Twisted bilayer U(1) Dirac spin liquids

Zhu-Xi Luo [1,*] Urban F. P. Seifert,[1,*] and Leon Balents[1,2]

[1]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA*
[2]*Canadian Institute for Advanced Research, Toronto, ON, Canada M5G 1M1*

When two layers of two-dimensional materials are assembled with a relative twist, moiré patterns arise, inducing a tremendous wealth of exotic phenomena. In this work, we consider twisting two triangular lattices hosting Dirac quantum spin liquids. A single decoupled layer is described by compact quantum electrodynamics in 2+1 dimensions (QED$_3$) with an emergent U(1) gauge field, which is assumed to flow to a strongly interacting fixed point in the IR with conformal symmetry. We use recent results for the quantum numbers of monopole operators, which tunnel 2π fluxes of the compact gauge field. It is found that, in the bilayer system, interlayer monopole tunneling is a symmetry-allowed relevant perturbation which induces an (ordering) instability. We show using perturbation theory that upon twisting the two layers the system remains unstable under the interlayer interaction, but any finite twist angle softens this instability compared to the untwisted case. To analyze the resulting phase induced by the (twisted) interlayer tunneling, we use "conformal mean-field theory", which reduces the interacting bilayer system to two copies of QED$_3$ coupled to background fields which are to be determined self-consistently. In the weak-coupling regime, where the interlayer coupling is weak compared to the energy scale set by the moiré lattice constant, we solve the self-consistency equations perturbatively. In the limit of strongly coupled layers, a local scaling approximation is used, and we find that the magnetically ordered state exhibits a lattice of magnetic vortices, with the lattice constant tunable through the twisting angle.

I. INTRODUCTION

A. Motivation

Van der Waals materials are two-dimensional atomic crystals with strong in-plane covalent bonds and weak interlayer van der Waals (vdW) interactions, which can be exfoliated down to monolayer. In recent years, numerous experiments have strongly advanced capabilities to prepare and control these materials with atomic precision, and assemble them like quantum Legos [1,2].

A particularly fruitful avenue in this regard consists in exploiting moiré physics, which arises from relative twisting or lattice mismatches between different layers, leading to quantum interferences which quench the energy scales in the system, often permitting a strongly coupled, interaction-dominated regime. One leading example of this manipulation of electronic properties, or twistronics [3] in vdW heterostructures, is that of the twisted bilayer graphene [4,5], where exotic superconductivity and correlated insulating behaviors have been observed.

Recently, the study of moiré heterostructures has been extended to systems with magnetic ordering, where interesting noncollinear magnetic ordering, magnon behaviors, and topological spin textures have been found [6–15]. A natural followup question is then as follows: What happens when one twists quantum spin liquids, i.e., strongly correlated spin systems that intrinsically lack magnetic ordering [16]?

In this work, we will focus on the twisted bilayer of U(1) Dirac spin liquids (DSL) described in the low-energy limit by $N = 4$ flavors of Dirac fermions coupled minimally to an emergent U(1) gauge field. This effective theory is known as QED$_3$, quantum electrodynamics in 2+1 dimensions. In the context of U(1) DSL, the Dirac fermions are fractionalized quasiparticles which carry the spin degrees of freedom of the electrons, thereby named spinons. In the absence of monopole and SU(N) symmetry breaking (which is avoided for a sufficiently large number of fermion flavors N), QED$_3$ is assumed to be stable and to flow to a strongly coupled conformal fixed point in the IR [17–22].

Our motivation for focusing on U(1) DSL is twofold: On the one hand, there is evidence which suggests that such a state could be the ground state in realistic and experimentally relevant microscopic spin models, thus constituting a prime example of an exotic highly entangled strongly correlated magnetic state of matter. Dirac spin liquids were originally studied in the context of high-T_c superconductors [23] on the square lattice, and later examined on other lattices [24–27] as well. In particular, on the triangular lattice with both nearest and next nearest neighbor spin couplings, there is considerable numerical evidence suggesting the presence of a Dirac spin liquid [28–31]. Material candidates have been proposed as well, including Ba$_8$CoNb$_6$O$_{24}$ [32] and 1T-TaS$_2$ [33], the latter being a van der Waals material.

On the other hand, U(1) DSLs have commonly been described as "parent states" of competing orders, as they offer a unified framework to describe various (seemingly unrelated) magnetic (and valence-bond paramagnetic) ordered

---

[*]These authors contributed equally to this work.

144437-1

states [25] by inducing instabilities of the U(1) DSL. Our present study hence can also be understood as exploring the viability of the concept of the DSL as a parent state, both in homogeneous layered systems and upon spatial modulations of the interlayer coupling as an additional control knob, potentially stabilizing various exotic competing ordered states with moiré supermodulations.

For the spin liquid on the triangular lattice, recent works by Song *et al.* have shown that all monopoles in the effective QED₃ theory carry nontrivial symmetry quantum numbers [34,35]. The lowest-order symmetry-allowed term is a triple monopole [34,35], very likely to be irrelevant based on large-N expansion [36]. There can also be four fermion terms that break the SU(4) flavor symmetry, which are found to be irrelevant in ϵ expansion [20], but relevant in large-N expansions [37]. While the stability of the U(1) Dirac spin liquid on the triangular lattice is hence an open question, any instabilities inherent in a single two-dimensional layer are relatively weak, if they are present at all. By contrast, we observe in this work a strong instability (associated with highly relevant interlayer interactions) of the bilayer system, which can be tuned through twisting.

B. Summary of results and outline

We briefly summarize the model and our results in this subsection. In each layer, the effective low-energy theory is that of the $N_f = 4$ QED₃, i.e., four flavors of Dirac fermions coupled to a U(1) gauge field, which we assume to flow to a conformal fixed point [17–20], facilitating our use of conformal data to constrain correlation functions and construct effective actions.

We consider the interlayer coupling of fermion bilinear masses in the two layers as well as interlayer tunneling of monopoles of the U(1) gauge fields. Guided by latest conformal bootstrap results [38,39], we take the monopole tunneling terms to be the most relevant. For a relative rigid twist of angle θ, symmetry analysis predicts monopole tunneling terms to be of the form

$$\mathcal{L}_1 = J_v \sum_{a=1}^{3} f_v^a(x)\Phi_{1a}^{\dagger}\Phi_{2a} + J_s f_s(x) \sum_{a=4}^{6} \Phi_{1a}^{\dagger}\Phi_{2a} + \text{H.c.},$$

$$f_v^a(x) = e^{i\vec{q}_a \cdot \vec{x}/2}, \quad \text{and} \quad f_s(x) = \sum_{i=1}^{3} e^{i\vec{q}_i \cdot \vec{x}}, \tag{1}$$

where Φ_{la} is the monopole annihilation operator in layer l with flavor a. There are altogether six different flavors, which fall into two classes labeled by the indices v ($a = 1, 2, 3$) and s ($a = 4, 5, 6$) standing for valence bond solid (VBS) and spin channels, respectively. Proliferating the corresponding monopoles (or a linear combination thereof) yields a corresponding VBS or magnetic order parameter. The finite twist angle leads to supermodulations of the interlayer couplings, i.e., the functions $f_\alpha(x)$ are periodic on the *moiré scale* with moiré reciprocal vectors $\vec{q}_a = \theta\hat{z} \times \vec{K}_a$, where $\vec{K}_a$ is the three C_3-symmetry-related Brillouin zone corner vectors of the parent layers.

