1) are the visons[?], i.e. π flux excitations in each plaquette. More precisely, Majorana fermions $\{c_{\nu,i}^{x,y}\}$ hop in a background flux, where the flux in the plaquette $\langle i, j, k, l \rangle$ is given by

$$F_{\nu,ijkl} = u_{\nu,ij} u_{\nu,jk} u_{\nu,kl} u_{\nu,li} = \pm 1 \quad (21)$$

In a fixed flux configuration where both layers share the same flux in every plaquette, we can always choose a gauge so that $u_{\nu,ij} \equiv u_{ij} = \pm 1$ are fixed as constants. Therefore the physical spectrum with fixed fluxes (i.e. without vison excitations) is given by a Hubbard model of Majorana fermions $\{c_{\nu,i}^{x,y}\}$ with a hopping strength K and a Hubbard interaction strength J . The effective symmetry for a fixed flux configuration is generated by the spinless time reversal symmetry

$$\mathcal{T} = \prod_{j,\nu} (i c_{\nu,j}^x c_{\nu,j}^y)^{j_x + j_y} \cdot \mathcal{K} \quad (22)$$

and the well-known unitary $SO(4) = SU(2) \times SU(2)/\mathbb{Z}_2$ symmetry of a 2-flavor Hubbard model[?], where the layer index ν plays the role of the pseudo-spin. We have used $j = (j_x, j_y) \in \mathbb{Z}^2$ to label a site on the square lattice. Therefore the emergent symmetry group for a fixed flux configuration is given by

$$G' = SO(4) \times \mathbb{Z}_2^T \quad (23)$$

It is most convenient to represent the symmetry in the $u_{\nu,ij} \equiv u_{ij} = \pm 1, \forall \nu = 1, 2$ gauge, where the $SO(4)$ symmetry is simply implemented by an $SO(4)$ unitary rotation on the Majorana basis

$$\psi_j \equiv (c_{1,j}^x, c_{1,j}^y, c_{2,j}^x, c_{2,j}^y)^T \quad (24)$$

since the fixed-flux Hamiltonian in this gauge is written as

$$\hat{H}_{\text{fixed flux}} = -K \sum_{\langle i,j \rangle} u_{ij} \sum_{\nu=1,2} \sum_{\alpha=x,y} i c_{\nu,i}^\alpha c_{\nu,j}^\alpha + J \sum_i c_{1,i}^x c_{1,i}^y c_{2,i}^x c_{2,i}^y \quad (25)$$

Meanwhile, under the spinless time reversal symmetry operation, the Majorana fermions (24) transform as

$$\psi_j \xrightarrow{\mathcal{T}} (-1)^j \psi_j \quad (26)$$

The $SO(4)$ symmetry can also be understood in the basis of complex fermions $f_{\nu,j}$ defined in (14), where $SU(2)_s$ pseudo-spin rotation and $SU(2)_c$ charge rotation are realized as left and right $SU(2)$ rotations on the matrix in (17):

$$F_j \equiv \begin{pmatrix} f_{1,j} & f_{2,j}^\dagger \\ f_{2,j} & -f_{1,j}^\dagger \end{pmatrix} \rightarrow M_s F_j M_c^\dagger, \quad M_{c,s} \in SU(2) \quad (27)$$

Physically, at an energy scale much smaller than the vison gap, vison excitations are suppressed and the low energy subspace has a fixed flux configuration. As proven by Lieb[?], a uniform π flux per plaquette has the lowest energy and the low energy physics is described by Majoranas ψ_j hopping in a uniform background of π flux per plaquette. Therefore, the effective (or emergent) symmetry of the system below the vison excitation gap is exactly given by the symmetry of a fixed flux configuration as we described above.

C. Remnant symmetry of the Néel order

In a Hubbard model on the square lattice, as the onsite repulsive Hubbard interaction strength J/K increases beyond a critical value J_c/K , the system enters a Néel ordered phase with the long-range antiferromagnetic order. Now that we have obtained the symmetry of the full Hamiltonian (1) and of any fixed flux configuration, next we analyze the remnant (unbroken) symmetry of the Néel ordered ground state in our model (1).

The Néel order in the Hubbard model is characterized by the staggered magnetization as its order parameter, represented by the Néel vector:

$$\mathbf{n} = \frac{1}{N} \sum_j \mathbf{n}_j, \quad \mathbf{n}_j = (-1)^{j_x+j_y} \langle f_{\mu j}^\dagger \boldsymbol{\sigma}_{\mu\nu} f_{\nu j} \rangle = (-1)^{j_x+j_y} \text{Tr}(F_j^\dagger \boldsymbol{\sigma} F_j) \quad (28)$$

Note that the Néel vector always changes sign when transformed under the spinless time reversal symmetry $\mathcal{T} = \mathcal{K}$, therefore it always spontaneously breaks the antiunitary time reversal symmetry \mathcal{T} . From (28) it is clear that the Néel vector is invariant under charge rotations $M_c \in SU(2)_c$ in (14), therefore it always preserves the $O(2)_c$ subgroup of “charge” symmetries of the full Hamiltonian (1), or the $SU(2)_c$ subgroup of the emergent $G' = SO(4) \times \mathbb{Z}_2^{\mathcal{T}}$ symmetry for a fixed flux configuration. In fact, independent of the Néel vector orientation, in a fixed flux configuration, the Néel order always spontaneously breaks the emergent symmetry group $G' = SO(4) \times \mathbb{Z}_2^{\mathcal{T}}$ down to a subgroup

$$H' = SU(2)_c \times U(1)_{\mathbf{n}} \rtimes \mathbb{Z}_2^{\mathcal{T}_{\mathbf{n}}} \quad (29)$$

where in the complex fermion basis (17) we have defined

$$U(1)_{\mathbf{n}} = \{e^{i\theta \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}_L} | 0 \leq \theta < \pi\}, \quad R(\mathcal{T}_{\mathbf{n}}) \equiv (i \hat{\mathbf{m}} \cdot \boldsymbol{\sigma}_L) \cdot Y_L Y_R \cdot \mathcal{K} \quad (30)$$

Here $\hat{\mathbf{m}} \parallel \hat{z} \times \mathbf{n}$ is an in-plane unit vector perpendicular to the unit Néel vector $\hat{\mathbf{n}}$.

Below we analyze three cases with different orientations of the Néel vector \mathbf{n} , where the symmetry group G in (20) of the full Hamiltonian is spontaneously broken down to different subgroups.

(1) If the Néel vector \mathbf{n} is parallel to the \hat{z} axis, the $U(1)_s$ symmetry is preserved while the layer exchange \mathcal{X} is broken. As a result, the remnant symmetry group is

$$H(\mathbf{n} \parallel \hat{z}) = O(2)_c \times U(1)_s \rtimes \mathbb{Z}_2^{\mathcal{X} \cdot \mathcal{T}} \quad (31)$$

The associated order parameter manifold is given by

$$\mathcal{M} = G/H = \mathbb{Z}_2 \quad (32)$$

(2) If the Néel vector \mathbf{n} lies in the plane perpendicular to the \hat{z} axis, the $U(1)_s$ spin rotation is broken while a layer exchange symmetry $\mathcal{X}_{\mathbf{n}}$ defined by

$$R(\mathcal{X}_{\mathbf{n}}) = i \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}_L \quad (33)$$

in the complex fermion basis (17) is preserved. As a result, the remnant symmetry group is

$$H(\mathbf{n} \perp \hat{z}) = O(2)_c \times \mathbb{Z}_2^{\mathcal{X}_{\mathbf{n}}} \times \mathbb{Z}_2^{\mathcal{T}} \quad (34)$$

Néel vector	Unbroken subgroup H	G/H	Topological defects	Gauge group
$\mathbf{n} \parallel \hat{z}$	$O(2)_c \times U(1)_s \times \mathbb{Z}_2^{\mathcal{X} \cdot \mathcal{T}}$	\mathbb{Z}_2	Domain wall	$\mathbb{Z}_2 \times \mathbb{Z}_2$
$\mathbf{n} \perp \hat{z}$	$O(2)_c \times \mathbb{Z}_2 \times \mathbb{Z}_2^{\mathcal{T}}$	S^1	Vortex	\mathbb{Z}_2
$n_z \neq 0, n_x n_y \neq 0$	$O(2)_c \times \mathbb{Z}_2^{\mathcal{T}^n}$	$O(2)$	Domain wall + Vortex	\mathbb{Z}_2

TABLE I. A summary of distinct long range orders for different orientations of the Néel vector \mathbf{n} , and the associated gauge group for their topological orders.

where the preserved anti-unitary symmetry is defined as

$$R(\tilde{\mathcal{T}}) = X_L Y_R \cdot \mathcal{K} \quad (35)$$

As a result, the associated order parameter manifold is

$$\mathcal{M} = G/H = U(1) \simeq S^1 \quad (36)$$

(3) If the Néel vector has both in-plane and \hat{z} components, both $U(1)_s$ rotation and layer exchange \mathcal{X} are broken, leaving anti-unitary \mathcal{T}_n defined in (30) an unbroken symmetry. Therefore the remnant symmetry group is

$$H(n_z \neq 0, n_x n_y \neq 0) = O(2)_c \times \mathbb{Z}_2^{\mathcal{T}^n} \quad (37)$$

and the associated order parameter manifold is

$$\mathcal{M} = G/H = O(2) \quad (38)$$

A summary of the long range orders and associated topological defects of the local order parameters is shown in Table I.

