ARTICLE   OPEN

Kitaev spin-orbital bilayers and their moiré superlattices

Emilian Marius Nica [1,2 ✉], Muhammad Akram[1,3], Aayush Vijayvargia[1], Roderich Moessner[4] and Onur Erten[1]

We determine the phase diagram of a bilayer, Yao-Lee spin-orbital model with inter-layer interactions (J), for several stackings and moiré superlattices. For AA stacking, a gapped $\mathbb{Z}_2$ quantum spin liquid phase emerges at a finite J_c. We show that this phase survives in the well-controlled large-J limit, where an isotropic honeycomb toric code emerges. For moiré superlattices, a finite-$\mathbf{q}$ inter-layer hybridization is stabilized. This connects inequivalent Dirac points, effectively 'untwisting' the system. Our study thus provides insight into the spin-liquid phases of bilayer spin-orbital Kitaev materials.

INTRODUCTION

Quantum spin liquids (QSLs) are disordered phases of magnetic systems with emergent exotic properties arising from their underlying topological character[1–5]. The Kitaev model on the honeycomb lattice[6,7] is of particular significance as the first member of a family of exactly-solvable models. Recent years witnessed experimental progress in identifying candidate materials which include a number of iridates[8] and a-RuCl$_3$[9]. Kitaev interactions can also be strong in other van der Waals (vdW) materials such as CrI$_3$[10,11]. Moreover, vdW materials can be arranged in stacking patterns and twisted to form moiré superlattices, potentially leading to new phases. Indeed, recent theoretical studies[12–17] predict several magnetic phases in twisted vdW magnets, partially realized experimentally[18,19].

We study the zero-temperature phase diagram of bilayer versions of Kitaev spin-orbital models, initially proposed by Yao and Lee[20], with additional inter-layer Heisenberg interactions. Spin-orbital models are generalizations of the original Kitaev model with extra local orbital degrees of freedom (DOF) and Kugel-Khomskii interactions for spin and orbital sectors[20–30]. Much like Kitaev's original proposal, spin and orbital DOF can each be represented in terms of three-flavored sets of Majorana fermions. The Yao-Lee model exhibits an emergent $\mathbb{Z}_2$ gauge symmetry with gapped flux excitations (visons) defined exclusively in terms of the orbital DOF[20]. The inter-layer spin-exchange interactions commute with the intra-layer flux operators, in contrast to the Kitaev model and subsequent bilayer realizations[31–34]. We take advantage of this unique feature by considering only the lowest-energy, zero-flux sector. Furthermore, we treat the spin-exchange interactions in the Hartree approximation. This introduces an effective inter-layer hybridization for the itinerant Majorana fermions associated with the spin DOF. A non-zero expectation value indicates the formation of inter-layer spin-singlets, as shown in Supplementary Note 1. The conservation of the fluxes in the Yao-Lee bilayer, which are defined exclusively in terms the orbital DOF, stands in clear contrast to bilayers based on Kitaev's original model. As shown below, this leads to distinct phase diagrams and to an enhanced stability of topological QSL phases in Yao-Lee bilayers.

We focus on AA stacking and moiré superlattices, which exhibit fully-gapped spectra, but also briefly cover the gapless, AB stacking case. For AA stacking, the effective hybridization becomes non-zero at a finite value of the inter-layer exchange coupling, and opens a

gap in the itinerant Majorana fermion spectrum. This signals a topological phase transition to a gapped $\mathbb{Z}_2$ QSL. We support our Hartree approximation with two additional considerations. First, we show that the bilayer model is equivalent to an attractive Hubbard model with three flavors of complex fermions, for our choice of gauge. Previous quantum Monte Carlo (QMC) studies have shown that the Hubbard model exhibits a single transition to a charge density wave (CDW) phase[35], which is equivalent to the bilayer with a non-zero inter-layer hybridization. Secondly, we show that in the limit of large inter-layer interactions, the bilayer model maps onto Kitaev's toric code[36], which is gapped and exhibits topological order. This naturally suggests that the gapped phase predicted by the Hartree approximation is adiabatically connected to the toric code. However, first-order transitions, possibly involving changes in the flux configurations, cannot be completely excluded. For AB stacking, the formation of inter-layer spin singlets leaves the itinerant Majorana fermions gapless with quadratic band touching, in analogy with bilayer graphene[37]. For moiré superlattices, we consider both uniform ($\mathbf{q} = 0$) and modulated inter-layer effective hybridizations ($\mathbf{q} \neq 0$). In contrast to the $\mathbf{q} = 0$ case, the finite-$\mathbf{q}$ hybridization connects inequivalent Dirac points, effectively 'untwisting' the system, and opening a gap. This leads to the emergence of a gapped $\mathbb{Z}_2$ QSL, as for AA stacking.

Kitaev spin-orbital models can be realized in spin-orbit coupled $4d$ and $5d$ Mott insulators, as predicted by several recent studies[26,38–40]. For instance, an enhanced SU(4) symmetry[41] has been advanced for a-ZrCl$_3$.

RESULTS
Model
Our models include intra-layer Yao-Lee[20] interactions on a honeycomb lattice (H_v), and inter-layer, antiferromagnetic Heisenberg interactions (H_l):

$$H = H_v + H_l \tag{1}$$

$$H_v = \sum_{a-\text{links},\langle ij\rangle} K^{(a)}\left(\tau_{v,i}^{(a)}\tau_{v,j}^{(a)}\right)\left(\boldsymbol{\sigma}_{v,i}\cdot\boldsymbol{\sigma}_{v,j}\right) \tag{2}$$

$$H_l = \sum_{ij} J_{ij}\boldsymbol{\sigma}_{1i}\cdot\boldsymbol{\sigma}_{2j}. \tag{3}$$