We first analyze the interlayer interaction using perturbation theory, using the fact that monopole correlation functions are strongly constrained by conformal symmetry. For $\theta = 0$,

one can predict simply from the scaling dimension Δ_Φ that this perturbation theory is divergent, and the uniform coupling of layers is a relevant perturbation in the renormalization group sense. This divergence appears already at quadratic order in the couplings. For $\theta > 0$, the couplings oscillate spatially and have zero mean. This immediately eliminates any divergence arising at quadratic order, which suggests that, keeping θ fixed and taking J_v, J_s arbitrarily small, the coupling of layers might be irrelevant.

A key result of our analysis is that this is not the case. We find that even for $\theta > 0$, a perturbative divergence persists, but appears only at higher order in the couplings. This leads to a breakdown of perturbation theory which we identify as an instability. That is, even at nonzero θ (finite a_m), we find that the twisted bilayer system is unstable in the thermodynamic limit (i.e., sample size) $L \to \infty$ for any infinitesimally small J. The transfer of the instability from quadratic to higher order with the introduction of a twist leads to the conclusion that the instability is *softened* upon introducing a nonzero twist angle θ; in this case, the power-law finite-size scaling of the critical J_c reads as

$$J_c \sim (a_m L)^{-(3/2-\Delta_\Phi)}, \tag{2}$$

compared to $J_c \sim L^{-(3-2\Delta_\Phi)}$ in the homogeneous case (absence of any twist angle). Hence, the spatial modulation of the interlayer coupling due to the twisting effectively *renders the interlayer coupling less relevant*.

To study the nature of the phase resulting from the instability, we make a mean-field approximation by replacing the interlayer interactions with a mean-field action of the form

$$\Phi_{1a}^{\dagger}\Phi_{2a} + \text{H.c.} \to \langle\Phi_{1a}^{\dagger}\rangle_{\text{mf}}\Phi_{2a} + \Phi_{1a}^{\dagger}\langle\Phi_{2a}\rangle_{\text{mf}} + \text{H.c.}, \tag{3}$$

where the mean fields $\langle\Phi_{l,a}\rangle$ are to be determined self-consistently. This approach is in direct analogy to "chain mean-field theory" [40–42] used very successfully to describe coupled one-dimensional spin chains.

Crucially, the presence of both an energy and length scale, namely, the interlayer coupling J and the moiré lattice scale, allows us to distinguish between *weak-coupling* $J \ll a_m^{2\Delta_\Phi-3}$ and *strong-coupling* $J \gg a_m^{2\Delta_\Phi-3}$ regimes. Here, $\Delta_\Phi \simeq 1.02$ is the scaling dimension of the monopole operator at the QED₃ fixed point, estimated from large-N and bootstrap methods [39].

On one hand, considering the mean-field theory in the weak-coupling limit, one may use conformal perturbation theory to quadratic order to obtain above monopole expectation values in each layer, and perform a qualitative analysis of quartic terms to fix accidental degeneracies. We find that in this perturbative regime, nontrivial solutions to the mean-field equations emerge when $J \sim (a_m L)^{-(\frac{3}{2}-\Delta_\Phi)}$, which is precisely the same scaling as obtained in perturbation theory without any mean-field approximation [see Eq. (2)].

To determine the form of the order parameter in the ordered phase, we first consider spin monopoles $\{\Phi_{l4}, \Phi_{l5}, \Phi_{l6}\}$, which are generated by spin-spin interactions in a microscopic theory and are thus expected to be dominant. At the finite-size critical point, the spatial dependence of the monopole expectation values (which correspond to the Néel order parameters

on the two parent triangular lattices) read as

$$\langle \boldsymbol{\Phi}_1(\vec{x}) \rangle_{\mathrm{mf}} \approx -\frac{A}{|J_{\mathrm{s}}|}\tilde{\boldsymbol{h}}(0) - \frac{f_{\mathrm{s}}(x)}{3J_{\mathrm{s}}}\tilde{\boldsymbol{h}}(0),$$

$$\langle \boldsymbol{\Phi}_2(\vec{x}) \rangle_{\mathrm{mf}} \approx \frac{A}{J_{\mathrm{s}}}\tilde{\boldsymbol{h}}(0) + \frac{f_{\mathrm{s}}^*(x)}{3|J_{\mathrm{s}}|}\tilde{\boldsymbol{h}}(0). \qquad (4)$$

Here $\tilde{\boldsymbol{h}}(0)$ is a constant three-dimensional complex vector. $A \sim L/a_{\mathrm{m}} \gg 1$ is a constant proportional to the ratio between the IR cutoff (sample size L) and the moiré lattice constant ($a_{\mathrm{m}} \sim a_0/\theta$). Therefore, the first uniform terms in the two equations above dominate, and there are some corrections with modulations of moiré scale. The coupling of VBS monopoles $\{\Phi_{l1}, \Phi_{l2}, \Phi_{l3}\}$ can be analyzed similarly.

On the other hand, in the strong-coupling regime of $J_a \gg a_{\mathrm{m}}^{2\Delta_\Phi - 3}$, the spatial modulations of the tunneling amplitudes are slow enough such that locally, the bilayer system can be well approximated by that of a uniform stacking. This limit therefore facilitates a scaling ansatz for the mean-field free energy from which correlation functions can be determined (the resulting local degeneracy is lifted by symmetry-allowed gradient terms). Again focusing on the interlayer coupling of spin monopoles, we find that the monopole expectation values, and thus the magnetic order parameter, vanish at the centers of the triangular plaquettes in the moiré triangular lattice, precisely where the interlayer-coupling function $f_{\mathrm{s}}(x)$ possesses zeros, with $|\langle \boldsymbol{\Phi}_l(x) \rangle| \sim |f_{\mathrm{s}}(x)|^{\frac{\Delta_\Phi}{3-2\Delta_\Phi}}$. Around these zeros, $f_{\mathrm{s}}(x)$ possesses a nonzero winding number (by $\pm 2\pi$ going anticlockwise around rightward- and leftward-facing triangles), which leads to the expectation values of the spin monopoles in the two layers to differ by a spatially modulating phase $\langle \boldsymbol{\Phi}_2 \rangle_{\mathrm{mf}} / \langle \boldsymbol{\Phi}_1 \rangle_{\mathrm{mf}} = e^{i \arg f_{\mathrm{s}}(x)}$. Note that this implies that the interlayer exchange symmetry of the system is broken.

We have thus found that, in the limit of strong coupling (to which the *relevant* interlayer interaction inevitably flows), the model *realizes an exotic magnetic vortex lattice on the moiré scale*. The latter is experimentally tunable by varying the twist angle θ. An exemplary configuration for such a moiré vortex lattice, showing the order parameter configurations in both layers along with microscopic spin configurations, is depicted in Fig. 1.