D. Local and string order parameters for the long range Néel order

According to Lieb's theorem[?], for the model (25) of interacting fermions with bipartite hopping and particle-hole symmetry, a uniform π flux per plaquette has the lowest energy. As a result, in the low-energy manifold below the vison gap, we can consider a π -flux Hubbard model of fermions $\{f_{\nu,j}\}$ on the square lattice. To understand the topological nature of the ground state and its excitations, we need to understand the wavefunctions of the bilayer spin liquid model.

In the low-energy manifold, since the flux (21) in every plaquette is fixed to $F_{\nu,ijkl} \equiv -1, \forall \nu$, we can choose a gauge so that

$$i b_i^{\gamma ij} b_j^{\gamma ij} = u_{ij}, \quad \forall \nu = 1, 2 \quad (39)$$

where u_{ij} is fixed as shown in Fig. 1. In the fixed gauge, let's denote the eigenstates for the Hubbard model (25) of fermions $\{f_{\nu,j}\}$ (or equivalently of $\{\psi_j\}$) as $|\psi\rangle$. Then the mean-field ansatz $|MF\rangle$ of all fermions $\{c_{\nu,j}^{x,y}, b_{\nu,j}^{\gamma}\}$ can be written as

$$|MF\rangle = |\psi\rangle \otimes |i b_i^{\gamma ij} b_j^{\gamma ij} = u_{ij}\rangle \quad (40)$$

where $|i b_i^{\gamma ij} b_j^{\gamma ij} = u_{ij}\rangle$ denotes the unique eigenstate of fermions $\{b_{\nu,j}^{\gamma}\}$ that satisfies the fixed gauge (39). The physical eigenstate of the bilayer spin liquid model (1) is then obtained by enforcing the onsite constraint (7) on the mean-field ansatz:

$$|QSL\rangle = \prod_{j,\nu} \frac{1 + D_{\nu,j}}{2} |MF\rangle \quad (41)$$

via the projection operator $\frac{1 + D_{\nu,j}}{2}$ on every site.

In the Néel order of the square-lattice Hubbard model, there is an off-diagonal long range order for the Néel vector (28) in the mean-field ansatz:

$$\lim_{|i-j| \rightarrow \infty} (-1)^{i_x + i_y + j_x + j_y} \langle MF | f_{\mu,i}^{\dagger} \vec{\sigma}_{\mu\nu} f_{\nu,i} \cdot f_{a,j}^{\dagger} \vec{\sigma}_{ab} f_{b,j} | MF \rangle \neq 0 \quad (42)$$

However, this does not guarantee the same local order parameter to exist in the physical ground state (41). More precisely, while the \hat{z} -component of the Néel vector

$$n_j^z \propto \sum_{\mu=1,2} \mu f_{\mu,j}^\dagger f_{\mu,j} = \sum_{\nu} \kappa \Gamma_{\nu,j}^5 \quad (43)$$

where $\kappa = -1, 1$ for $\nu = 1, 2$ respectively. n_j^z is a local operator and, the in-plane components

$$n_j^+ = n_j^x + i n_j^y \propto f_{2,j}^\dagger f_{1,j} \quad (44)$$

do not commute with the onsite constraint (7). As a result, the off-diagonal long range order of local order parameters in the mean-field ansatz will translate into a string order parameter in the physical wavefunction:

$$\lim_{|i-j| \rightarrow \infty} (-1)^{i_x+i_y+j_x+j_y} \langle QSL | f_{2,i}^\dagger f_{1,i} \left(\prod_{\nu} \prod_{a=1}^M i b_{\nu,j_a}^{\gamma_{j_a,j_{a+1}}} b_{\nu,j_{a+1}}^{\gamma_{j_a,j_{a+1}}} \right) f_{1,j}^\dagger f_{2,j} | QSL \rangle \neq 0 \quad (45)$$

where (j_1, j_2, \dots, j_M) is a string of consecutive sites that connect $j_1 = i$ and $j_M = j$. It is straightforward to check that the string order parameter above commutes with the onsite constraint (7). Although the in-plane component n^+ of the Néel vector is not a local operator, there is a local order parameter for the in-plane Néel order:

$$S_{\langle ij \rangle}^+ \equiv f_{2,i}^\dagger f_{1,i} \left(\prod_{\nu} i b_{\nu,i}^{\gamma_{i,j}} b_{\nu,j}^{\gamma_{i,j}} \right) f_{2,j}^\dagger f_{1,j} \quad (46)$$

where $\langle i, j \rangle$ is a pair of nearest neighbor sites. And one can check that the in-plane Néel order in the mean-field ansatz $|MF\rangle$ implies the following long range order of the physical wavefunction $|QSL\rangle$:

$$\lim_{|i-i'| \rightarrow \infty} \langle QSL | S_{\langle ij \rangle}^+ S_{\langle i'j' \rangle}^- | QSL \rangle \neq 0 \quad (47)$$

E. Topological orders in the Néel ordered phase

For a small value of J/K , the π -flux Hubbard model (25) has a gapless spectrum with massless Dirac fermion excitations. Each layer hosts one type of vison (i.e. π -flux) excitations, giving rise to a $\mathbb{Z}_2 \times \mathbb{Z}_2$ gauge field. The low energy physics is therefore described by the two decoupled layers of gapless \mathbb{Z}_2 spin liquids, each layer described by two branches of massless Dirac fermions coupled to a \mathbb{Z}_2 gauge field.

In the large J/K case, a long-range Néel order develops in the ground states of the π -flux Hubbard model (25), where the Dirac fermions acquires a mass, giving rise to a gapped spectrum for all fermions. Different orientations of the Néel vector (28) not only give rise to distinct long range orders with different order parameters, but also distinct topological orders with different gauge groups. Below we analyze the 3 different orientations of the Néel vector summarized in Table I.

1. $\mathbf{n} \parallel \hat{z}$

As previously discussed, when the Néel vector aligns along \hat{z} -axis, the \hat{z} -component of the Néel vector is a local order parameter for the Ising-type Néel order, with a order parameter manifold of $G/H = \mathbb{Z}_2$. Since the fermions are all gapped, the ground state exhibits a $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order, due to deconfined vison excitations from each layer.

The only gapless excitations in this case are the Goldstone modes of the Néel order, described by the spin flip operator in (44). However, the operator $n_j^+ \sim (-1)^{j_x+j_y} f_{2,j}^\dagger f_{1,j}$ is not a gauge invariant operator as it violates the onsite constraint (7). The gapless Goldstone modes of the Hubbard model lead to a power-law decaying correlation of the string order parameter (45) in this case.

If the system is gapped, it will naively be an Abelian $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order described by a 4-component Chern-Simons theory with the following \mathbf{K} matrix

$$\mathbf{K} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \quad (48)$$

Since both $U(1)_c$ and $U(1)_s$ symmetries are preserved, they are captured by charge and pseudo-spin vectors

$$\mathbf{q}_c = (2, 0, 2, 0)^T, \quad \mathbf{q}_s = (2, 0, -2, 0)^T. \quad (49)$$

However, the nature of the gapless ‘‘Goldstone modes’’ here, plays a crucial role to interpret the physical ground state (41) after the projection $\frac{1+D_{\nu,i}}{2}$ on each site. Notice, the two real Goldstone modes together form a complex bosonic mode

$$n^x + in^y \sim b_{\tilde{\mathbf{k}}=(\pi,\pi)} \sim \sum_i (-1)^{i_x+i_y} f_{2,i}^\dagger f_{1,i} \quad (50)$$

which is gapless, with a linear dispersion. This mode carries a unit gauge charge of both \mathbb{Z}_2 gauge fields, one from each layer, and is hence not a local excitation. Therefore, this ground state should be understood as an algebraic $\mathbb{Z}_2 \times \mathbb{Z}_2$ spin liquid, where the bound state (50) (with quasiparticle vector $(1, 1, -1, 1)^T$ in the Abelian Chern-Simons theory) of two types of fermionic spinons f_2 (with quasiparticle vector $(1, 1, 0, 0)^T$) and f_1 (with quasiparticle vector $(0, 0, -1, 1)^T$) becomes gapless with an algebraic correlation in (45). The mode b obeys bosonic self statistics, but is really an anyon excitation with semionic mutual braiding with fermions $f_{\nu,j}$ from either layer. Therefore the system exhibits fractionalized Goldstone modes that obey anyonic statistics.

After incorporating the gapless anyon b in (50) into the low energy description, the effective field theory for this algebraic spin liquid writes

$$\mathcal{L}_{\text{ASL}} = \sum_{I,J} \frac{\epsilon^{\mu\nu\rho}}{4\pi} a_\mu^I \mathbf{K}_{I,J} \partial_\nu a_\rho^J - \sum_{\alpha=c,s} \sum_I \frac{\epsilon^{\mu\nu\rho}}{2\pi} A_\mu^\alpha \mathbf{q}_I^\alpha \partial_\nu a_\rho^I + |(-i\partial_\mu - 2A_\mu^s - a_\mu^1 - a_\mu^2 + a_\mu^3 - a_\mu^4)^2 b|^2 + \dots \quad (51)$$

where b denotes the critical anyon mode out of the Goldstone modes of the broken emergent symmetry. Typically, the robust presence of a gapless anyon excitation in a system requires a mechanism to protect it. In this case, it is the emergent pseudo-spin $SU(2)_s$ symmetry that protects the criticality of this anyon mode.