[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]Department of Physics, Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta 87300, Pakistan. [4]Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Strasse 38, 01187 Dresden, Germany. ✉email: en5@rice.edu
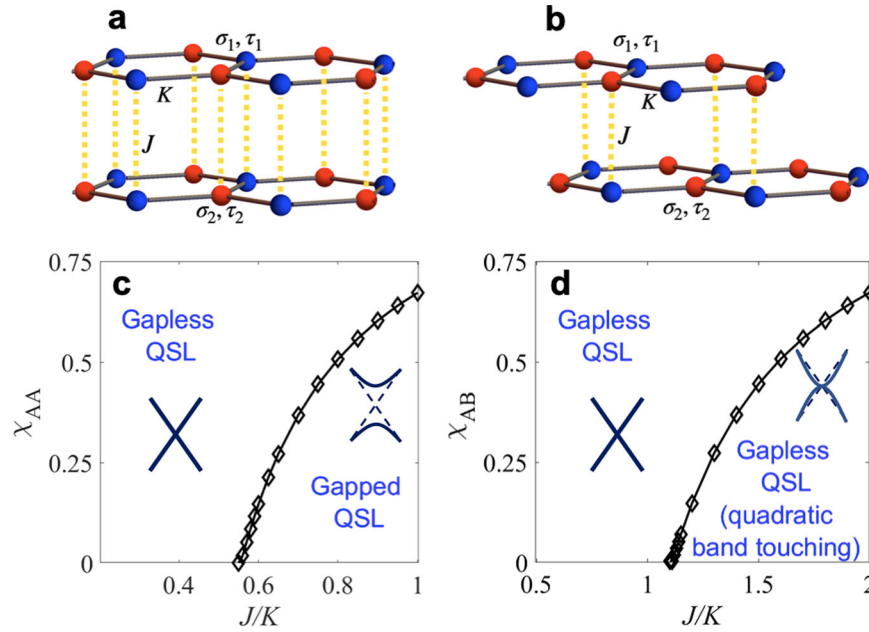
Fig. 1 Yao-Lee bilayer without twisting. Illustration of the model for (**a**) AA and (**b**) AB stacking patterns. K and J are the intra-layer Kitaev and inter-layer Heisenberg exchange terms, respectively. (**c**) Effective inter-layer hybridization for AA stacking. Finite $\langle \chi_{AA} \rangle$ indicates the formation of inter-layer singlets and leads to gapped itinerant Majorana fermions. (**d**) Same for AB stacking, leading to quadratic band touching.

We first focus on H_ν, where $K^{(a)}$ is the nearest neighbor coupling (NN) constant for type-a links ($a \in \{x, y, z\}$) (Fig. 1a, b). The lattice sites are labeled by i and j, while $\nu \in \{1, 2\}$ denotes the two layers. An exact solution is obtained by introducing Majorana fermion representations for the spin and orbital DOF in each layer: $\sigma_{\nu,j}^{(a)} = -i\epsilon^{\alpha\beta\gamma} c_{\nu,j}^{(\beta)} c_{\nu,j}^{(\gamma)}/2$ and $\tau_{\nu,j}^{(a)} = -i\epsilon^{\alpha\beta\gamma} d_{\nu,j}^{(\beta)} d_{\nu,j}^{(\gamma)}/2$ [20]. Note that we use a normalization convention for the Majorana fermions where $\{c_{\mu,i}^{(a)}, c_{\nu,j}^{(\beta)}\} = 2\delta_{\alpha\beta}\delta_{\mu\nu}\delta_{ij}$, and similarly for the b's. These representations are redundant and the physical states in each layer must be restricted to the eigenstates of $D_{\nu,i} = -i c_{\nu,i}^{(x)} c_{\nu,i}^{(y)} c_{\nu,i}^{(z)} d_{\nu,i}^{(x)} d_{\nu,i}^{(y)} d_{\nu,i}^{(z)}$ operators with eigenvalues 1. As in Kitaev's model, these constraints can be imposed via projection operators $P_\nu = \prod_i (1 + D_{\nu,i})/2$. The intra-layer Hamiltonians in the Majorana representation can be expressed as $H_\nu = P_\nu \mathcal{H}_\nu P_\nu$, where

$$\mathcal{H}_\nu = \sum_{\langle ij \rangle} K^{(a)} u_{\nu,ij}^a \left[i c_{\nu,i}^{(x)} c_{\nu,j}^{(x)} + i c_{\nu,i}^{(y)} c_{\nu,j}^{(y)} + i c_{\nu,i}^{(z)} c_{\nu,j}^{(z)} \right]. \quad (4)$$

The bond operators $u_{\nu,ij}^{(a)} = -i d_{\nu,i}^{(a)} d_{\nu,j}^{(a)}$, where i, j are on the A and B sublattices, respectively, commute with $\mathcal{H}_\nu$, and are therefore conserved with eigenvalues ± 1. Both $\mathcal{H}_\nu$ are invariant under separate $\mathbb{Z}_2$ gauge transformations $c_{\nu,i}^{(a)} \rightarrow -c_{\nu,i}^{(a)}$; $u_{ij}^{(a)} \rightarrow -u_{ij}^{(a)}$ with flux operators which are defined by the product of the $u_{ij}^{(a)}$ around hexagonal plaquettes.

Lieb's theorem[42] predicts that the ground state lies in the zero-flux sector, with a finite vison gap. We can obtain the itinerant Majorana spectrum by choosing a gauge where $u_{ij} = 1 \, \forall \langle ij \rangle$ in both layers. Unless otherwise stated, we use this choice throughout. The three flavors of itinerant Majorana fermions have identical spectra, which are gapless for $K_x + K_y > K_z$. We consider the symmetric, gapless case with $K_x = K_y = K_z = K$.

We now consider the inter-layer interactions in H_I. Unlike in the Kitaev model, the visons in the Yao-Lee model are defined exclusively in terms of the orbital DOF, while the itinerant Majorana excitations stem from the spin DOF alone. Consequently, additional terms involving the spin DOF only, including a bilayer coupling, commute with the flux operators. The resulting spectrum can a priori preserve the gapped flux excitations, in

contrast to the original Kitaev model[31,33,34]. Consequently, we consider Yao-Lee bilayers coupled via inter-layer antiferromagnetic Heisenberg interactions in H_I. Note that we allow for general inter-layer J_{ij} coupling beyond NN. The bilayer Hamiltonian in the Majorana representation is $\mathcal{H} = \sum_\nu \mathcal{H}_\nu + \mathcal{H}_I$ where

$$\mathcal{H}_I = \sum_{i,j;\alpha \neq \beta} \frac{J_{ij}}{2} \left(c_{1i}^{(\alpha)} c_{2j}^{(\alpha)} c_{1i}^{(\beta)} c_{2j}^{(\beta)} \right). \quad (5)$$

Self-consistent solutions for AA and AB stacking patterns

The inter-layer interactions are bi-quadratic in the itinerant Majorana operators, thus precluding a closed-form solution. Instead, we treat the inter-layer interactions within a Hartree approximation. This approach is supported by additional considerations, as discussed below.