One may argue that one could obtain bilayer ordered states qualitatively resembling our results by coupling together order parameters for *classically ordered* phases, such as the 120° Néel antiferromagnetic phases, in the two layers [resulting bilayer ordered states are expected to be similar to what we found but with explicitly fixed order parameter magnitudes, while within our framework they vary with critical exponents related to operator scaling dimensions of the (conformal) fixed point theory]. However, we emphasize that it is *a priori* unclear if the (twisted) bilayer coupling induces an ordering instability of the two Dirac spin liquids (which, in a single layer, are stable phases). The problem of stability of the bilayer DSL thus necessitates studying perturbations to the IR fixed point, which is described by a conformal field theory specified by scaling dimensions and an operator algebra. In contrast, a nonlinear sigma model formalism for the intertwined order parameters of valence bond solid and magnetic
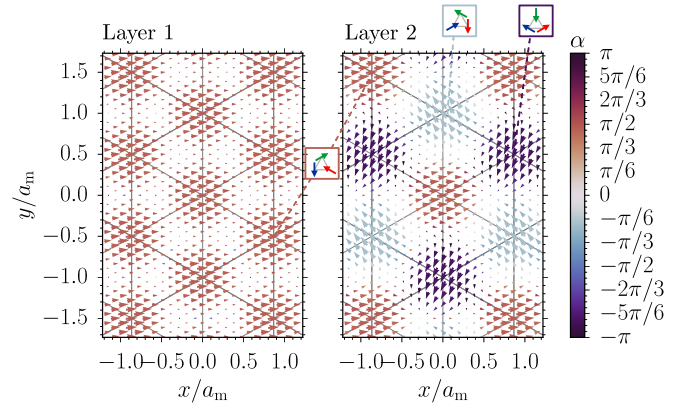

FIG. 1. Schematic example order parameters $\langle \boldsymbol{\Phi}_l \rangle$ (on the moiré lattice, indicated by gray lines) for the two layers in the strong-coupling limit, where $J_{\mathrm{s}} \gg a_{\mathrm{m}}^{2\Delta_\Phi - 3}$. The length of the plotted arrows corresponds to the magnitude $|\langle \boldsymbol{\Phi}_l \rangle|$, the direction of the plotted arrows corresponds to $\mathrm{Re}[\langle \boldsymbol{\Phi}_l \rangle / |\langle \boldsymbol{\Phi}_l \rangle|]$, and the angle $\alpha \equiv \arg \mathrm{Im}[\langle \boldsymbol{\Phi}_l \rangle / |\langle \boldsymbol{\Phi}_l \rangle|]$ is color coded. The insets depict the *spin configurations* on the A (blue), B (red), and C (green) sublattices of the parent layers, determined via Eq. (10). As visible, the magnetic order parameter in both layers vanishes at the centers of the moiré triangles, leading to *moiré vortices* with a nontrivial winding number of the relative angle of the order parameters between the two layers. Note that this configuration spontaneously breaks the interlayer exchange symmetry.

phases would not be able to describe the physics at the fixed point.

The remainder of this paper is organized as follows. We review Dirac spin liquids (i.e., the physics of a single layer) and their conformal low-energy field theory in Sec. II. In Sec. III, we derive a continuum description for the interlayer interactions based on symmetry principles. We use perturbation theory to analyze the stability of the system under the (twisted) interlayer interaction in Sec. IV. In Sec. V, we develop a mean-field treatment for the interlayer monopole tunneling term, and solve the resulting mean-field equations perturbatively in Sec. V, followed by a strong-coupling analysis in Sec. VII. Finally, in Sec. VIII, we present a conformal renormalization group analysis for homogeneous interlayer coupling along with the summary and outlook.

II. U(1) DIRAC SPIN LIQUIDS

In this work, we will be interested in spinons dispersing on the triangular lattice with a staggered $\pi/0$ flux background, leading to two Dirac cones per spin at momenta $\vec{K}_1 = 2\pi(1/3, -1/\sqrt{3})^\top/a_0$ and $\vec{K}_2 = (-4\pi/3, 0)^\top/a_0$ at zero energy. In the long-wavelength (continuum) limit, the system can be described in terms of (2+1)-dimensional quantum electrodynamics (QED$_3$) with the Euclidean Lagrangian

$$\mathcal{L}_{\mathrm{QED}_3} = \sum_{i=1}^{N=4}[-\bar{\psi}_i \gamma^\mu (\partial_\mu - i a_\mu)\psi_i] + \frac{1}{4g^2} f_{\mu\nu}f^{\mu\nu}, \qquad (5)$$

where each ψ_i is a two-component Dirac fermion with $N = 4$ flavors (2 spin×2 valley degrees of freedom), γ^μ are the gamma matrices, a_μ denotes the emergent U(1) gauge field

with field strength tensor $f_{\mu\nu}$, and g is a coupling constant. As visible from (5), the $SU(2)_{\text{spin}} \times SU(2)_{\text{valley}}$ symmetry becomes enhanced to SU(4) at low energies. In contrast to quantum electrodynamics in 3+1 dimensions, the coupling constant g^2 has unit mass dimension and thus the theory flows to strong coupling in the IR. In particular, studies of QED_3 with N flavors have shown that for sufficiently large N, the IR fixed point possesses conformal symmetry, while the fate of the theory for small N is uncertain [43–46].

The Dirac fermions in (5) can become gapped upon adding (or spontaneously generating) bilinear masses $M^{\mu\nu} = \bar{\psi}\sigma^{\mu}\tau^{\nu}\psi$, with σ^{μ} (τ^{ν}) denoting Pauli matrices acting on spin (valley) components of the Dirac fermions (note that, as we work in a flat Euclidean spacetime, we are free to use upper and lower indices for notational convenience). Note that $M^{\mu\nu}$ transforms in a reducible representation of SU(4) which splits as $16 = 1 \oplus 15$ into the singlet and 15-dimensional adjoint irreducible representations. While adding mass terms explicitly to (5) is forbidden by symmetry, a scenario of spontaneous chiral symmetry breaking has been suggested for a small enough flavor number N_c, with the precise nature of the phase structure and the value of N_c under investigation [43–46]. In the following, we will not consider such scenario and instead assume that the theory at $N = 4$ flows to the conformal IR fixed point, as supported by recent estimates that give $N_c < 4$ [18,47].

We further note that in writing the action $S_{QED_3} = \int d^3x \, \mathcal{L}_{QED_3}$ as the (continuum) low-energy theory for the bilayer U(1) Dirac spin liquid, one implicitly assumes the presence of a UV cutoff scale given by the (single-layer) lattice spacing a_0 beyond which nonuniversal microscopic (lattice) details of the interacting spin system are of importance. In real space, the UV cutoff gives a lower bound $a_0 \leqslant |x - y|$ on the separation of two operator insertions at points x, y.