2. $\mathbf{n} \perp \hat{z}$

It’s straightforward to show that a nonzero in-plane Néel component $\langle b \rangle \neq 0$ corresponds to condensing the anyon characterized by vector $(1, 1, -1, 1)$ in the Chern-Simons theory, and drives the system into a usual \mathbb{Z}_2 topological order described by a 2-component Abelian Chern-Simons theory with $\mathbf{K} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$.

This can be understood by considering the topological defects of the local order parameter (46): since the order parameter manifold for an in-plane Néel order ($\mathbf{n} \perp \hat{z}$) is $G/H = U(1) \simeq S^1$, the long-range order supports point defects (i.e. vortices) classified by an integer-valued vorticity $\pi_1(G/H) = \mathbb{Z}$. The vortex of local order parameter $S^+ = S^x + iS^y$ in (46) is nothing but a π flux from either layer, characterized by quasiparticle vector $(1, 0, 0, 0)$ or $(0, 0, 1, 0)$ in the 4-component Abelian Chern-Simons theory. Since the vortices are logarithmically confined in two spatial dimensions, the π flux from either layer is confined with an in-plane Néel order. As a result, the ground state features a \mathbb{Z}_2 topological order, whose vison excitation is the bound state of a π flux from both layers.

Next we describe the Goldstone modes in the in-plane Néel order. Without loss of generality, we assume the Néel vector to point along \hat{x} axis in the ground state, i.e. $\langle n^x \rangle \neq 0$ in the ground state of the gauge-fixed π -flux Hubbard model (25). In this case, we can write down the following local operator that creates the Goldstone mode n^y :

$$\mathcal{N}^y = \sum_i n_i^x \prod_\nu (i b_{\nu,i}^\gamma b_{\nu,i+\hat{e}_\gamma}^\gamma) n_{i+\hat{e}_\gamma}^y \quad (52)$$

where \hat{e}_γ is the unit lattice vector along bond γ direction. Due to the long-range string order (45) in the ground state, it is straightforward to show that

$$\mathcal{N}^y |QSL\rangle \sim \left(\prod_{\nu,i} \frac{1+D_{\nu,i}}{2} \right) n^y |MF\rangle \quad (53)$$

3. $n_z \neq 0, n_x n_y \neq 0$

When the Néel vector has both in-plane and \hat{z} components, the ground state also exhibits a \mathbb{Z}_2 topological order, for the same reason as discussed above. With an order parameter manifold of $G/H = O(2)$, the topological defects in this case include both domain walls classified by $\pi_1(O(2)) = \mathbb{Z}_2$, and vortices classified by $\pi_1(O(2)) = \mathbb{Z}$.

IV. EFFECTIVE $\eta - \rho$ HAMILTONIAN IN THE LARGE- J LIMIT

In this section, we provide the details of the derivation for the effective Hamiltonian to fourth-order in $|K/J| \ll 1$, which we label as an “ $\eta - \rho$ ” model. For $K = 0$, the ground state (GS) is doubly degenerate, $|\uparrow_1\uparrow_2\rangle, |\downarrow_1\downarrow_2\rangle$ at each site. These doublets can be represented by the pseudospins $|\uparrow\rangle, |\downarrow\rangle$ with associated operators

$$\begin{aligned}\eta_i^z &= \frac{1}{4}(\sigma_{1i}^z + \sigma_{2i}^z) \\ \eta_i^\pm &= \frac{1}{4}\sigma_{1i}^\pm\sigma_{2i}^\pm\end{aligned}\quad (54)$$

which obey an SU(2) algebra. Note that η_i^z is a dipolar operator whereas η_i^x and η_i^y ,

$$\begin{aligned}\eta_i^x &= \frac{1}{4}(\sigma_{1i}^x\sigma_{2i}^x - \sigma_{1i}^y\sigma_{2i}^y) \\ \eta_i^y &= \frac{1}{4}(\sigma_{1i}^x\sigma_{2i}^y + \sigma_{1i}^y\sigma_{2i}^x)\end{aligned}\quad (55)$$

are quadrupole operators. Apart from the pseudospin doublet, there is an additional four-fold degeneracy per site, arising from the orbital DOF. We introduce an operator

$$P_0 = \prod_i \frac{1 + \sigma_{1i}^z\sigma_{2i}^z}{2}, \quad (56)$$

which projects out the states $|\uparrow_1\downarrow_2\rangle, |\downarrow_1\uparrow_2\rangle$ at each site.

In the following, we derive the effective Hamiltonian perturbatively to fourth-order in K . We also note that all odd-ordered terms vanish since H_K only connects states inside the GS manifold to states outside of it.

A. Second-order effective Hamiltonian H_{g_2}

The effective Hamiltonian to second order is

$$H^{(2)} = P_0 H_K S H_K P_0 \quad (57)$$

where

$$S = \frac{(1 - P_0)}{(E_0 - H_J)}, \quad (58)$$

and H_J is the inter-layer coupling. In order to evaluate eq. 57, we introduce the operators

$$X_{ij} = \sigma_{1i}^+ \sigma_{1j}^- \tau_{1i}^\gamma \tau_{1j}^\gamma + \sigma_{2i}^- \sigma_{2j}^+ \tau_{2i}^\gamma \tau_{2j}^\gamma \quad (59)$$

and re-write $H_K = \sum_{\langle ij \rangle_\gamma} X_{ij} + X_{ji}$. Since $P_0 X_{ij}^2 P_0 = 0$, we obtain

$$H^{(2)} = \frac{K^2}{4J} \sum_{\langle ij \rangle_\gamma} P_0 \{X_{ij}, X_{ji}\} P_0. \quad (60)$$

Furthermore, we use the identities

$$\begin{aligned}P_0 \sigma_{1i}^+ \sigma_{2i}^+ P_0 &= 4\eta_i^+ \\ P_0 \sigma_{1i}^- \sigma_{2i}^- P_0 &= 4\eta_i^- \\ P_0 \sigma_{\nu i}^+ \sigma_{\nu i}^- P_0 &= 2(1 + 2\eta_i^z) \\ P_0 \sigma_{\nu i}^- \sigma_{\nu i}^+ P_0 &= 2(1 - 2\eta_i^z),\end{aligned}\quad (61)$$

to express

$$H^{(2)} = -g_2 \sum_{\langle ij \rangle_\gamma} \left[-\eta_i^z \eta_j^z + \frac{1}{2}(\eta_i^+ \eta_j^- + \eta_i^- \eta_j^+) \tau_{1i}^\gamma \tau_{1j}^\gamma \tau_{2i}^\gamma \tau_{2j}^\gamma \right], \quad (62)$$

where $g_2 = K^2/4J$. We perform rotations by π about the z -axis on the η pseudo-spins at every other site to obtain the effective Hamiltonian at g_2 level

$$H_{g_2} = g_2 \sum_{\langle ij \rangle_\gamma} \left[\eta_i^z \eta_j^z + \frac{1}{2} (\eta_i^+ \eta_j^- + \eta_i^- \eta_j^+) \tau_{1i}^\gamma \tau_{1j}^\gamma \tau_{2i}^\gamma \tau_{2j}^\gamma \right]. \quad (63)$$

Note that $\tau_{1i}^\gamma \tau_{1j}^\gamma \tau_{2i}^\gamma \tau_{2j}^\gamma = \rho_{ij}^\gamma$, where $\rho_{ij}^\gamma = q_i^\gamma q_j^\gamma$.

B. Fourth-order effective Hamiltonian H_{g_4}

The general expression to fourth order reads

$$H^{(4)} = P_0 H_K S H_K S H_K S H_K P_0 - \frac{1}{2} (P_0 H_K S^2 H_K P_0 H_K S H_K P_0 + P_0 H_K S H_K P_0 H_K S^2 H_K P_0) \quad (64)$$

We focus on the first term, which, as described below, includes coupled pseudo-spin and orbital DOF around the p/p' plaquettes. These play a critical role in lifting the extensive degeneracy of the GS of the leading H_{g_2} , due to the orbital DOF, as implied by $\rho_{ij}^\gamma = 1 \forall \gamma, \langle ij \rangle$ (Supplementary section VI). By contrast, the remaining terms in eq. 64 are either independent of the orbital DOF or involve two pairs of NN pseudo-spins coupled to $\rho_{ij}^\gamma \rho_{jk}^{\gamma'}$, where $\gamma \neq \gamma'$ for the two connected bonds. These involve subleading corrections to the pseudo-spin configuration, but do not lift the GS degeneracy of H_{g_2} . Consequently, we ignore these contributions and focus instead on the plaquette terms in the following. To illustrate, we consider a single p plaquette (Supplementary figure 1) in the first term as

$$P_0 H_K S H_K S H_K S H_K P_0 = \frac{K^4}{8J^3} P_0 [-2\{X_{ij} X_{kl}, X_{jk} X_{li}\} - \{\{X_{ij}, X_{jk}\}, \{X_{kl}, X_{li}\}\} - \{\{X_{jk}, X_{kl}\}, \{X_{li}, X_{ij}\}\}] P_0 + \dots \quad (65)$$

where \dots includes all remaining p/p' plaquettes. We simplify the expressions by using the identities in eq. 61 and by rotating the pseudo-spins such that $\eta_i^\pm \rightarrow -\eta_i^\pm$ at every other site. Carrying out similar operations for the remaining plaquettes, we obtain the effective Hamiltonian at fourth order

$$H_{g_4} = g_4 \sum_p [\eta_{ijkn}^\square (W_{1p} + W_{2p}) + \eta_{klmn}^\square (W_{1p'} + W_{2p'})], \quad (66)$$