For the purpose of illustration, we restrict the inter-layer coupling to NN pairs. We do not expect that weaker couplings beyond NNs will alter our conclusions. The on-site mean-field (MF) parameters $\langle \chi_i^{(a)} \rangle = \langle i c_{1i}^{(a)} c_{2i}^{(a)} \rangle$ preserve the SO(3) spin symmetry, and, we drop the corresponding flavor indices for most of the following discussion.

Before proceeding with a detailed presentation of the results, we first clarify the nature of the MF parameters. In the absence of intra-layer Yao-Lee interactions ($K^{(a)} = 0$), the decoupled, inter-layer, spin-singlet states for overlapping sites can equally be described by two eigenstates of $\chi_i^{(a)}$, with eigenvalues ± 1 for each a, as shown in Supplementary Note 1. The Ising-like nature of these states stems from a redundancy in the representation of the decoupled singlets in terms of the c Majorana fermions. Once the intra-layer interactions are turned on, and a set of bond variables ($u_{ij}^{(a)}$) is chosen, we obtain a unique MF solution with $\langle \chi_i^{(a)} \rangle \neq 0$, which is identical for the three flavors. These finite MF parameters likewise indicate the formation of inter-layer spin-singlets in the physical ground-state (GS). However, the Ising-like nature of these parameters is not immediately physical, since the non-trivial phases that we find are not described in terms of a local order parameter. We further elucidate these aspects in the following.

As previously mentioned, we carry out the Hartree approximation in a gauge where all $u_{ij} = 1$ in both layers, and obtain the GS

$$|\Psi\rangle = \left|\forall\, u_{v,ij}^{(a)} = 1\right\rangle \otimes \left|\left\langle \chi_i^{(a)} \neq 0\right\rangle\right\rangle. \tag{6}$$

Importantly, $\langle \chi_i^{(a)}\rangle_\Psi$ is not a well-defined, Landau-Ginzburg order parameter for the bilayer. Indeed, any gauge transformation, implemented for instance by $D_{v,i}|\Psi\rangle$, changes the sign of the associated $\langle \chi_i^{(a)}\rangle$ together with those of the three bonds extending from i in layer v. Furthermore, the physical GS is obtained by applying the projector P to $|\Psi\rangle$ as

$$|\Psi\rangle_{\text{Phys}} = P|\Psi\rangle. \tag{7}$$

$|\Psi\rangle_{\text{Phys}}$ amounts to a linear superposition of all gauge-symmetrized states which preserve a net zero flux, as shown in Supplementary Note 2 A. States with finite $\pm|\langle \chi_i^{(a)}\rangle|$ occur with equal weight, implying that $\langle \chi_i^{(a)}\rangle_{\text{Phys}} = 0$.

In order to characterize transitions in the physical GS, we instead consider a gauge-invariant correlator

$$\left\langle C_{ij}^{(a)}\right\rangle_{\text{Phys}} = \left\langle\left(\prod_{a-\text{links},\langle i'j'\rangle}' u_{1,i'j'}^{(a)} u_{2,i'j'}^{(a)}\right)\chi_i^{(a)}\chi_j^{(a)}\right\rangle_{\text{Phys}}, \tag{8}$$

where the strings of bonds connect operators at the end sites i, j. In Supplementary Note 2 B, we show that the expectation value of $C_{ij}^{(a)}$ in the physical GS matches that of a two-point correlator for $\chi^{(a)}$ in $|\Psi\rangle$.

$$\left\langle C_{ij}^{(a)}\right\rangle_{\text{Phys}} = \left\langle \chi_i^{(a)}\chi_j^{(a)}\right\rangle_\Psi. \tag{9}$$

From this expression, long-range order in $|\Psi\rangle$ is equivalent to $\langle \chi_i^{(a)}\rangle_\Psi \neq 0$. It follows that non-vanishing MF parameters imply a finite $\langle C_{ij}^{(a)}\rangle_{\text{Phys}}$, in the limit of infinite separation. In Supplementary Note 2 C, we express $\langle C_{ij}^{(a)}\rangle_{\text{Phys}}$ in terms of the spin and orbital operators of the bilayer, and show that it signals a topological phase transition to a gapped, $\mathbb{Z}_2$ QSL for the AA-stacked case, which involves the formation of inter-layer spin-singlets. We note that all subsequent conclusions regarding the MF parameters, obtained in the Hartree approximation and in a fixed gauge, are to be understood in the present context.

We now discuss our results in the Hartree approximation. For the AA stacking pattern, the A and B sublattice sites overlap (Fig. 1a). The inter-layer interactions involve two pairs of sites per unit cell: $\mathcal{H}_I = -2J[\sum_{i\in A,a}\langle \chi_{\text{AA}}\rangle(ic_{1i}^a c_{2i}^a) + \sum_{j\in B,a}\langle \chi_{\text{BB}}\rangle(ic_{1j}^a c_{2j}^a)]$. Solutions which are both uniform and symmetric in the sublattice index ($\langle \chi_{\text{AA}}\rangle = \langle \chi_{\text{BB}}\rangle$) amount to gapless itinerant Majorana fermions, with shifted Dirac cones. In contrast, when the hybridization has an alternating sign on the two sublattices ($\langle \chi_{\text{AA}}\rangle = -\langle \chi_{\text{BB}}\rangle$), the spectrum is gapped, leading to a lower ground-state energy. Our self-consistent solutions are shown in Fig. 1c as functions of J/K. We find that the critical value for this transition is $J_c/K = 0.55$.