A. Monopole operators

As written, the theory $\mathcal{L}_{QED_3}$ is endowed with a *topological* conserved current $j^{\mu}_{\text{top}} = \epsilon^{\mu\nu\rho}\partial_{\nu}a_{\rho}/(2\pi)$ corresponding to the conservation of the magnetic flux of the emergent U(1) gauge field. However, this conservation law is an artifact of the continuum limit, considering that (5) follows from a lattice gauge theory with a *compact* U(1) gauge field, we may consider operators which insert a 2π (emergent) magnetic flux or a multiple thereof. These operators are commonly referred to as monopole operators $\mathcal{M}(x)$, which carry unit charge under the $U(1)_{\text{top}}$ symmetry. As these operators cannot be written in terms of the fermions $\psi(x)$ or the gauge field a_{μ}, it is convenient to employ the theory's conformal symmetry. This allows us to characterize these operators through the state-operator correspondence as scalar primary operators with some scaling dimension Δ_{Φ} which may be evaluated in a controlled manner by quantizing the theory on $S^2 \times \mathbb{R}$ with a given monopole configuration, and performing a large-N expansion [48]. To leading order in $1/N$, this amounts to quantizing the Dirac operator on a sphere which is pierced by a 2π magnetic flux, yielding $N = 4$ fermionic zero modes (one zero mode per flavor). By gauge invariance, the four zero modes must be half-filled, so that there are $\binom{4}{2} = 6$
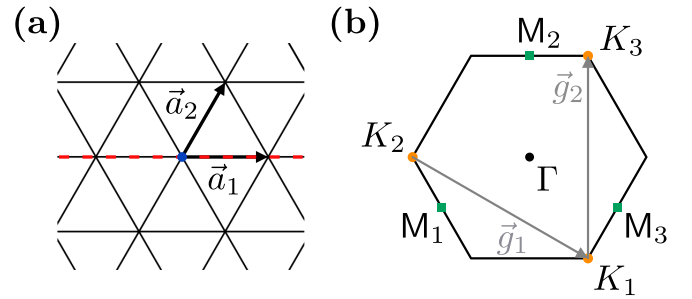

FIG. 2. (a) Basis vectors and symmetry operations on the triangular lattice. The blue dot is the rotation center for C_6, and the red dashed line is the reflection axis. (b) Hexagonal Brillouin zone on the triangular lattice with the high-symmetry Γ, K, and M points and reciprocal lattice vectors $\vec{g}_1, \vec{g}_2$.

distinct physical states resulting from filling the fermionic zero modes associated with a monopole operator. Schematically, these physical monopole operators may be written as [34,35]

$$\Phi_{\alpha\beta} \sim f_{\alpha}^{\dagger}f_{\beta}^{\dagger}\mathcal{M}^{\dagger}, \qquad (6)$$

where $f_{\alpha}^{\dagger}$ are fermionic zero-mode creation operators, so that $\Phi_{\alpha\beta}$ must transform in the six-dimensional antisymmetric representation of SU(4). It is convenient to employ the isomorphism $SO(6) = SU(4)/\mathbb{Z}_2$ such that the monopole operators Φ_a with $a = 1, \ldots, 6$ are taken to transform in the defining (vector) representation of SO(6). This also implies that the $SU(2)_{\text{spin}}$, $SU(2)_{\text{valley}}$ subgroups of SU(4) are isomorphic to the respective $SO(3)_{\text{s/v}}$ subgroups of SO(6). In addition, the monopole operator is odd under both the SO(6) center and a π rotation in $U(1)_{\text{top}}$, so the global symmetry group of the low-energy theory is $SO(6) \times U(1)_{\text{top}}/\mathbb{Z}_2$, together with charge conjugation, time reversal $\mathcal{T}$, and Lorentz symmetries. The discrete microscopic (UV) symmetries are shown in Fig. 2 and include translations $T_{1,2}$, reflection R, sixfold rotation C_6, and time reversal $\mathcal{T}$, which act nontrivially on the operators of the continuum (IR) field theory. While the transformation properties of the singlet and adjoint masses straightforwardly follow from the microscopic implementation of above symmetries, the appropriate transformations of monopole operators are given by combinations of discrete Lorentz symmetries, SO(6) transformations (due to the zero modes), and $U(1)_{\text{top}}$ rotations resulting from the microscopic symmetries acting on the filled Dirac sea with a $\pm 2\pi$ background flux. In this spirit, the action of microscopic symmetries on the monopole operator has been found and tabulated by Song *et al.* in Refs. [34,35] using both numerical methods as well as a study of the band topology of spinons. For reference, we reproduce their results in Table I.

We conclude this section by relating the continuum field theory monopole operators to symmetry-equivalent microscopic observables. From Table I, the first three monopoles $\langle \Phi_{1,2,3} \rangle$ are time-reversal invariant. Further imposing spin-rotational invariance, the corresponding microscopic operators at lowest order are of the form $S_i \cdot S_j$, consistent with the expectation that these three monopoles describe VBS order.

TABLE I. Transformation laws of the monopoles under microscopic symmetries. The second and third columns represent the transformation under translation along $\vec{a}_1/\vec{a}_2$ in Fig. 2, fourth column the reflection symmetry, fifth column the sixfold rotation symmetry, and last column the time-reversal symmetry. This table is adapted from Table 2 in Ref. [34].

	T_1	T_2	R	C_6	$\mathcal{T}$
Φ_1	$e^{i\frac{\pi}{3}}\Phi_1$	$e^{-i\frac{\pi}{3}}\Phi_1$	$-\Phi_3$	$\Phi_2^\dagger$	$\Phi_1^\dagger$
Φ_2	$e^{-i\frac{2\pi}{3}}\Phi_2$	$e^{-i\frac{\pi}{3}}\Phi_2$	Φ_2	$-\Phi_3^\dagger$	$\Phi_2^\dagger$
Φ_3	$e^{i\frac{\pi}{3}}\Phi_3$	$e^{i\frac{2\pi}{3}}\Phi_3$	$-\Phi_1$	$-\Phi_1^\dagger$	$\Phi_3^\dagger$
$\Phi_{4/5/6}$	$e^{-i\frac{2\pi}{3}}\Phi_{4/5/6}$	$e^{i\frac{2\pi}{3}}\Phi_{4/5/6}$	$\Phi_{4/5/6}$	$-\Phi_{4/5/6}^\dagger$	$-\Phi_{4/5/6}^\dagger$

The simplest choice that can recover Table I is

$$S_i \cdot S_{i+\delta_1} = \text{Re}[e^{i\vec{k}_1 \cdot \vec{R}_i}\Phi_1(\vec{R}_i)],$$

$$S_i \cdot S_{i+\delta_2} = \text{Re}[e^{i\vec{k}_2 \cdot \vec{R}_i}\Phi_2(\vec{R}_i)],$$

$$S_i \cdot S_{i+\delta_3} = -\text{Re}[e^{i\vec{k}_3 \cdot \vec{R}_i}\Phi_3(\vec{R}_i)], \quad (7)$$

where $S_i \cdot S_{i+\delta_j}$ with $j = 1, 2, 3$ labeling the bonds along the directions $\vec{a}_1$, $\vec{a}_2 - \vec{a}_1$ and $-\vec{a}_1 - \vec{a}_2$, respectively, with

$$\vec{a}_1 = a_0(1, 0)^\top \quad \text{and} \quad \vec{a}_2 = a_0(1/2, \sqrt{3}/2)^\top \quad (8)$$

being two lattice vectors, and the corresponding reciprocal lattice vectors are given by $\vec{g}_1 = (2\pi, -2\pi/\sqrt{3})^\top/a_0$ and $\vec{g}_2 = (4\pi/\sqrt{3}, 0)^\top/a_0$.