where

$$\begin{aligned} \eta_{ijkn}^\square = & [5\eta_i^z \eta_j^z \eta_k^z \eta_n^z + 5/2(\eta_i^+ \eta_j^- \eta_k^+ \eta_n^- + \eta_i^- \eta_j^+ \eta_k^- \eta_n^+) \\ & - 5/2(\eta_i^+ \eta_j^- \eta_k^z \eta_n^z \rho_{ij} + 11 \text{ permutations}) + \\ & [(\eta_i^z \eta_j^z + 1/2(\eta_i^+ \eta_j^- + \eta_i^- \eta_j^+) \rho_{ij}) + \\ & (\eta_j^z \eta_k^z + 1/2(\eta_j^+ \eta_k^- + \eta_j^- \eta_k^+) \rho_{jk}) + \\ & (\eta_k^z \eta_n^z + 1/2(\eta_k^+ \eta_n^- + \eta_k^- \eta_n^+) \rho_{kn}) + \\ & (\eta_n^z \eta_i^z + 1/2(\eta_n^+ \eta_i^- + \eta_n^- \eta_i^+) \rho_{ni}) + \\ & (\eta_i^z \eta_k^z + 1/2(\eta_i^+ \eta_k^- + \eta_i^- \eta_k^+) \rho_{ik}) + \\ & (\eta_j^z \eta_n^z + 1/2(\eta_j^+ \eta_n^- + \eta_j^- \eta_n^+) \rho_{jn}) + 1]. \end{aligned}$$

Note that these are due to the terms in eq. 65.

V. THE EFFECTS OF $\tau_{\nu,i}$ AND $W_{\nu p/p'}$ OPERATORS IN THE $\eta - \rho$ BASIS

We first consider $\tau_{\nu,i}$ operating on any local orbital state determined by the set $\{q_i^x, q_i^y, q_i^z\}$ subject to the constraint $q_i^x q_i^y q_i^z = -1$:

$$\begin{aligned}
\tau_1^x |-, -, -\rangle &= -|-, +, +\rangle \\
\tau_2^x |-, -, -\rangle &= +|-, +, +\rangle \\
\tau_1^y |-, -, -\rangle &= i|+, -, +\rangle \\
\tau_2^y |-, -, -\rangle &= -i|+, -, +\rangle \\
\tau_1^z |-, -, -\rangle &= +|+, +, -\rangle \\
\tau_2^z |-, -, -\rangle &= -|+, +, -\rangle \\
\tau_1^x |-, +, +\rangle &= -|-, -, -\rangle \\
\tau_2^x |-, +, +\rangle &= +|-, -, -\rangle \\
\tau_1^y |-, +, +\rangle &= i|+, +, -\rangle \\
\tau_2^y |-, +, +\rangle &= i|+, +, -\rangle \\
\tau_1^z |-, +, +\rangle &= +|+, -, +\rangle \\
\tau_2^z |-, +, +\rangle &= +|+, -, +\rangle \\
\tau_1^x |+, -, +\rangle &= +|+, +, -\rangle \\
\tau_2^x |+, -, +\rangle &= +|+, +, -\rangle \\
\tau_1^y |+, -, +\rangle &= -i|-, -, -\rangle \\
\tau_2^y |+, -, +\rangle &= +i|-, -, -\rangle \\
\tau_1^z |+, -, +\rangle &= +|-, +, +\rangle \\
\tau_2^z |+, -, +\rangle &= +|-, +, +\rangle \\
\tau_1^x |+, +, -\rangle &= +|+, -, +\rangle \\
\tau_2^x |+, +, -\rangle &= +|+, -, +\rangle \\
\tau_1^y |+, +, -\rangle &= -i|-, +, +\rangle \\
\tau_2^y |+, +, -\rangle &= -i|-, +, +\rangle \\
\tau_1^z |+, +, -\rangle &= +|-, -, -\rangle \\
\tau_2^z |+, +, -\rangle &= +|-, -, -\rangle
\end{aligned} \tag{67}$$

In summary, $\tau_{1/2,i}^x$ preserve q_i^x while changing the signs of $q_i^{y/z}$, and so on for all permutations. In terms of the $\rho_{ij}^\gamma = q_i^\gamma q_j^\gamma$ bonds, the same operator changes the signs of ρ_{ij}^y and ρ_{ik}^z , where ij and ik are NN pairs.

The flux operators are defined by

$$W_{\nu p} = \sigma_{\nu k}^z \sigma_{\nu n}^z \tau_{\nu i}^x \tau_{\nu j}^y \tau_{\nu k}^x \tau_{\nu n}^y \tag{68}$$

$$W_{\nu p'} = \sigma_{\nu k}^z \sigma_{\nu n}^z \tau_{\nu l}^x \tau_{\nu m}^y \tau_{\nu k}^x \tau_{\nu n}^y, \tag{69}$$

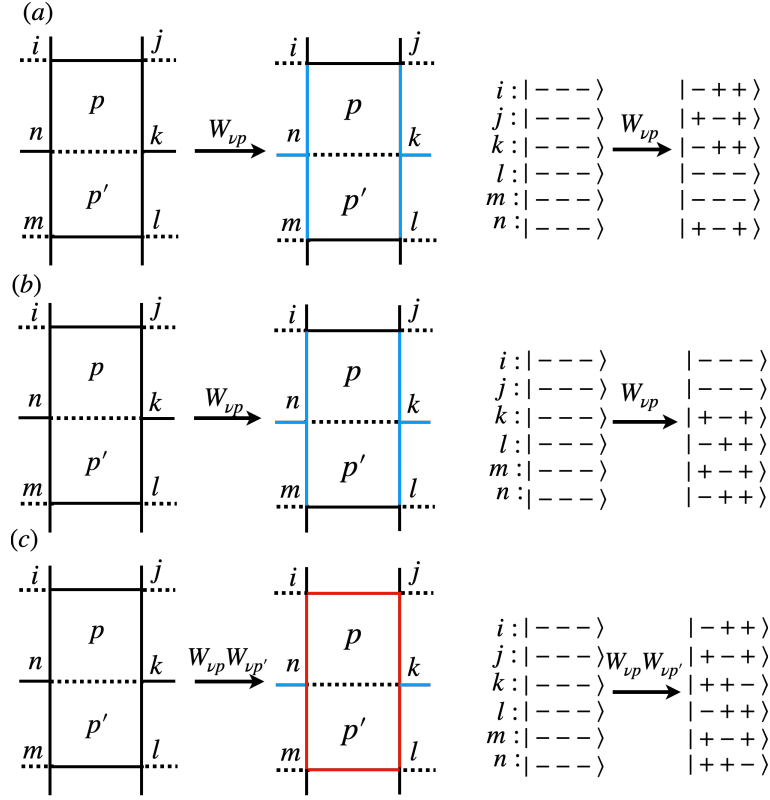
where the index convention is illustrated in Supplementary figure 1. These operators act both on the spin (σ), as well as on the orbital (τ) DOF. The labels p, p' label the plaquettes, as illustrated in Supplementary figure 1. By applying eq. 68 around a single p plaquette, we see that W_{1p} flips the q^γ 's at four of the six sites of the unit cell. Consequently, six ρ_{ij}^γ bonds change signs, as shown in Supplementary figure 2 (a). The effect of $W_{1p'}$ is similar (Supplementary figure 2 (b)) although the end state differs in the q^γ values at each of the six sites. Due to the presence of the σ_k^z and σ_n^z operators, both W_{1p} and $W_{1p'}$ also rotate the in-plane pseudo-spin components at sites k and n as $\eta_k^{x,y} \rightarrow -\eta_k^{x,y}$ and $\eta_n^{x,y} \rightarrow -\eta_n^{x,y}$.

VI. EXACT DIAGONALIZATION OF H_{g_2}

The effective Hamiltonian

$$H_{g_2} = g_2 \left[\sum_{\langle ij \rangle} \eta_i^z \eta_j^z + \sum_{\langle ij \rangle_\gamma} (\eta_i^x \eta_j^x + \eta_i^y \eta_j^y) \rho_{ij}^\gamma \right] + \sum_i (-1)^{i_x + i_y} (h_x \eta_i^x + h_z \eta_i^z), \tag{70}$$

ouples the in-plane pseudo-spin components $\eta_i^{x/y}$ with the orbital bonds $\rho_{ij}^\gamma = q_i^\gamma q_j^\gamma$. The latter are preserved and H_{g_2} can be solved for the pseudo-spins in any given bond configuration which is consistent with the local constraint



Supplementary figure 2. The effects of the flux operators on the q_i^γ orbital states and ρ_{ij}^γ bond configurations. Black lines indicate positive bonds with all $q_i^\gamma = -1$ around the unit cell. Blue lines indicate negative bonds, while red lines are positive bonds with non-trivial q_i^γ values. The right panels show how the latter transform. We illustrate for (a) W_{1p} , (b) $W_{1p'}$, and (c) $W_{1p}W_{1p'}$.

$\prod_\gamma q_i^\gamma = -1$. Here, we consider H_{g_2} on a 4×4 torus with both $h_{x/z} = 0$. We determine the GS energy via exact diagonalization in the pseudo-spin sector for various bond configurations. We find that the uniform $\rho_{ij}^\gamma = 1$ configuration has the lowest energy. In this sector, H_{g_2} maps onto a 2D Heisenberg model with periodic boundary conditions. In contrast, all states with one or more strings of $\rho_{ij}^\gamma = -1$ bonds have higher energy, as shown in Supplementary figure 3.