To establish the stability of our solutions beyond the Hartree approximation, we map $\mathcal{H}$ to an equivalent form by using complex fermions $f_i^a = (c_{1i}^a + ic_{2i}^a)/2$:

$$\mathcal{H} = 2K\sum_{\langle ij\rangle,a}(if_{A,i}^{a\dagger}f_{B,j}^a + \text{H.c.}) - 2J\sum_i\left(n_i - \frac{3}{2}\right)^2, \tag{10}$$

where $n_i = \sum_a f_i^{a\dagger}f_i^a$. For $J > 0$, Eq. (10) describes an attractive Hubbard model with three flavors of complex fermions. This model exhibits a single, broken-symmetry CDW phase with finite $\langle n_A\rangle = -\langle n_B\rangle$, as determined by QMC[35]. This Ising order parameter acts as a mass term for the complex fermions, and gaps their spectrum. It is equivalent to a solution in which $\langle \chi\rangle$ alternates between sublattices in the Yao-Lee bilayer. Importantly,

the Hubbard model and CDW order parameter were obtained by fixing the gauge. While the CDW breaks inversion symmetry in the Hubbard model, the same cannot be said of the physical GS of the bilayer model. As previously mentioned, the order parameters obtained in a fixed gauge are physically meaningful only in relation to the gauge-invariant correlator $\langle C_{ij}^{(a)}\rangle_{\text{Phys}}$ defined in Eq. (9).

The GS obtained in the Hartree approximation for AA stacking has a fourfold topological degeneracy, as shown in Supplementary Note 9. This result is corroborated by the perturbative analysis in the large-J limit discussed in the following.

For AB stacking, the A sublattice sites of layer 1 lie directly on top of the B sublattice sites of layer 2, with a single bond per unit cell, (Fig. 1b). Therefore, for finite $\langle \chi\rangle$ beyond $J_c/K \simeq 1.1$, the itinerant Majorana spectrum is similar to that of AB-stacked bilayer graphene with quadratic band touching[37]. The self-consistent solutions for $\langle \chi_{\text{AB}}\rangle$ are shown in Fig. 1d. A mapping to an equivalent model as in the AA case is not apparent here.

The choice of uniform $u_{ij} = 1$ for both layers implies that the system persists in a zero-flux sector. This is supported by additional MF calculations with several distinct non-zero flux patterns (Supplementary Note 3), which indicate that the zero-flux states are lower in energy. Furthermore, the effective Hamiltonian in the large-J limit (see below) similarly prefers this configuration.

We comment on the stability of the phases obtained in the Hartree approximation in the presence of additional inter-layer, NN, spin-exchange interactions, which we realistically expect to be subleading. For the AA-stacked bilayer, the gapped phase obtained for $J > J_c$ is stable with respect to additional, infinitesimal, NN interactions. For the AB-stacked bilayer with $J > J_c$, our Hartree approximation predicts quadratic band touching, which implies a finite density of states for the itinerant Majorana fermions at zero energy. Additional inter-layer, NN interactions are therefore likely relevant in a renormalization-group sense. Establishing the nature of the low-energy phases in these cases requires further analysis, at Hartree level and beyond, and we reserve such questions for future study.

The GSs obtained in the fixed gauge survive projection onto the physical sector, as shown in Supplementary Note 2 A.

## Limit of large inter-layer interactions with AA stacking pattern

We consider the AA stacking pattern in the limit of large J/K. To zeroth order in the intra-layer (K) terms, the GS manifold consists of a collection of independent inter-layer spin singlets with degenerate orbital states. We derive an effective Hamiltonian on the GS manifold, perturbatively up to 6th order in K/J

$$\begin{aligned} H_{\text{eff}} = &\sum_{\substack{a-\text{links}\\\langle ij\rangle}}\left(g_2\tau_{1i}^{(a)}\tau_{2i}^{(a)}\tau_{1j}^{(a)}\tau_{2j}^{(a)} + g_3\sum_{v=1,2}\tau_{vi}^{(a)}\tau_{vj}^{(a)}\right)\\ &+ g_6\sum_{\bigcirc,v}W_p^v \end{aligned} \tag{11}$$

where $W_p^{1(2)}$ is the flux operator defined on the honeycomb plaquettes on layer 1(2) as $W_p^v = \tau_{vi}^{(z)}\tau_{vj}^{(y)}\tau_{vk}^{(x)}\tau_{vl}^{(z)}\tau_{vm}^{(y)}\tau_{vn}^{(y)}$ (Fig. 2d). Please consult Supplementary Notes 4 and 5 for additional details. The coupling constants are $g_2 = -K^2/4J$, $g_3 = -K^3/J^2$, and $g_6 = -K^6/(8J)^5$. The g_2 term describes Kitaev interactions around inter-layer plaquettes while the g_3 term is a standard Kitaev interaction in each layer.

Note that g_6 terms promote uniform $W_p^{1(2)} = 1$ corresponding to a zero-flux low-energy manifold. This configuration is preserved by the remaining terms which commute with the $W_p^{1(2)} = 1$.