The bonds are separated from each other by $120°$ rotations. The momenta $\vec{k}_i = -\vec{K}_i/2$ in (7) are given by half of the Brillouin zone's K points,

$$\vec{K}_{1,3} = \frac{4\pi}{3a_0}\left(\frac{1}{2}, \mp\frac{\sqrt{3}}{2}\right)^\top, \quad \vec{K}_2 = \frac{4\pi}{3a_0}(-1, 0)^\top. \quad (9)$$

Here we use the vector symbol $\vec{}$ to label two-dimensional vectors in space/reciprocal space, and use boldface to denote $SO(3)_{\text{v/s}}$ [or $SO(6)$] vectors.

The vector $\boldsymbol{\Phi}_s = (\Phi_4, \Phi_5, \Phi_6)$ corresponds to the three components the Néel order parameter which determines the spin density in the magnetically ordered phase as

$$S_i = \text{Re}\left[ie^{i\vec{K}_1 \cdot \vec{R}_i}\boldsymbol{\Phi}_s\right]. \quad (10)$$

From now on, we will refer to $\{\Phi_1, \Phi_2, \Phi_3\}$ as "VBS monopoles" and $\{\Phi_4, \Phi_5, \Phi_6\}$ as "spin monopoles" as they can be understood as order parameters for valence bond solid and magnetically ordered phases, respectively.

B. Conformal data and operator product expansions

In general, a conformal field theory is fully specified by its conformal data, containing the operator spectrum as well as the *operator product expansion* (OPE) coefficients. The two lowest-lying sets of primaries in the spectrum are given by the 15 adjoint masses M^{i0}, M^{0i}, and M_{ij} (where $i, j = 1, 2, 3$) as well as monopole operators Φ_a. Their scaling dimensions to first subleading order in $1/N$ are given by $\Delta_M \approx 1.46$ and $\Delta_\Phi \approx 1.02$ [49]. A recent comprehensive conformal bootstrap study [39] finds a striking match (under certain CFT bootstrap

assumptions), placing the monopole operator scaling dimension within the range $\Delta_\Phi \in (1.02, 1.04)$. Assuming $\Delta_\Phi = 1.02$ further yields a fermion adjoint mass scaling dimension of $\Delta_M \in (1.33, 1.66)$. Another recent bootstrap work [38] explicitly accounting for the fate of the UV symmetries of the Dirac spin liquid on the triangular lattice found that stability demands $\Delta_\Phi > 1.046$. Note that Monte Carlo simulations [18] appear to be consistent with the latter study, but have claimed to be ruled out by the aforementioned bootstrap study [39].

Note that the precise numerical values of Δ_Φ and Δ_M are inconsequential to our study below. However, given the recent results referenced above, it appears to be justified to take $\Delta_\Phi < \Delta_M \lesssim 1.5$ when analyzing scaling behavior and (ir)relevance of operators.

Crucially, conformal invariance determines the two-point functions of two primary operators $\mathcal{O}_i$, $\mathcal{O}_j$ up to a global constant, which may be absorbed into the normalization of the operators,

$$\langle \mathcal{O}_i(x)\mathcal{O}_j(y) \rangle = \frac{\delta_{ij}}{|x-y|^{2\Delta_i}}. \quad (11)$$

Further, CFTs admit operator product expansions: two operators approaching each other may be expanded in primaries $\mathcal{O}_k$ as

$$\lim_{x \to y} \mathcal{O}_i(x)\mathcal{O}_j(y) = \lim_{x \to y} \sum_k \frac{C_{ij}^k}{|x-y|^{\Delta_i + \Delta_j - \Delta_k}}\mathcal{O}_k(y), \quad (12)$$

where the C_{ij}^k are the OPE coefficients. Note that, as written, (12) corresponds to an asymptotic form of the operator product expansion as $x \to y$, while in general any OPE for distinct x, y can be shown to be a *convergent* series (extending over primaries and descendants) in a conformal field theory, implying that all correlation functions are fully determined by OPE coefficients and scaling dimensions $\Delta_\mathcal{O}$ [50].

The form of the OPE in (12) is strongly constrained by symmetry considerations: If the two operators $\mathcal{O}_i$ and $\mathcal{O}_j$ transform in some irreducible representation of a symmetry group $\mathcal{G}$, the operators on the right-hand side of (12) can be classified according to the irreducible representations of $\mathcal{G} \times \mathcal{G}$. Using the results by Song *et al.* [34,35], we can hence formulate the operator product expansions for the most relevant monopole and mass operators up to global constants (with $\sim$ denoting the asymptotic character of the expressions)

$$\Phi_a^\dagger(x)\Phi_b(y) \sim \frac{\delta_{ab}}{|x-y|^{2\Delta_\Phi}} + \frac{ic_{\Phi\Phi}^M \mathcal{F}_{\mu\nu}^{ab}}{|x-y|^{2\Delta_\Phi - \Delta_M}}M_{\mu\nu}(y) + \cdots, \quad (13a)$$

$$\Phi_a(x)M_{\mu\nu}(y) \sim \frac{ic_{\Phi M}^\Phi \bar{\mathcal{F}}_{ab}^{\mu\nu}}{|x-y|^{\Delta_M}}\Phi_b(y) + \cdots, \quad (13b)$$

where $\mathcal{F}_{\mu\nu}^{ab}$ is a tensor which maps elements of the 15-dimensional adjoint representation (indexed by μ, ν, with $\mu = \nu = 0$ excluded from any summation) of $SU(4)$ to the rank-2 antisymmetric representation of $SO(6)$ (indexed by a, b), and $\bar{\mathcal{F}}_{ab}^{\mu\nu}$ is the corresponding inverse tensor satisfying $\mathcal{F}_{\mu\nu}^{ab}\bar{\mathcal{F}}_{ab}^{\rho\lambda} = \delta_\mu^\rho\delta_\nu^\lambda$. An explicit construction of $\mathcal{F}$ shows that one can take $\bar{\mathcal{F}}_{ab}^{\mu\nu} = \mathcal{F}_{\mu\nu}^{ab}/2$, and we further note the identity

$\mathcal{F}_{\mu\nu}^{ab}\mathcal{F}_{\mu\nu}^{cd} = \delta^{ac}\delta^{bd} - \delta^{ad}\delta^{bc}$. We give details on the derivation of Eqs. (13a) and (13b) in the Appendix. The real OPE coefficients $c_{\Phi\Phi}^M$ and $c_{\Phi M}^\Phi$ cannot be determined from symmetry considerations, but are constrained by consistency relations encoding the associativity of various four-point functions. Determining numerical values from these sets of constraints is a key element of the conformal bootstrap program [51].