Note that exceptions to these conclusions occur for $h_z > 0$, $h_x = 0$, or when the pseudo-spin spontaneously orders along the z direction only, in the infinite-size system. As shown in Supplementary section VIII, a non-contractible string of $\rho_{ij}^\gamma = -1$ bonds becomes degenerate with the uniform $\rho_{ij}^\gamma = 1$ state.

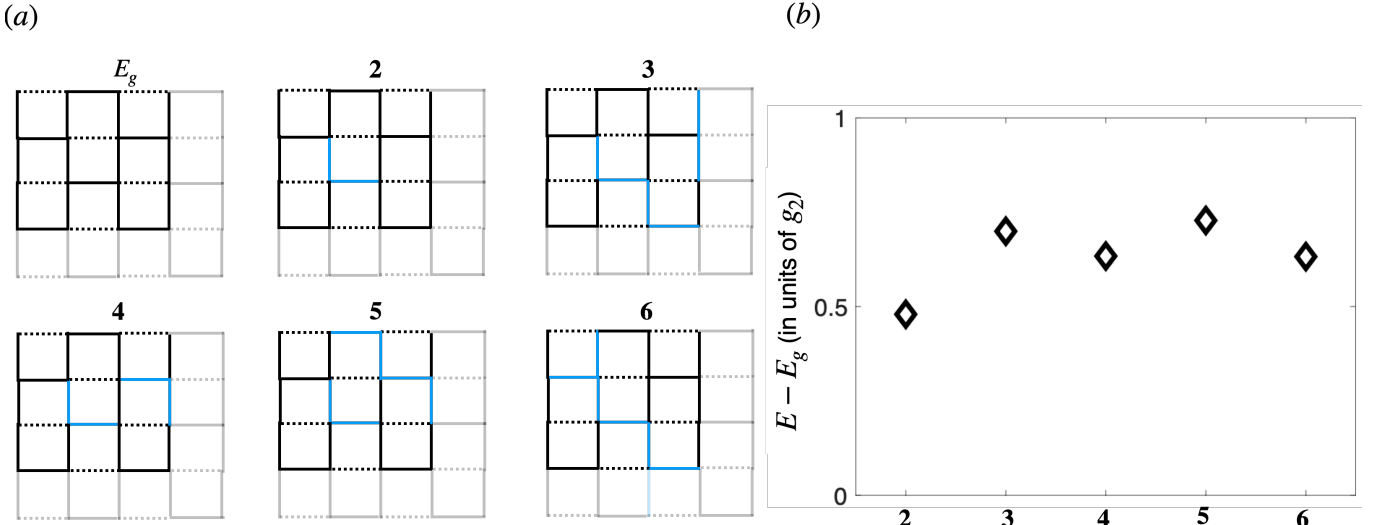
VII. EXACT DIAGONALIZATION OF H_{g_4}

As discussed in Supplementary section VI, a state with uniform $\rho_{ij}^\gamma = 1$ bonds and zero net pseudo-spin, as in a 2D Heisenberg model, has the lowest energy at H_{g_2} level. For $h_x = 0$, H_{g_2} commutes with the flux operators $W_{\nu p/p'}$. Therefore, the states

$$\prod_{\{\nu, p\}, \{\nu', p'\}} W_{\nu p} W_{\nu' p'} |\eta_{\text{Heisenberg}}, \forall \rho_{ij}^\gamma = 1\rangle, \quad (71)$$

where the ket is the GS at H_{g_2} level, are degenerate with the latter. The products can include any combinations of $W_{\nu p/p'}$ operators and this implies an extensive degeneracy of the GS. For $h_x > 0$, H_{g_2} only commutes with $W_{\nu, p} W_{\nu, p'}$, where p/p' are in the same unit cell, such that the degeneracy is partially lifted. We shall assume a small h_x in the following and ignore this additional complication.

H_{g_4} lifts the extensive degeneracy. To see how, we first project onto states of definite flux as



Supplementary figure 3. Bond configurations and ground-state energy for a 4×4 lattice on a torus. (a) Dotted bonds correspond to the identity bond while gray bonds illustrate wrapping around the torus. The configurations shown correspond to (1) Uniform $\rho_{ij}^\gamma = 1$, and (2-6) strings of negative bonds. (b) Difference in the ground-state energies corresponding to (a). The uniform configuration is lowest in energy.

$$|\eta; \{a_{\nu p}, a_{\nu p'}\}\rangle = \prod_{\nu, p, p'} \frac{1}{2^4} (1 + a_{\nu p} W_{\nu p}) (1 + a_{\nu p'} W_{\nu p'}) |\eta_{\text{Heisenberg}}, \forall \rho_{ij}^\gamma = 1\rangle, \quad a_{\nu p/p'} = \pm 1. \quad (72)$$

$|\eta; \{a_{\nu p}, a_{\nu p'}\}\rangle$ are eigenstates of $W_{\nu p/p'}$ with eigenvalue $a_{\nu p}$, corresponding to a $\pi(1 - a_{\nu p})/2$ flux per plaquette p in layer ν , and similarly for p' . We can consider the effect of H_{g_4} on the degenerate GS manifold of H_{g_2} . The former commutes with the fluxes and its matrix elements are diagonal in the projected basis:

$$\langle \eta; \{a_{\nu p}, a_{\nu p'}\} | H_{g_4} | \eta; \{b_{\nu p}, b_{\nu p'}\} \rangle = \frac{g_4}{2^8} \left[\sum_{\nu, p} \langle \eta_p^\square \rangle a_{\nu p} + \sum_{\nu, p'} \langle \eta_{p'}^\square \rangle a_{\nu p'} \right] \delta_{\{a_{\nu p}\}, \{b_{\nu p}\}} \delta_{\{a_{\nu p'}\}, \{b_{\nu p'}\}}, \quad (73)$$

where

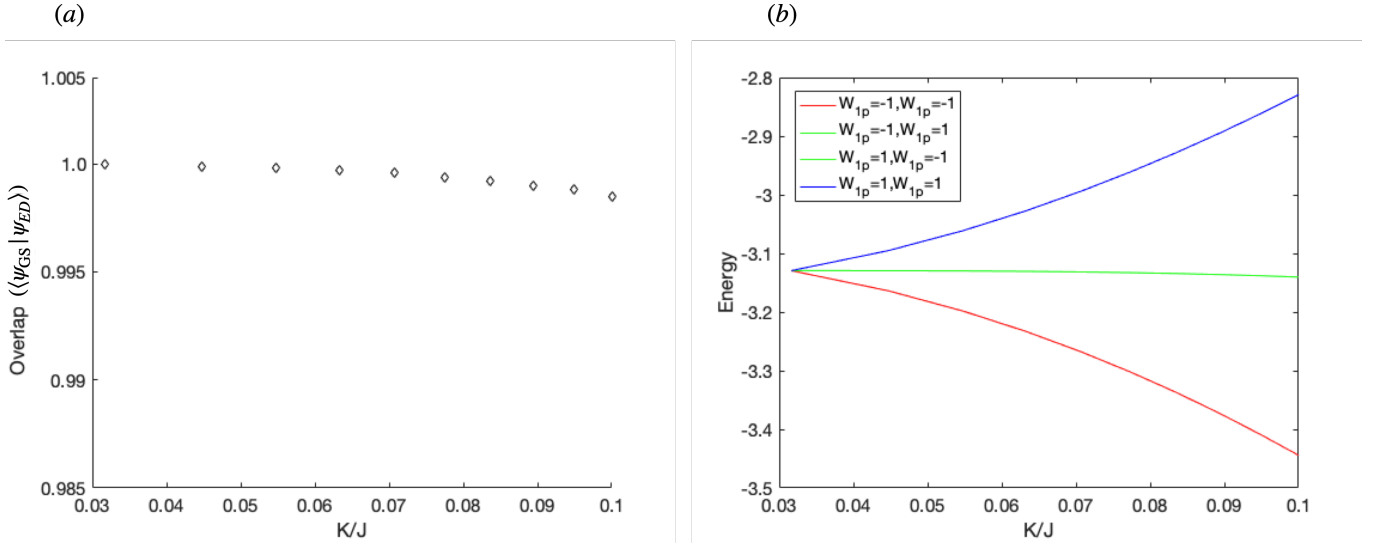
$$\langle \eta_p^\square \rangle = \langle 5(\boldsymbol{\eta}_i \cdot \boldsymbol{\eta}_j)(\boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_n) + 5(\boldsymbol{\eta}_i \cdot \boldsymbol{\eta}_n)(\boldsymbol{\eta}_j \cdot \boldsymbol{\eta}_k) - 5(\boldsymbol{\eta}_i \cdot \boldsymbol{\eta}_k)(\boldsymbol{\eta}_j \cdot \boldsymbol{\eta}_n) - (\boldsymbol{\eta}_i \cdot \boldsymbol{\eta}_j) - (\boldsymbol{\eta}_j \cdot \boldsymbol{\eta}_k) - (\boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_n) - (\boldsymbol{\eta}_i \cdot \boldsymbol{\eta}_n) - (\boldsymbol{\eta}_i \cdot \boldsymbol{\eta}_k) - (\boldsymbol{\eta}_j \cdot \boldsymbol{\eta}_n) + 1 \rangle_{\eta_{\text{Heisenberg}}}. \quad (74)$$

This expression is the expectation value of pseudo-spin ring- and pair-exchange terms on plaquette p , in the GS of the 2D Heisenberg model. A similar expression holds for the p' plaquettes. Depending on the sign of the expectation value, H_{g_4} lifts the extensive degeneracy at H_{g_2} level to leading order, and selects a GS of uniform flux. We confirmed the previous conclusion via exact diagonalization of $H_{\eta-\rho}$ in a unit cell. The Hilbert space includes all pseudo-spins for uniform $\rho_{ij}^\gamma = 1$ bonds and all states obtained by applying $W_{\nu p/p'}$ on the latter. The resulting Hamiltonian in the pseudo-spin basis was solved numerically. We find that the GS is unique and corresponds to a state of uniform π flux for both p/p' plaquettes (Supplementary figure 4(b)). We also find that the GS overlaps significantly with $|\eta; \{a_{\nu p}, a_{\nu p'}\}\rangle$ for $g_4 \ll g_2$ as shown in Supplementary figure 4(a).