We first focus on the the leading g_2 terms, and define new operators $p_i^{(a)} = \tau_{i1}^{(a)}\tau_{2i}^{(a)}$, which unlike the τ's, all commute with each other. Furthermore their product amounts to $p_i^{(x)}p_i^{(y)}p_i^{(z)} = -1$.
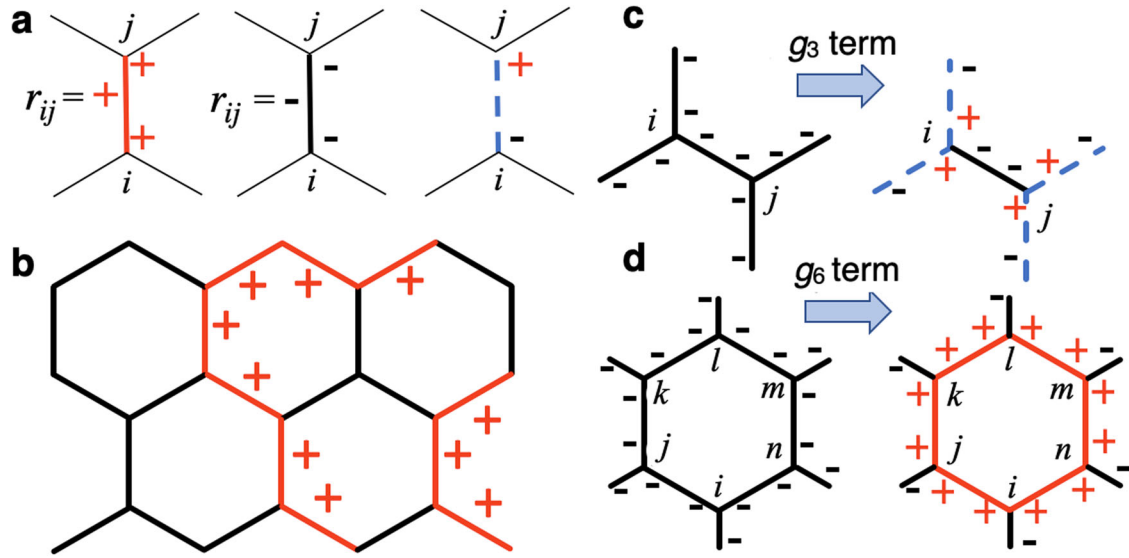
Fig. 2 Effective Hamiltonian in the large-J limit. a The bond operator $r_{ij} = \text{sgn}(p_i^{(a)}) = \text{sgn}(p_j^{(a)})$, which is defined in the ground-state manifold of the g_2 terms in Eq. (11). Red (black) solid lines correspond to $+(-)$ bonds. Conversely, bond configurations which include NNs with $\text{sgn}(p_i^{(a)}) \neq \text{sgn}(p_j^{(a)})$, shown here with blue dashed lines, are not labeled by r_{ij} bonds. These configurations correspond to excited states. **b** A ground-state manifold configuration which minimizes the g_2 terms and which also obeys the local product constraint, equivalent to an Ising Gauss's law. **c** g_3 terms on sites i, j flip four adjacent bonds resulting in an excited state. **d** g_6 terms change bond configurations from $r_{ij} = 1$ to $r_{ij} = -1$ around each plaquette and thus connect configurations which belong to the Ising Gauss' Law manifold.

Therefore, we use local basis states which are eigenstates of all p operators and which also satisfy the product rule: $\{|-,-,-\rangle, |-,+,+\rangle, |+,-,+\rangle, |+,+,-\rangle\}$ where $\pm$ denotes the eigenvalue of $p^{(a)}$, ($a = x, y, z$). The $g_2 < 0$ terms favor equal-$p^{(a)}$ states on NN sites. Therefore, in the GS manifold of the g_2 terms, it is possible to define bond variables $r_{ij} = \pm 1$ for pairs of $(+, +)$ and $(-, -)$ eigenvalues of $p_{i/j}^{(a)}$, respectively. For configurations that do not minimize the g_2 terms, r_{ij} is not defined (Fig. 2a). In addition to minimizing the g_2 terms, the GS manifold must also satisfy the local constraint due to $p_i^{(x)} p_i^{(y)} p_i^{(z)} = -1$. Taken together, these conditions are equivalent to bond configurations which obey an Ising Gauss's law $G_i^P = \prod_\gamma r_{ij} = -1$ (Fig. 2b). We stress that the r_{ij} bond variables and Gauss's law are only defined in the GS manifold of the g_2 terms.

Next, we examine the effect of g_3 and g_6 terms acting on the GS manifold obtained from the combined effects of the g_2 terms and local product constraints. Each $\tau_{1,2}^{(a)}$ acting on $|p_x, p_y, p_z\rangle$ preserves the corresponding a eigenvalues but flips the remaining two (see Supplementary Note 5). Therefore, the g_6 terms acting on a plaquette flips all of the r_{ij} bond variables therein (Fig 2d), leading to an effective term

$$-\kappa \sum_{\bigcirc} (|\bigcirc\rangle\langle\overline{\bigcirc}| + \text{H.c.}) \qquad (12)$$

In contrast, the single g_3 term on sites $\langle ij \rangle$, connects a ground-state configuration to excited states (Fig. 2 (c)). Consecutive application of g_3 terms around a plaquette leads to plaquette flips, but these processes are subdominant with respect to those due to the g_6 term.

As shown in Supplementary Note 6, the resonance term in Eq. (12), along with Gauss' law, describe Kitaev's toric code[36] on a honeycomb lattice. We thus conclude that the bilayer model in the limit of large inter-layer spin exchange interactions is in a gapped abelian $\mathbb{Z}_2$ topological QSL phase.

Self-consistent solutions for moiré superlattices

We generalize the Hartree approximation to include the effects of small-angle twists. We follow Ref. [43] to derive a low-energy theory

defined on the moiré extended BZ, as shown in Supplementary Note 7.

To allow for non-vanishing inter-layer interactions under arbitrary, small twist angles, we extend the former beyond overlapping NN pairs and allow for an implicit decay with increasing pair separation. In general, this entails a decay of the Fourier components $J(\mathbf{k})$ with $|\mathbf{k}|$, and involves interactions which are delocalized in the extended BZ. In the low-energy limit, the interactions are naturally limited to the vicinity of a discrete set of equivalent crystal momenta throughout the extended BZ. In practice, we keep only $J(\mathbf{k})$ with $|\mathbf{k}| \lesssim |2\mathbf{K}_{00}|$, or twice the distance from the origin to the nearest Dirac point (Eq. 59 in Supplementary Note 7). We also assume that the retained Fourier components are all comparable in magnitude. The restrictions on the values of $J(\mathbf{k})$ allow us to explicitly consider the Yao-Lee bilayer analogs of flat bands in twisted bilayer graphene[43]. However, our conclusions are independent of this approximation, as discussed in the following. We also limit the hybridization to states in the vicinity of Dirac points in neighboring moiré reciprocal unit cells. This truncation is justified in the low-energy limit, where small-momentum scattering processes are dominant.