III. CONTINUUM THEORY FOR BILAYER U(1) DIRAC SPIN LIQUIDS

In this section, we derive a continuum model for the bilayer Dirac spin liquids with arbitrary elastic deformations. We start by assuming that the two layers are only slightly deformed with respect to each other, so that the interlayer interaction can be described as the integral of a local Lagrangian density. To make the locality manifest, the following Eulerian coordinates will be used:

$$\vec{x}_l = \vec{R}_l + \vec{u}_l(\vec{x}) + \vec{z}_l, \tag{14}$$

where $l \in \{1, 2\}$ is the layer index, $\vec{R}$ describes the coordinates in the parent triangular lattice, $\vec{u}$ is the smooth deformation satisfying $\partial \vec{u} \ll 1$, $\vec{x}$ is the actual position in the bilayer system, and $\vec{z}$ represents the vertical displacement between the two layers. In the case of a relative rigid twist of angle θ, we have

$$\vec{u}_{1/2} = \pm\frac{\theta}{2}\hat{z} \times \vec{x}. \tag{15}$$

As usual [52], such a rigid relative twist leads to the formation of a *moiré superlattice* with reciprocal lattice vectors $\vec{g}_i^{(m)} = -\theta\hat{z} \times \vec{g}_i$, and the corresponding lattice vectors for the moiré superlattice read as $\vec{a}_1^{(m)} = (0, -1)^\top a_0/\theta$ and $\vec{a}_2^{(m)} = (\sqrt{3}/2, -1/2)^\top a_0/\theta$, which implies that the moiré lattice constant $a_m = a_0/\theta$ is inverse proportional to the twist angle θ.

In the following, we will use the physical symmetries to constrain the possible interlayer interactions. Considering the scaling dimensions review in Sec. II, two types of interactions are most relevant: (i) interlayer tunneling terms of the monopoles, which keep the *total* emergent magnetic flux conserved in the full bilayer system; and (ii) interlayer couplings of the bilinear masses from the two layers. We will discuss these two cases separately in the following two subsections.

A. Interlayer monopole couplings

We have seen in Sec. II A that within each layer, there are three monopoles $\Phi_{l,a}$ with $a \in \{1, 2, 3\}$ that carry the valley degrees of freedom and transforms in the defining representation of $SO(3)_v$, while other three with $a \in \{4, 5, 6\}$ carry the spin degrees of freedom and transform under $SO(3)_s$. Assuming locality, Hermiticity, and spin-rotation invariance, the most general monopole tunneling terms are

$$\mathcal{L}_1 = J_v \sum_{a,b=1,2,3} f_v^{ab}[\vec{u}_1, \vec{u}_2]\,\Phi_{1a}^\dagger(x)\Phi_{2b}(x) + \text{H.c.}$$

$$+ J_s\, f_s[\vec{u}_1, \vec{u}_2] \sum_{a=4,5,6} \Phi_{1a}^\dagger(x)\Phi_{2a}(x) + \text{H.c.} \tag{16}$$

In the above, $\vec{u}_l$ are spatial vectors while x is a spacetime 3-vector with spatial components $\vec{x}$. We will use this labeling method throughout the paper. Since the interlayer physics should not be modified by a simultaneous translation of both layers by a lattice vector, the complex functions f_s and f_v^{ab} can depend only on the relative displacements between the two layers $\vec{u} \equiv \vec{u}_1 - \vec{u}_2$. One can also observe that $f_v^{ab} = f_v^a \delta_{ab}$ should be diagonal because the translation symmetries act as diagonal matrices on the monopoles, which can be observed from the second and third columns of Table I. Only in the case $a = b$, the phase factors coming from the two layers that are gained from translations will cancel each other.

In addition, if the relative displacement field is shifted by a Bravais lattice vector of the parent triangular lattice, the interlayer physics should be invariant. For example, under a shift of the basis vector $\vec{a}_i$ defined in Fig. 2, $\vec{u}_1 \to \vec{u}_1 + \vec{a}_i$, one can compensate by $\vec{R}_l \to \vec{R}_l - \vec{a}_i$ to return to the original labeling of fields at the same locations. Consequently, we should also make the transformations on the monopole operators as in Table I. Considering this transformation in $\mathcal{L}_1$, we find the periodicity conditions $f_v^a[\vec{u} + \vec{a}_1] = e^{i\vec{k}_a \cdot \vec{a}_1} f_v^a[\vec{u}]$ and $f_s[\vec{u} + \vec{a}_1] = e^{i\vec{K}_1 \cdot \vec{a}_1} f_s[\vec{u}]$, which implies that the interlayer tunneling functions can be written in terms of a plane wave multiplied by a (Fourier-expanded) periodic function

$$f_v^a[\vec{u}] = e^{i\vec{k}_a \cdot \vec{u}} \sum_{\vec{Q}} f_{\vec{k}_a + \vec{Q}}\, e^{i\vec{Q} \cdot \vec{u}},$$

$$f_s[\vec{u}] = e^{i\vec{K}_1 \cdot \vec{u}} \sum_{\vec{Q}'} f_{\vec{K}_1 + \vec{Q}'}\, e^{i\vec{Q}' \cdot \vec{u}} \tag{17}$$

with $\vec{Q}$ and $\vec{Q}'$ being the reciprocal lattice vectors and $\vec{k}_a$ the momentum of the monopole operator $\Phi_{la}^\dagger$ as defined in Eqs. (9) and (10).

Next, we impose rotation symmetries. The threefold rotation $C_3 = C_6^2$ leaves invariant the spin monopoles with $a \in \{4, 5, 6\}$, thus relating different Fourier components in $f_s(\vec{u})$. Keeping only the lowest harmonics, we get

$$f_s(\vec{u}) = f_{0,s} \sum_{i=1}^{3} e^{i\vec{K}_i \cdot \vec{u}}, \tag{18}$$

with $\vec{K}_i$ defined in (9) and $f_{0,s}$ an arbitrary complex number for now. A single sixfold rotation C_6 takes $\Phi_{la}^\dagger \to -\Phi_{la}$, giving $f_s(\vec{u})\Phi_{1a}^\dagger(x)\Phi_{2a}(x) \to f_s(\vec{u}')\Phi_{1a}^\dagger(x')\Phi_{2a}^\dagger(x')$, with $\vec{x}' = \hat{R}_{\pi/3}\vec{x}$ and $\vec{u}' = \hat{R}_{\pi/3}^{-1}\vec{u}$. ($\hat{R}_{\pi/3}$ is the counterclockwise 2D rotation matrix of angle $\pi/3$.) For this to be a symmetry, we need $f_s[\vec{u}'] = f_s^*[\vec{u}]$, which, upon plugging in (18), requires a *real* $f_{0,s} = f_{0,s}^*$.