VIII. ρ_{ij}^γ STRING DEFECTS

In this section, we consider

$$H_{g_2} = \sum_i (-1)^{i_x + i_y} (h_x \eta_i^x + h_z \eta_i^z) + g_2 \left[\sum_{\langle ij \rangle} \eta_i^z \eta_j^z + \sum_{\langle ij \rangle_\gamma} (\eta_i^x \eta_j^x + \eta_i^y \eta_j^y) \rho_{ij}^\gamma \right], \quad (75)$$



Supplementary figure 4. (a) The overlap between $|\psi_{GS}\rangle$ and the wavefunction obtained from ED as a function of K/J . The overlap decreases slowly as the perturbation gets stronger. (b) The lowest energy eigenstate is the state with both $W_{1p} = -1$ and $W_{1p'} = -1$ up to $K/J = 0.1$.

where $\gamma \in \{x, y, z, I\}$, on a torus. We distinguish two regimes for a) $h_x \gg g_2, h_z = 0$ and b) $h_z \gg g_2, h_x = 0$, respectively. H_{g_2} preserves ρ_{ij}^γ bonds and can be solved separately in any configuration of the latter, which is consistent with the local constraint $\prod_i q_i^\gamma = -1$. Here, we consider the uniform $\rho_{ij}^\gamma = 1$ configuration, together with related states which include non-contractible strings of n bonds along which $\rho_{ij}^\gamma = -1$. We present a semi-classical analysis and perturbation theory results, which show that the string has infinite energy cost for a), while for b) it vanishes, in the infinite-size system.

We first proceed by singling out n , ρ_{ij}^γ bonds along a non-contractible loop $C^{x/y}$ which winds once around the torus (Supplementary figure 5). We re-write the Hamiltonian as

$$H_{g_2} = H_{\text{String}} + H_{\text{Background}} + H_{\text{Coupling}}, \quad (76)$$

where H_{String} includes all on-site and NN exchange terms along the path. For the latter, we consider two configurations with ρ_{ij}^γ bonds either all negative or all positive. Similarly, $H_{\text{Background}}$ includes the remaining sites, with bonds trivially equal to 1. H_{Coupling} includes all of the exchange terms which connect string and background sites, with associated $\rho_{ij}^\gamma = 1$ bonds. We calculate corrections to the GS energy in perturbation theory for the g_2 terms.

The string of negative bonds around $C^{x/y}$ loops are equivalent to periodic boundary conditions (PBC) for the pseudo-spins along the direction of the loops. By contrast, the boundary conditions in the direction which crosses $C^{x/y}$ are

$$\eta_i^{x/y} = -\eta_{i+N_{2/1}}^{x/y} \quad (77)$$

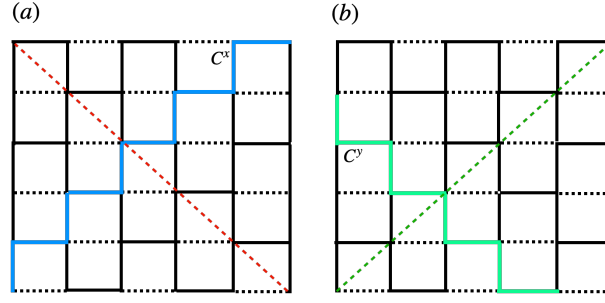
$$\eta_i^z = \eta_{i+N_{2/1}}^z. \quad (78)$$

for both sublattices. $N_{2/1}$ is the number of unit cells along the two directions implied by $C^{y/x}$. The $\eta^{x/y}$ components obey anti-periodic BC (ABC), while the z -components are subject to PBC.

A. $h_x > 0, h_z = 0$

It is convenient to apply a global rotation $\eta_i^x \rightarrow \tilde{\eta}_i^x, \eta_i^y \rightarrow \tilde{\eta}_i^y, \eta_i^z \rightarrow -\tilde{\eta}_i^z$ which maps H_{g_2} to

$$\tilde{H}_{g_2} = h_x \sum_i (-1)^{i_x + i_y} \tilde{\eta}_i^z + g_2 \sum_{\langle ij \rangle_\gamma} [\tilde{\eta}_i^z \tilde{\eta}_j^z \rho_{ij}^\gamma + \tilde{\eta}_i^x \tilde{\eta}_j^x + \tilde{\eta}_i^y \tilde{\eta}_j^y \rho_{ij}^\gamma], \quad (79)$$



Supplementary figure 5. Two non-contractible loops $C^{x,y}$ for the lattice on a torus. Solid black lines represent non-trivial $\rho_{ij}^\gamma = 1$ bonds while the dashed black lines are identity bonds. $C^{x/y}$ only include non-trivial bonds. Dashed red and green lines indicates paths used in constructing the Wilson loop operators in eq. 109

For a uniform $\rho_{ij}^\gamma = 1$ configuration, this reduces to a 2D AFM Heisenberg model which has finite staggered pseudo-spins in the GS for any $h_x \neq 0$. A non-contractible loop implies ABC for $\tilde{\eta}_i^z = -\tilde{\eta}_{i+N_{1/2}}^z$. At a semi-classical level, the presence of the staggered field h_x implies that such a twist costs an amount of energy which scales with $N_{2/1}$. Moreover, this energy cost diverges for an infinite-size system. This also holds in the subsequent $h_x \rightarrow 0$ limit.

The same conclusion can be reached using a perturbative analysis. It is convenient to re-write

$$\tilde{H}_{g_2} = H_0 + g_2 \sum_{\langle ij \rangle_\gamma} V_{z,ij}^\gamma + g_2 \sum_{\langle ij \rangle_\gamma} V_{\perp,ij}^\gamma, \quad (80)$$

where

$$H_0 = h_x \sum_i (-1)^{i_x + i_y} \tilde{\eta}_i^z \quad (81)$$

$$V_{z,ij}^\gamma = \tilde{\eta}_i^z \tilde{\eta}_j^z \rho_{ij}^\gamma \quad (82)$$

$$V_{\perp,ij}^\gamma = \tilde{\eta}_i^x \tilde{\eta}_j^x + \tilde{\eta}_i^y \tilde{\eta}_j^y \rho_{ij}^\gamma. \quad (83)$$

When considering matrix elements, we shall label the expectation value in the GS of \tilde{H}_0 as $\langle \rangle_{00}$. $\langle \rangle_{01}$ indicates matrix elements between GS and an excited state with a NN pair of flipped pseudo-spins, and similarly $\langle \rangle_{12}$ for the matrix elements between two excited states with 1 and 2 pairs of flipped pseudo-spins, respectively. The GS has staggered pseudo-spins to zeroth order in g_2 . To first order, the correction to the GS energy is

$$E_{n,\rho}^{(1)} = g_2 \left\langle \sum_{\langle ij \rangle_\gamma} V_{z,ij}^\gamma \right\rangle_{00} \quad (84)$$

$$= -g_2 \times n \times \rho + \text{Background}, \quad (85)$$

where we explicitly indicated the number of contributing configurations and their weights. These are due to the Ising exchange terms along the (rotated) z axis, and amount to taking into account all of the bonds in $\tilde{H}_{\text{String}}$ and $\tilde{H}_{\text{Coupling}}$. ρ stands for the two cases with either all bonds positive or negative on the string. The remaining contributions from $\tilde{H}_{\text{Background}}$ are independent of ρ . The corrections to second order are due exclusively to the (rotated) in-plane terms, which introduce pseudo-spin flips for NNs. Importantly, each flip carries a $(1 + \rho_{ij}^\gamma)$ factor as

$$E_{n,\rho}^{(2)} = g_2^2 \frac{\sum_{\langle ij \rangle_\gamma} \langle V_{\perp,ij}^\gamma \rangle_{01}^2}{\Delta E_{01}} \quad (86)$$

$$= -\frac{g_2^2}{32h_x} \times n \times (1 + \rho)^2 + \text{Background}, \quad (87)$$

where $\Delta E_{01} = -2h_x$ is the energy cost of an excited state with a single pair of flipped pseudo-spins. Also note that $\langle V_{\perp,ij}^\gamma \rangle = 2^{-2}(1 + \rho_{ij}^\gamma)$. To third order, we obtain

$$\begin{aligned}
E_{n,\rho}^{(3)} &= g_2^3 \frac{\sum_{\langle ij \rangle_\gamma} \langle V_{\perp,ij}^\gamma \rangle_{01} \left(\langle \sum_{\langle ab \rangle_\gamma} V_{z,ab}^\gamma \rangle_{11} - \langle \sum_{\langle ab \rangle_\gamma} V_{z,ab}^\gamma \rangle_{00} \right) \langle V_{\perp,ij}^\gamma \rangle_{10}}{(\Delta E_{01})^2} \\
&= \frac{g_2^3}{64h_x^2} [n \times (1 + \rho)^2(8 + 4\rho) + 2n \times 4(8 + 4\rho)] + \text{Background}. \tag{88}
\end{aligned}$$