The intra-layer terms amount to the usual Dirac fermions for the two layers, which are shifted with respect to each other due to twisting. The inter-layer interactions together with the approximations discussed previously can be written as

$$\mathcal{H}_I = -\frac{4J}{N} \sum_{nm} \left[\langle \chi_{00}^\dagger(\mathbf{q}) \rangle \chi_{nm}(\mathbf{q}) + \langle \chi_{00}(\mathbf{q}) \rangle \chi_{nm}(-\mathbf{q}) \right] + \text{H.c.} \qquad (13)$$

where

$$\begin{aligned}
\chi_{nm}(\mathbf{q}) = i \sum_{\mathbf{k}} \Big[& c_{1,\alpha;nm}^{\dagger,(\mu)}(\mathbf{k}) c_{2,\beta;nm}^{(\mu)}(\mathbf{k} - \mathbf{q}) \\
& + e^{-i\mathcal{G}_2 \cdot (\boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_\beta)} c_{1,\alpha;nm}^{\dagger,(\mu)}(\mathbf{k}) c_{2,\beta;n+1m}^{(\mu)}(\mathbf{k} - \mathbf{q}) , \\
& + e^{-i\mathcal{G}_3 \cdot (\boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_\beta)} c_{1,\alpha;nm}^{\dagger,(\mu)}(\mathbf{k}) c_{2,\beta;nm+1}^{(\mu)}(\mathbf{k} - \mathbf{q}) \Big]
\end{aligned} \qquad (14)$$

while

$$c_{a,1/2}^{\dagger,(\mu)}(\mathbf{K}_{00} + \mathbf{k} - n\mathbf{b_2} - m\mathbf{b_3}) = c_{a,1/2;nm}^{\dagger,(\mu)}(\mathbf{k}) \qquad (15)$$

are states with an effective Dirac dispersion which is shifted by the moiré reciprocal vectors $\mathbf{b}_{2,3}$ with respect to the Dirac point centered on the moiré first BZ at $n = m = 0$. $\mathbf{K}_{00}$ is the position of the Dirac point of layer 1 in the first BZ while $\mathcal{G}_{2,3}$ are the reciprocal unit vectors of layer 1. The sums over momenta $\mathbf{k}$ cover the extended moiré BZ, with an implicit cutoff. The vectors $\boldsymbol{\tau}_{a,\beta}$ denote the shift of the A, B sublattices in layers 1 and 2, respectively. $\mathbf{q}$ is a vector contained within a single moiré reciprocal unit cell. As already mentioned, our approximations, and the cutoff for $J(\mathbf{k})$ in particular, ensure that the form of the effective hybridization in Eqs. (13) and (14) bears a close resemblance to that of twisted bilayer graphene[43]. For more details on the MF procedure, please see Supplementary Note 7.

We consider two cases, one for $\mathbf{q} = 0$ corresponding to a uniform inter-layer hybridization, and another for finite $\mathbf{q} = \mathbf{q}_1$ where $\mathbf{q}_1 = -8\pi/3 \sin(\theta/2)\hat{y}$[43] which denotes the shift between the Dirac points in layers 1 and 2 in the first BZ due to twisting (see Fig. 3d). In both cases, $\langle \chi_{AA} \rangle = -\langle \chi_{BB} \rangle$ acquire finite expectation values whereas $\langle \chi_{AB} \rangle$ and $\langle \chi_{BA} \rangle$ remain pinned to zero. Our self-consistent calculations indicate that the critical coupling J_c/K for the $\mathbf{q} = \mathbf{q}_1$ solution is below its $\mathbf{q} = 0$ counterpart for the entire range of twist angles (Fig. 3a), indicating that the modulated hybridization is energetically favored. A finite-$\mathbf{q}$ hybridization connects states near inequivalent Dirac points in the moiré BZ and gaps the spectrum, as illustrated in Fig. 3b, effectively 'untwisting' the system. In contrast, for $\mathbf{q} = 0$, the spectrum remains gapless, (Fig. 3c). Consequently, the finite-$\mathbf{q}$ solution is preferred for any non-zero twist angle. The two solutions merge smoothly as $\theta \rightarrow 0$ since $\mathbf{q}_1$ vanishes in this limit, at which point the low-energy sectors match the self-consistent solutions of the un-twisted bilayer with AA stacking.

At the level of the Hartree approximation, our results indicate that the gap remains open as the small-angle twisting is turned on. Within the same approximation, we conclude that resulting phases are adiabatically connected with the AA-stacked bilayer in the large-J limit. Our results suggest that, beyond the Hartree approximation, the gap in the spin excitations of the bilayer survives, and that the GS remains in a net zero-spin state for small-angle twisting. We expect that the intra-layer interactions lift the extensive degeneracy of the orbitals, resulting in a gapped, $\mathbb{Z}_2$ QSL, as for the case with AA stacking.

For the gauge choice of uniform and identical bonds in both layers, the $\mathbf{q} = \mathbf{q}_1$ incommensurate, inter-layer hybridization breaks the translation symmetry of simple moiré pattern but preserves all other symmetries. It consequently triples the size of the moiré unit cell (Fig. 3e). Figure 3d shows the moiré (black) and folded (red) BZ's, respectively. The rotated Dirac cones at the corners of the moiré BZ are folded onto the Γ point. However, since the effective hybridization is not gauge invariant, this does not imply a true translation symmetry breaking, but instead demonstrates that small-angle twisting preserves the gapped $\mathbb{Z}_2$ QSL.

We note that the main conclusion of the preceding paragraphs, that twisting the AA-stacked bilayer by small angles preserves the gapped spectrum, does not rely on our assumptions concerning the cutoff in $J(\mathbf{k})$. Indeed, keeping only the leading $J(0)$ terms in the expression for the self-consistent hybridization (Eqs. 82-84 in Supplementary Note 7), which likewise connect pairs of Dirac points in the moiré BZ, leads to a similar conclusion in the Hartree approximation.