As for the VBS monopoles with $a \in \{1, 2, 3\}$, rotation symmetry relates different f_v^a's: Under C_3, $f_v^a[\vec{u}]\Phi_{1a}^\dagger\Phi_{2a} \to f_v^a[\vec{u}']\Phi_{1,a+2}^\dagger\Phi_{2,a+2}$, with $a + 2$ understood as modulo 3 and $\vec{u}' = \hat{R}_{2\pi/3}^{-1}\vec{u}$. This requires $f_v^a[\vec{u}'] = f_v^{a+2}[\vec{u}]$. Keeping only the lowest harmonics closed under the symmetry, we arrive at

$$f_v^a[\vec{u}] = f_{0,v}\, e^{i\vec{k}_a \cdot \vec{u}}. \tag{19}$$

Similarly, a single C_6 rotation constrains $f_{0,v}$ to be real. The reflection symmetry does not give rise to additional constraints
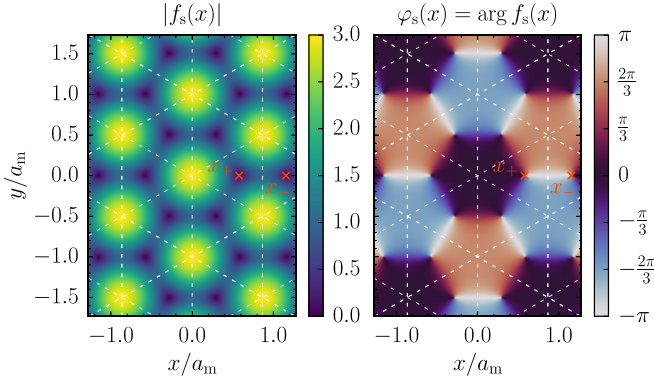
FIG. 3. Left: Spatial dependence of $|f_s(x)|$ (on the moiré lattice scale). At the center of each moiré triangle, the norm of f_s vanishes. Two singular points with positive (x_+) and negative (x_-) are marked by a red cross. Right: The phase $\phi_s = \arg f_s(x)$ which is seen to wind by 2π when encircling a moiré triangle.

on the hopping amplitudes in both spin and VBS cases. From now on, we will absorb the real constants $f_{0,s}$ and $f_{0,v}$ into the coupling constants J_v and J_s in Eq. (16).

In the case of a trivial stacking with $\vec{u} = 0$, the monopole tunneling terms simply reduce to

$$\mathcal{L}_1^{\text{triv}} = J_v \sum_{a=1,2,3} \Phi_{1a}^\dagger \Phi_{2a} + J_s \sum_{a=4,5,6} \Phi_{1a}^\dagger \Phi_{2a} + \text{H.c.} \quad (20)$$

For a rigid twist of angle θ with $\vec{u} = \theta \hat{z} \times \vec{x}$, we can write $e^{i\theta \vec{K}_i \cdot (\hat{z} \times \vec{x})} = e^{i\theta \vec{x} \cdot (\vec{K}_i \times \hat{z})}$ and thus it becomes clear that the function $f_s(\vec{x}) = f_s[\vec{u}(x)]$ is periodic with wave vectors $\vec{q}_a = -\theta \hat{z} \times \vec{K}_a$ which lie on the corners of the lattice's moiré Brillouin zone. In particular, we have

$$\vec{q}_{1,3} = \frac{4\pi\theta}{3a_0} \left(\mp \frac{\sqrt{3}}{2}, -\frac{1}{2} \right)^\top \quad \text{and} \quad \vec{q}_2 = \frac{4\pi\theta}{3a_0}(0,1)^\top. \quad (21)$$

The $\vec{q}_a$ are hence reciprocal lattice vectors for a *magnetic* moiré Brillouin zone (which is a factor 3 smaller than the lattice's moiré Brillouin zone), and the lattice vectors for this magnetic moiré Brillouin zone read as

$$\vec{b}_1^{(\text{m})} = (-\sqrt{3}, 0)^\top \frac{a_0}{\theta} \quad \text{and} \quad \vec{b}_2^{(\text{m})} = \left(-\frac{\sqrt{3}}{2}, \frac{3}{2} \right)^\top \frac{a_0}{\theta}. \quad (22)$$

For a rigid relative twist by angle θ, we thus find that the interlayer monopole tunneling terms read as

$$\mathcal{L}_1 = J_v \sum_{a=1}^3 f_v^a(x) \Phi_{1a}^\dagger \Phi_{2a} + J_s f_s(x) \sum_{a=4}^6 \Phi_{1a}^\dagger \Phi_{2a} + \text{H.c.},$$

$$f_v^a(x) = e^{i\vec{q}_a \cdot \vec{x}/2}, \quad f_s(x) = \sum_{i=1}^3 e^{i\vec{q}_i \cdot \vec{x}}. \quad (23)$$

Note that for convenience of notation, we will define momenta in Euclidean spacetime as $q \equiv (0, \vec{q})$, i.e., all interactions are at zero frequency. We emphasize that, while the functions $f_v^a(x)$ are *pure phases*, the function $f_s(x)$ has a spatially varying magnitude and features zeros at the the centers of the moiré triangles, as shown in Fig. 3.

TABLE II. Transformation laws of the bilinear masses under microscopic symmetries. The second and third columns represent the transformation under translation along $\vec{a}_1/\vec{a}_2$, fourth column the reflection symmetry, fifth column the sixfold rotation symmetry, and last column the time-reversal symmetry. This table is adapted from Table 2 in Ref. [34].

	T_1	T_2	R	C_6	$\mathcal{T}$
M_{00}	+	+	−	+	−
M_{i0}	+	+	+	−	+
M_{01}	−	−	M_{03}	$-M_{02}$	+
M_{02}	+	−	$-M_{02}$	M_{03}	+
M_{03}	−	+	M_{01}	M_{01}	+
M_{i1}	−	−	$-M_{i3}$	M_{i2}	−
M_{i2}	+	−	M_{i2}	$-M_{i3}$	−
M_{i3}	−	+	$-M_{i1}$	$-M_{i1}$	−

B. Interlayer mass couplings

In this part, we use the microscopic symmetries to constrain the possible interlayer coupling terms of bilinear masses $M_{l,\mu\nu} = \bar{\psi}_l \sigma^\mu \tau^\nu \psi_l$, with $\mu, \nu \in \{0, 1, 2, 3\}$. The general Lagrangian is

$$\mathcal{L}_2 = \sum_{\mu,\nu,\rho,\sigma} w_{\mu\nu\rho\sigma}[\vec{u}_1, \vec{u}_2] M_{1,\mu\nu}(\vec{x}) M_{2,\rho\sigma}(\vec{x}). \quad (24)$$

To avoid potential confusion, we will write out all summations explicitly. Spin-rotational invariance requires $w_{\mu\nu\rho\sigma} = \delta_{\mu\rho} w_{\mu\nu\rho\sigma}$ and $w_{1\mu1\nu} = w_{2\mu2\nu} = w_{3\mu3\nu}$, thus splitting the interlayer coupling terms into two sets $\sum_i M_{1,i\mu}(\vec{x}) M_{2,i\nu}(\vec{x})$ and $M_{1,0\mu}(\vec{x}) M_{2,0\nu}(\vec{x})$.