These corrections are due to the energy cost introduced by V_z^γ terms in the presence of a pair of flipped pseudo-spins in the string and connecting bonds. To fourth order, the corrections read

$$\begin{aligned}
E_{n,\rho}^{(4)} &= g_2^4 \left(\frac{\sum_p \langle V_{\perp,ij}^\gamma \rangle_{01} \langle V_{\perp,jk}^\gamma \rangle_{11} \langle V_{\perp,kn}^\gamma \rangle_{11} \langle V_{\perp,ni}^\gamma \rangle_{10}}{(\Delta E_{01})^3} + \frac{\sum_p \langle V_{\perp,ij}^\gamma \rangle_{01} \langle V_{\perp,nk}^\gamma \rangle_{12} \langle V_{\perp,kj}^\gamma \rangle_{21} \langle V_{\perp,ni}^\gamma \rangle_{10}}{(\Delta E_{01})^2 \Delta E_{02}} + (\text{Permut.}/p \leftrightarrow p') \right. \\
&\quad + \frac{\sum_{\langle ij \rangle_\gamma} \sum_{\langle jk \rangle_\gamma, k \neq i} \langle V_{\perp,ij}^\gamma \rangle_{01} \langle V_{\perp,jk}^\gamma \rangle_{11} \left(\langle V_{\perp,jk}^\gamma \rangle_{11} \langle V_{\perp,ij}^\gamma \rangle_{10} + \langle V_{\perp,ij}^\gamma \rangle_{11} \langle V_{\perp,jk}^\gamma \rangle_{10} \right)}{(\Delta E_{01})^3} \\
&\quad + \frac{\sum_{\langle ij \rangle_\gamma} \sum_{\langle ab \rangle_\gamma, ab \neq ij} \langle V_{\perp,ij}^\gamma \rangle_{01} \langle V_{\perp,ab}^\gamma \rangle_{12} \left(\langle V_{\perp,ab}^\gamma \rangle_{21} \langle V_{\perp,ij}^\gamma \rangle_{10} + \langle V_{\perp,ij}^\gamma \rangle_{21} \langle V_{\perp,ab}^\gamma \rangle_{10} \right)}{(\Delta E_{01})^2 \Delta E_{02}} \\
&\quad - \frac{\sum_{\langle ij \rangle_\gamma} \sum_{\langle jk \rangle_\gamma, k \neq i} \langle V_{\perp,ij}^\gamma \rangle_{01}^2 \langle V_{\perp,jk}^\gamma \rangle_{01}^2}{(\Delta E_{01})^3} - \frac{\sum_{\langle ij \rangle_\gamma} \sum_{\langle ab \rangle_\gamma} \langle V_{\perp,ij}^\gamma \rangle_{01}^2 \langle V_{\perp,ab}^\gamma \rangle_{01}^2}{(\Delta E_{01})^3} \\
&\quad \left. + \frac{\sum_{\langle ij \rangle_\gamma} \langle V_{\perp,ij}^\gamma \rangle_{01} \left(\langle \sum_{\langle ab \rangle_\gamma} V_{z,ab}^\gamma \rangle_{11} - \langle \sum_{\langle ab \rangle_\gamma} V_{z,ab}^\gamma \rangle_{11} \right)^2 \langle V_{\perp,ij}^\gamma \rangle_{10}}{(\Delta E_{01})^3} \right) \tag{89}
\end{aligned}$$

The first line includes all pseudo-spin flips around p/p' plaquettes. The next three lines amount to flips on single bonds and on pairs of connected bonds, while the last line contains corrections similar to those at third order. Their effective contribution is

$$\begin{aligned}
E_{n,\rho}^{(4)} &= -\frac{g_2^4}{2048h_x^3} \left(20 \times n \times 4(1 + \rho)^2 + 2 \times n \times (1 + \rho)^4 + 2 \times 4n \times 4(1 + \rho)^2 + 2 \times 2n \times 16 \right. \\
&\quad \left. + n \times (1 + \rho)^2(8 + 4\rho)^2 + 2n \times 4(8 + 4\rho)^2 - n \times (1 + \rho)^4 - 2n \times 16 \right) + \text{Background}. \tag{90}
\end{aligned}$$

The energy cost of the string to fourth order is

$$E_{n,\rho=-1} - E_{n,\rho=1} = n \left(2g_2 + \frac{g_2^2}{8h_x} - \frac{7g_2^3}{4h_x^2} + \frac{129g_2^4}{128h_x^3} \right) + O(g_2^5). \tag{91}$$

It is apparent that this diverges as the system size, and thus n , grows to infinity. Note that this also implies that the cost of an open string of infinite extent likewise diverges. An open string of length $(n - 2)$ is obtained from the non-contractible loop by changing the sign of two neighboring bonds. This is a local perturbation, which costs a finite amount of energy. The energy of the string therefore diverges as that of the parent state. As discussed in the main text, this implies that the visons are confined.

B. $h_z > 0, h_x = 0$

In contrast to the case with $h_x > 0, h_z = 0$, here the BCs associated with the loop do not imply a finite energy cost in a semi-classical approximation. This is due to the fact that the pseudo-spins are staggered along z , and η^z does not couple to the bonds. We further illustrate by considering the Holstein-Primakoff representation for the pseudo-spins [?].

$$\eta_i^+ = (\sqrt{2\eta - n_{b,i}}) b_i \tag{92}$$

$$\eta_i^- = b_i^\dagger (\sqrt{2\eta - n_{b,i}}) \tag{93}$$

$$\eta_i^z = \eta - n_{b,i}. \tag{94}$$

Here, b_i, b_i^\dagger are bosons, $n_{b,i} = b_i^\dagger b_i$, and $\eta \gg n_{b,i}$ is the size of the pseudo-spin. The BCs for the pseudo-spins can be implemented by imposing ABCs for the bosons as $b_i = -b_{i+N_2}$. The remaining analysis follows the usual semi-classical approximation. In particular, as $h_z \rightarrow 0$, the zero-point energy of the GS involves a summation over the reciprocal unit cell as [?].

$$E' = 2 \sum_{\mathbf{k}} g_2 \eta \left(\sqrt{1 - \alpha_{\mathbf{k}}^2} - 1 \right), \quad (95)$$

where $\alpha_{\mathbf{k}}$ is the usual form factor for a square lattice. ABC amount to a trivial translation of the entire reciprocal unit cell in the extended BZ by $\mathbf{G}_{1/2}/2N_{1/2}$, where \mathbf{G}_2 is a reciprocal lattice vector. The correction to the classical GS energy is invariant. Consequently, at this level of approximation, the string is degenerate with the uniform $\rho_{ij}^\gamma = 1$ state.

We reach the same conclusion using a perturbative analysis. In contrast to the $h_x > 0$ case, the first-order corrections are independent of ρ_{ij}^γ :

$$\begin{aligned} E_{n,\rho}^{(1)} &= -g_2 \sum_{\langle ij \rangle} \\ &= -ng_2 + \text{Background}, \end{aligned} \quad (96)$$

since the Ising interaction along z does not couple to the bonds. The second-order contribution, due to pairs of NN pseudo-spin flips, reads

$$\begin{aligned} E_{n,\rho}^{(2)} &= -g_2^2 \frac{\sum_{\langle ij \rangle} (\rho_{ij}^\gamma)^2}{8h_z} \\ &= -ng_2^2 \frac{1}{8h_z} + \text{Background}. \end{aligned} \quad (97)$$

Note the important difference w.r.t. the $h_x > 0$ case, where each flip of NN pseudo-spins was multiplied by a $(1 + \rho_{ij}^\gamma)$ factor. Here, the corrections are independent of bond configuration. A similar conclusion holds at third order. At fourth order, we obtain the plaquette terms

$$E_{n,\rho}^{(4)} = -\frac{g_2^4}{32h_x^3} \left[\sum_p \left(\prod_{\langle ij \rangle \in p/p'} \rho_{ij}^\gamma \right) + p \leftrightarrow p' \right] + \dots, \quad (98)$$

where we omitted contributions similar to those at third order, which likewise are independent of the bonds. The terms shown above involve products of ρ 's around p/p' plaquettes. Since each plaquette contains at least two bonds on the string, these are also trivial. However, similar corrections involving pseudo-spin flips around closed loops, which intersect the string an odd number of times, are non-trivial. It is straightforward to check that only non-contractible loops, which wrap around the torus at least once, can contribute. These occur at leading n th order and scale with a factor of g_2^n/h_z^{n-1} . Consequently, each is exponentially suppressed as the system size, and therefore n , goes to infinity. The string becomes degenerate with the uniform $\rho_{ij}^\gamma = 1$ state in this limit. This also implies an upper bound on the energy cost of any open string of length $n - 2$. The latter can be obtained by changing the sign of any two NN bonds, which is a local perturbation, implying a finite energy cost. This also holds as the system size goes to infinity. As discussed in the main text, it follows that visons are deconfined.

IX. CONNECTION BETWEEN HUBBARD AND $\eta - \rho$ MODELS

In this section, we show that the GSs of the Hubbard model and $H_{\eta-\rho}$ models are equivalent.