DISCUSSION

It is instructive to contrast the bilayer Yao-Lee model considered here with the bilayer Kitaev models of earlier works. For a bilayer Kitaev model, a mean-field study predicts gapped QSL and trivial dimer phases for intermediate and large values of the inter-layer coupling, respectively[31]. However, an exact diagonalization
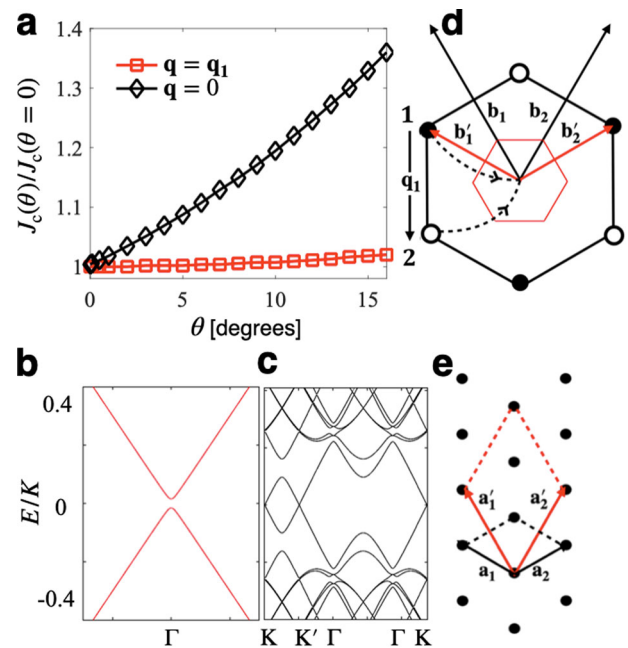
Fig. 3 Moiré superlattices of the Yao-Lee bilayer. a Critical inter-layer exchange $J_c(\theta)$, as a function of twisting angle θ, and for $\mathbf{q} = 0$ and $\mathbf{q} = \mathbf{q}_1$ in the effective interlayer hybridization $\langle \chi_{00}(\mathbf{q}) \rangle$ (Eq. (13)), in units of $J_c(\theta = 0)$. The $\mathbf{q} = \mathbf{q}_1$ solution is energetically preferred. The spectrum for (**b**) $\mathbf{q} = \mathbf{q}_1$ and (**c**) $\mathbf{q} = 0$ in units of the intra-layer coupling K (Eq. (2)). **d** Moiré reciprocal unit cells for $J < J_c$ (black) and $J > J_c$ (red). 1 and 2 denote the Dirac points of the two layers which are separated by $\mathbf{q}_1$ for $J < J_c$, corresponding to vanishing inter-layer hybridization. When the latter acquires a finite value for $J > J_c$ and $\mathbf{q} = \mathbf{q}_1$, the two Dirac points are shifted to the Γ point of the folded BZ, and are subsequently gapped. **e** Schematic illustration of the moiré lattice vectors for $J < J_c$ (black) and $J > J_c$ (red) corresponding to zero and non-zero ($\mathbf{q} = \mathbf{q}_1$) hybridization, respectively.

study[33] finds that a single phase transition between gapless QSL and trivial dimer phases occurs at a substantially weaker coupling $J/K \sim 0.06$. Our results indicate that the QSL phase remains stable in Yao-Lee bilayers for large but finite intra-layer couplings, while the trivial dimer phase emerges only in the absence of intra-layer terms ($K = 0$). The stability of the gapped QSL in the Yao-Lee bilayer can be attributed to the effect of the spin operators on the zero-flux GSs of the decoupled layers. In the Kitaev model, the spin operators create two visons, as shown in Supplementary Note 10. By contrast, the spin operators in the Yao-Lee model preserve the zero-flux GS manifold, since the spin and flux operators are associated with different DOF.

We studied the zero-temperature phase diagram of a bilayer Yao-Lee model with inter-layer interactions. For AA stacking, we determined that finite inter-layer singlet correlations gap the itinerant Majorana fermion spectrum. We also derived an effective Hamiltonian in the limit of large J/K, and demonstrated that it maps onto the toric code. In the absence of any additional transitions which close the gap, we concluded that the solutions obtained via the Hartree approximation are adiabatically connected to the large inter-layer interaction limit, leading to the stability of a topological gapped $\mathbb{Z}_2$ QSL. This phase persists for moiré superlattices under small-angle twisting. Detailed studies of the AB stacked phases and of the toric code models in the large inter-layer coupling limit are clearly desirable.

METHODS

For more details on the methods used to obtain the results shown in Figs. 1–3 please consult Supplementary Notes 7 and 8.