We start with the Hubbard model (eq. 3 in the main text). In general, the GS for large J is characterized by a finite expectation value of the gauge-invariant correlators

$$C_{ij}^x = \langle N_i^x \prod_{i'j' \in (i,j)} u_{1i'j'} u_{2i'j'} N_j^x \rangle \quad (99)$$

where

$$N_{i/j}^x = (-1)^{i_x + i_y} \langle f_{2i}^\dagger f_{1i} + f_{1i}^\dagger f_{2i} \rangle \quad (100)$$

are staggered, pseudo-spins along x , which are connected via a gauge string, and similarly for y components. We write the correlator in terms of Majorana fermions as

$$\mathcal{C}_{ij}^x = \prod_{i'j' \in (i,j)} (-1)^{i_x + i_y + i'_x + i'_y} \langle (c_{1i}^x c_{2i}^x + c_{1i}^y c_{2i}^y) u_{1i'j'} u_{2i'j'} (c_{1j}^x c_{2j}^x + c_{1j}^y c_{2j}^y) \rangle \quad (101)$$

We next consider a similar correlator in the π -flux GS of $H_{\eta-\rho}$, (eq. 19 in the main text)

$$\mathcal{D}_{ij}^x = (-1)^{i_x + i_y + i_{x'} + i_{y'}} \left\langle \eta_i^x \left(\prod_{i'j' \in C_{ij}} \rho_{ij}^\gamma \right) \eta_j^x \right\rangle. \quad (102)$$

The correlator commutes with $W_{\nu p/p'}$. It follows from the form of the GS of $H_{\eta-\rho}$ that \mathcal{D}_{ij}^x is finite for $h_x > 0$ and any i, j . We next perform a rotation on all of the spin (σ) DOF of the second layer as

$$\begin{aligned} \eta_i^x &\rightarrow \tilde{\eta}_i^x = \frac{1}{4}(\sigma_{1i}^x \sigma_{2i}^x + \sigma_{1i}^y \sigma_{2i}^y) \\ \eta_i^y &\rightarrow \tilde{\eta}_i^y = \frac{1}{4}(\sigma_{1i}^y \sigma_{2i}^x - \sigma_{1i}^x \sigma_{2i}^y) \\ \eta_i^z &\rightarrow \tilde{\eta}_i^z = \frac{1}{4}(\sigma_{1i}^z - \sigma_{2i}^z). \end{aligned} \quad (103)$$

We introduce a Majorana representation for the rotated pseudo-spins as

$$\tilde{\eta}_i^x = (-1)^{i_x + i_y} b_{1i}^4 b_{2i}^4 (c_{1i}^x c_{2i}^x + c_{1i}^y c_{2i}^y), \quad (104)$$

as well as for the string

$$\prod_{i'j' \in C_{ij}} \rho_{ij}^\gamma = b_{1i}^4 b_{2i}^4 c_{1i}^x c_{2i}^y c_{1i}^y c_{2i}^x \prod_{i'j' \in (i,j)} u_{1i'j'} u_{2i'j'} b_{1j}^4 b_{2j}^4 c_{1j}^x c_{2j}^y c_{1j}^y c_{2j}^x \quad (105)$$

where we used $\rho_{ij}^\gamma = \prod_\nu \tau_{\nu i}^\gamma \tau_{\nu j}^\gamma$ and $\tau_{\nu i}^\gamma = b_{\nu i}^4 b_{\nu i}^\gamma c_{\nu i}^x c_{\nu i}^y$. Substituting these in eq. 102, we obtain

$$\mathcal{D}_{ij}^x = \prod_{i'j' \in (i,j)} (-1)^{i_x + i_y + i'_x + i'_y} \langle (c_{1i}^x c_{2i}^x + c_{1i}^y c_{2i}^y) u_{1i'j'} u_{2i'j'} (c_{1j}^x c_{2j}^x + c_{1j}^y c_{2j}^y) \rangle_{\Psi_{\text{GS}}}. \quad (106)$$

This expression is formally identical to \mathcal{C}_{ij}^x .

X. TOPOLOGICAL DEGENERACY OF THE GROUND STATE OF $H_{\eta-\rho}$

In this section we show that the GS of $H_{\eta-\rho}$ exhibits \mathbb{Z}_2 topological order for $h_x > 0, h_z = 0$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order for $h_z > 0, h_x = 0$.

A. \mathbb{Z}_2 topological order

For $h_x > 0, h_z = 0$, the GS of $H_{\eta-\rho_{ij}}$ has the form

$$|\Psi_{\text{GS}}\rangle = \prod_{\nu,p} \frac{1}{4} (1 - W_{\nu p}) (1 - W_{\nu p'}) |\tilde{\eta}_{xx}; \phi_0\rangle + O\left(\frac{h_x}{g_4}\right). \quad (107)$$

where $|\tilde{\eta}_{xx}; \phi_0\rangle$ is the GS of the 2D Heisenberg model for finite h_x , while ϕ_0 is an orbital configuration with all $q_i^\gamma = -1$, which implies that it has uniform $\rho_{ij}^\gamma = 1$ bonds. Small corrections to this state occur because h_x does not commute with the flux operators $W_{\nu p/p'}$.

The projection of $|\tilde{\eta}_{xx}; \phi_0\rangle$ onto the π -flux state in eq. 107 contains factors of the form $W_{\nu p} W_{\nu p'}$, where p/p' are in the same unit cell. When operating on $|\tilde{\eta}_{xx}; \phi_0\rangle$ these preserve all pseudo-spin and ρ_{ij}^γ bonds. However, they change the sign of all pairs of NN pairs of q^γ 's around the unit cell. To illustrate, the NN pair $q_1^z = -1, q_j^z = -1$ is mapped

onto $q_1^z = 1, q_j^z = 1$, with $\rho_{ij}^\gamma = 1$ in both cases, and similarly for the remaining NN pairs. However, since these states are obtained from the ϕ_0 configuration via local changes in the set of q 's, they are not topologically distinct. The $W_{\nu p} W_{\nu p'}$ operators are in this sense the analogs of plaquette operators in the toric code, while the signs of NN q^γ 's can be used to define effective bonds.

We construct topologically distinct configurations by operating with

$$F^{x/y} = \prod_{i \in C^{x/y}} \tau_{1i}^{x/y}. \quad (108)$$

Here, the orbital operators $\tau_{1i}^{x/y}$ are applied on every site along non-contractible loops $C^{x/y}$, as illustrated in Supplementary figure 5. $F^{x/y}$ commutes with both $H_{\eta-\rho}$ and $W_{\nu p/p'}$, and has the effect of changing the signs of all NN q 's along the loop. Note that $\tau_{2i}^{x/y}$ would have the same effect. Any state thus obtained is degenerate with $|\Psi_{\text{GS}}\rangle$.

In order to capture the topology of the GS manifold, we introduce the Wilson loop operators

$$I^{x/y} = \prod'_{i,j \in C^{y/x}} q_i^{y/x} q_j^z, \quad (109)$$

where \prod' denotes every other site on $C^{y/x}$ (Supplementary figure 5). Since $[I^{x/y}, H_{\eta-\rho}] = 0$ and $[I^{x/y}, W_{\nu p/p'}] = 0$, $|\Psi_{\text{GS}}\rangle$ can be labeled by $\lambda_x = \lambda_y = 1$. More generally, the states

$$|\Psi_{\text{GS}}; \lambda_x, \lambda_y\rangle = (F^x)^{\frac{1+\lambda_x}{2}} (F^y)^{\frac{1+\lambda_y}{2}} |\Psi_{\text{GS}}\rangle. \quad (110)$$

belong to four distinct topological sectors. The four-fold degeneracy is characteristic of \mathbb{Z}_2 topological order. This procedure exhausts the degeneracy of the GS manifold, since any non-local changes in the pseudo-spin configurations cost finite energy for $h_x > 0$. Similarly, changes in the ρ^γ bonds are equivalent to the strings discussed in Supplementary section VIII and are likewise gapped for $h_x > 0$.

B. $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order

The previous arguments for $h_x > 0, h_z = 0$ also hold for $h_z > 0, h_x = 0$. The \mathbb{Z}_2 degeneracy due to configurations of the local orbital states q^γ consistent with a given set of ρ^γ bonds still holds. However, as shown in Supplementary section VIII, non-contractible loops of flipped ($\rho_{ij}^\gamma \rightarrow -\rho_{ij}^\gamma$) bonds becomes degenerate with $|\Psi_{\text{GS}}\rangle$ in the infinite-size limit. In order to account for this additional degeneracy, we can define the Wilson loop operators

$$I_A^{x/y} = \prod_{i,j \in C^{y/x}} \rho_{ij}^\gamma, \quad (111)$$

with eigenvalues $\lambda_{A,x/y} = \pm 1$. These commute with $H_{\eta-\rho}$ and $W_{\nu p/p'}$. For any eigenvalues of $I_{x/y}$ (eq. 109), we can therefore distinguish four additional topologically distinct states

$$|\Psi_{\text{GS}}; \lambda_x, \lambda_y, \lambda_{A,x}, \lambda_{A,y}\rangle = (L^x)^{\frac{1+\lambda_{A,x}}{2}} (L^y)^{\frac{1+\lambda_{A,y}}{2}} |\Psi_{\text{GS}}; \lambda_x, \lambda_y\rangle. \quad (112)$$

The operators

$$L^{x/y} = \prod_{i \in C^{x/y}} \tau_i^{x/y} \quad (113)$$

act along every other site on path $C^{x/y}$ to create non-contractible loops of flipped ρ^γ bonds. The GS thus exhibits a sixteen-fold degeneracy on the torus, as consistent with a $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological order.