DATA AVAILABILITY

CODE AVAILABILITY

REFERENCES

1. Balents, L. Spin liquids in frustrated magnets. *Nature* **464**, 199–208 (2010).
2. Zhou, Y., Kanoda, K. & Ng, T.-K. Quantum spin liquid states. *Rev. Mod. Phys.* **89**, 025003 (2017).
3. Wen, X.-G. Colloquium: Zoo of quantum-topological phases of matter. *Rev. Mod. Phys.* **89**, 041004 (2017).
4. Knolle, J. & Moessner, R. A Field Guide to Spin Liquids. *Annu. Rev. Condens. Matter Phys.* **10**, 451–472 (2019).
5. Broholm, C. et al. Quantum spin liquids. *Science* **367**, eaay0668 (2020).
6. Kitaev, A. Anyons in an exactly solved model and beyond. *Ann. Phys.* **321**, 2–111 (2006).
7. Hermanns, M., Kimchi, I. & Knolle, J. Physics of the Kitaev Model: Fractionalization, Dynamic Correlations, and Material Connections. *Annu. Rev. Condens. Matter Phys.* **9**, 17–33 (2018).
8. Hwan Chun, S. et al. Direct evidence for dominant bond-directional interactions in a honeycomb lattice iridate Na_2IrO_3. *Nat. Phys.* **11**, 462–466 (2015).
9. Takagi, H., Takayama, T., Jackeli, G., Khaliullin, G. & Nagler, S. E. Concept and realization of Kitaev quantum spin liquids. *Nat. Rev. Phys.* **1**, 264–280 (2019).
10. Lee, I. et al. Fundamental Spin Interactions Underlying the Magnetic Anisotropy in the Kitaev Ferromagnet CrI_3. *Phys. Rev. Lett.* **124**, 017201 (2020).
11. Blei, M. et al. Synthesis, engineering, and theory of 2D van der Waals magnets. *Appl. Phys. Rev.* **8**, 021301 (2021).
12. Tong, Q., Liu, F., Xiao, J. & Yao, W. Skyrmions in the moiré of van der Waals 2D Magnets. *Nano Lett.* **18**, 7194–7199 (2018).
13. Hejazi, K., Luo, Z.-X. & Balents, L. Noncollinear phases in moiré magnets. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10721–10726 (2020).
14. Hejazi, K., Luo, Z.-X. & Balents, L. Heterobilayer moiré magnets: Moiré skyrmions and commensurate-incommensurate transitions. *Phys. Rev. B* **104**, L100406 (2021).
15. Akram, M. & Erten, O. Skyrmions in twisted van der Waals magnets. *Phys. Rev. B* **103**, L140406 (2021).
16. Xiao, F., Chen, K. & Tong, Q. Magnetization textures in twisted bilayer $CrX_3(X = Br, I)$. *Phys. Rev. Research* **3**, 013027 (2021).
17. Akram, M. et al. Moiré Skyrmions and Chiral Magnetic Phases in Twisted CrX3 (X = I, Br, and Cl) Bilayers. *Nano Lett.* **21**, 6633–6639 (2021).
18. Xu, Y. et al. Coexisting ferromagnetic–antiferromagnetic state in twisted bilayer CrI_3. *Nat. Nanotechnol.* **17**, 143–147 (2021).
19. Song, T. et al. Direct visualization of magnetic domains and moiré magnetism in twisted 2D magnets. *Science* **374**, 1140–1144 (2021).
20. Yao, H. & Lee, D.-H. Fermionic Magnons, Non-Abelian Spinons, and the Spin Quantum Hall Effect from an Exactly Solvable Spin-1/2 Kitaev Model with SU(2) Symmetry. *Phys. Rev. Lett.* **107**, 087205 (2011).
21. Kugel, K. I. & Khomskii, D. I. The Jahn-Teller effect and magnetism: transition metal compounds. *Sov. Phys. Uspekhi* **25**, 231–256 (1982).
22. Yao, H., Zhang, S.-C. & Kivelson, S. A. Algebraic Spin Liquid in an Exactly Solvable Spin Model. *Phys. Rev. Lett.* **102**, 217202 (2009).
23. Wang, F. & Vishwanath, A. Z_2 spin-orbital liquid state in the square lattice Kugel-Khomskii model. *Phys. Rev. B* **80**, 064413 (2009).
24. Wu, C., Arovas, D. & Hung, H.-H. Γ-matrix generalization of the Kitaev model. *Phys. Rev. B* **79**, 134427 (2009).
25. de Carvalho, V. S., Freire, H., Miranda, E. & Pereira, R. G. Edge magnetization and spin transport in an SU(2)-symmetric Kitaev spin liquid. *Phys. Rev. B* **98**, 155105 (2018).
26. Seifert, U. F. P. et al. Fractionalized Fermionic Quantum Criticality in Spin-Orbital Mott Insulators. *Phys. Rev. Lett.* **125**, 257202 (2020).
27. Chulliparambil, S., Seifert, U. F. P., Vojta, M., Janssen, L. & Tu, H.-H. Microscopic models for Kitaev's sixteenfold way of anyon theories. *Phys. Rev. B* **102**, 201111 (2020).
28. Natori, W. M. H. & Knolle, J. Dynamics of a Two-Dimensional Quantum Spin-Orbital Liquid: Spectroscopic Signatures of Fermionic Magnons. *Phys. Rev. Lett.* **125**, 067201 (2020).
29. Chulliparambil, S., Janssen, L., Vojta, M., Tu, H.-H. & Seifert, U. F. P. Flux crystals, Majorana metals, and flat bands in exactly solvable spin-orbital liquids. *Phys. Rev. B* **103**, 075144 (2021).
30. Tsvelik, A. M. & Coleman, P. Order Fractionalization in a Kitaev-Kondo model. *Phys. Rev. B* **106**, 125144 (2022).
31. Seifert, U. F. P. et al. Bilayer Kitaev models: Phase diagrams and novel phases. *Phys. Rev. B* **98**, 155101 (2018).
32. Tomishige, H., Nasu, J. & Koga, A. Interlayer coupling effect on a bilayer Kitaev model. *Phys. Rev. B* **97**, 094403 (2018).
33. Tomishige, H., Nasu, J. & Koga, A. Low-temperature properties in the bilayer Kitaev model. *Phys. Rev. B* **99**, 174424 (2019).
34. May-Mann, J. & Hughes, T. L. Twisted Kitaev bilayers and the moiré Ising model. *Phys. Rev. B* **101**, 245126 (2020).
35. Xu, H., Zhou, Z., Wang, X., Wang, L. & Wang, Y. Quantum Monte Carlo simulations of the attractive SU(3) Hubbard model on a honeycomb lattice. Preprint at https://arxiv.org/abs/1912.11233 (2019).
36. Kitaev, A. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303**, 2–30 (2003).
37. Rozhkov, A., Sboychakov, A., Rakhmanov, A. & Nori, F. Electronic properties of graphene-based bilayer systems. *Phys. Rep.* **648**, 1 (2016). Electronic properties of graphene-based bilayer systems.
38. Natori, W. M. H., Nutakki, R., Pereira, R. G. & Andrade, E. C. SU(4) Heisenberg model on the honeycomb lattice with exchange-frustrated perturbations: Implications for twistronics and Mott insulators. *Phys. Rev. B* **100**, 205131 (2019).
39. Xu, C. et al. Possible Kitaev Quantum Spin Liquid State in 2D Materials with $S = 3/2$. *Phys. Rev. Lett.* **124**, 087205 (2020).
40. Stavropoulos, P. P., Pereira, D. & Kee, H.-Y. Microscopic Mechanism for a Higher-Spin Kitaev Model. *Phys. Rev. Lett.* **123**, 037203 (2019).
41. Yamada, M. G., Oshikawa, M. & Jackeli, G. Emergent SU(4) Symmetry in $\alpha − ZrCl_3$ and Crystalline Spin-Orbital Liquids. *Phys. Rev. Lett.* **121**, 097201 (2018).
42. Lieb, E. H. Flux Phase of the Half-Filled Band. *Phys. Rev. Lett.* **73**, 2158–2161 (1994).
43. Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12233–12237 (2011).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

E.M.N., R.M., and O.E. contributed to the design of this study, to the interpretation of the data, and to the drafting and substantial revision of the text. M.A. and A.V. contributed to the creation of new software used in the work, to the acquisition, analysis, and interpretation of the data, and to the drafting of the text.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41535-023-00541-2.

Correspondence and requests for materials should be addressed to Emilian Marius Nica.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.