Next, we impose the translation invariance similar as in the previous section: (1) Translation of the bilayer system by an arbitrary constant will not affect the interlayer physics, therefore, $w_{\mu\nu\rho\sigma}[\vec{u}_1, \vec{u}_2] = w_{\mu\nu\rho\sigma}[\vec{u}]$ with $\vec{u} = \vec{u}_1 - \vec{u}_2$. (2) Based on Table II, under translations T_i, all masses are mapped to themselves up to a sign. Therefore, we can only combine the mass terms in the two layers that obtain the same sign under these translations. Combining with the spin-rotation symmetry, we get simply $w_{\mu\nu\rho\sigma} = \delta_{\mu\rho}\delta_{\nu\sigma} w_{\mu\nu\rho\sigma} \equiv w_{\mu\nu}$. (3) We can translate only the first layer by $\vec{a}_1$. We see that, while $M_{\mu\nu}$ with $\nu \in 2\mathbb{Z}$ transform to itself under T_1, the other masses obtain a minus sign. In other words, $w_{\mu\nu}[\vec{u}] = w_{\mu\nu}[\vec{u} + \vec{a}_1]$ when ν is even, and $w_{\mu\nu}[\vec{u}] = w_{\mu\nu}[\vec{u} + 2\vec{a}_1]$ when ν is odd. Further, under T_2, we have $w_{\mu\nu}[\vec{u}] = w_{\mu\nu}[\vec{u} + \vec{a}_2]$ when $\nu \mod 3 = 0$, and $w_{\mu\nu}[\vec{u}] = w_{\mu\nu}[\vec{u} + 2\vec{a}_2]$ when $\nu \mod 3 = 1, 2$. With this information, one can make a Fourier expansion

$$w_{\mu\nu}[\vec{u}] = \sum_{\vec{k}}' e^{i\vec{k}\cdot\vec{u}} w_{\mu\nu}(\vec{k}), \quad (25)$$

where the prime is a reminder that the summation over $\vec{k}$ has different meanings for different μ and ν: In the basis of the reciprocal lattice vectors $\vec{k} = k_1 \vec{b}_1 + k_2 \vec{b}_2$ with $\vec{b}_1 = \frac{2\pi}{a_0}(1, -1/\sqrt{3})$ and $\vec{b}_2 = \frac{2\pi}{a_0}(0, 2/\sqrt{3})$, $k_1 \in \mathbb{Z}$ when ν is even and $k_1 \in (2\mathbb{Z}+1)/2$ when ν is odd; $k_2 \in \mathbb{Z}$ when $\nu \mod 3 = 0$ and $k_2 \in (2\mathbb{Z}+1)/2$ when $\nu \mod 3 = 1, 2$.

We now examine the $w_{\mu i}$ terms under the remaining spatial symmetries. Notice that under C_6, the quartic terms

$w_{\mu 1}[\vec{u}]M_{1,\mu 1}M_{2,\mu 1} \to w_{\mu 1}[\hat{O}_{C_6}^{-1}\vec{u}]M_{1,\mu 2}M_{2,\mu 2}$, and we therefore expect $w_{\mu 1}[\hat{O}_{C_6}^{-1}\vec{u}] = w_{\mu 2}[\vec{u}]$. Here $\hat{O}_{C_6}$ is the rotation matrix which sends $\vec{a}_1 \to \vec{a}_2, \vec{a}_2 \to \vec{a}_2 - \vec{a}_1$. Plugging into the Fourier series (25), we thus arrive at $w_{\mu 1}(\vec{k}\hat{O}_{C_6}^{-1}) = w_{\mu 2}(\vec{k})$. Similarly, one can easily derive that $w_{\mu 2}(\vec{k}\hat{O}_{C_6}^{-1}) = w_{\mu 3}(\vec{k})$ and $w_{\mu 3}(\vec{k}\hat{O}_{C_6}^{-1}) = w_{\mu 1}(\vec{k})$. From the discussions of translational invariance, we know that there is no uniform $\vec{k} = (0, 0)$ component of $w_{\mu i}$. The relations among the Fourier components of $w_{\mu i}$ with lowest momenta are thus

$$
\begin{aligned}
w_{\mu 1}\left(\frac{1}{2}, \frac{1}{2}\right) &= w_{\mu 2}\left(0, \frac{1}{2}\right) = w_{\mu 3}\left(\frac{1}{2}, 0\right) \\
&= w_{\mu 1}\left(-\frac{1}{2}, -\frac{1}{2}\right) = w_{\mu 2}\left(0, -\frac{1}{2}\right) \\
&= w_{\mu 3}\left(-\frac{1}{2}, 0\right).
\end{aligned}
\tag{26}
$$

The reflection symmetry does not add additional constraints.

For the $w_{\mu 0}$ terms, we have seen from spin conservation and translations that only $w_{00}M_{1,00}M_{2,00}$ and $w_{10}\sum_i M_{1,i0}M_{2,i0}$ are allowed. This is also confirmed by the fact that $M_{\mu 0}$ can at most change by a sign under all the remaining symmetry transformations in Table I. Furthermore, similar to the $w_{\mu i}$ case, we can derive $w_{\mu 0}[\vec{u}] = w_{\mu 0}[\hat{O}^{-1}\vec{u}]$ with $\hat{O}$ being the operator for either the C_6 rotation or the M reflection. The Fourier component with the lowest momentum is simply the uniform piece $w_{\mu 0}(0, 0)$.

Combining the discussions above and transforming back to Cartesian coordinates, we arrive at the minimal symmetry-allowed interlayer mass-mass couplings

$$
\begin{aligned}
\mathcal{L}_2 = {}& g_0 M_{1,00}M_{2,00} + g_1 \sum_{i=1}^{3} M_{1,i0}M_{2,i0} \\
&+ g_2 \sum_{i=1}^{3} \xi_i[\vec{u}]M_{1,0i}M_{2,0i} + g_3 \sum_{i=1}^{3}\sum_{j=1}^{3} \xi_j[\vec{u}]M_{1,ij}M_{2,ij},
\end{aligned}
\tag{27}
$$

where the functions $\xi_i[\vec{u}]$ are given by

$$
\xi_i[\vec{u}] = \cos(\vec{M}_i \cdot \vec{u}),
\tag{28}
$$

with $\vec{M}_i$ denoting the $\vec{M}$ points of the hexagonal Brillouin zone with $\vec{M}_1 = (-\pi, -\pi/\sqrt{3})^\top/a_0$, $\vec{M}_2 = (0, 2\pi/\sqrt{3})^\top/a_0$, and $\vec{M}_3 = (\pi, -\pi/\sqrt{3})^\top/a_0$.

The scaling dimension of the scalar mass M_{00} in large N is found to be $\Delta_{M_{00}} = 3.08$ (to first subleading order) and the g_0 term is thus expected to be irrelevant [49]. As discussed earlier, the scaling dimension of the adjoint mass in a single layer is unclear as of now. In large N, the interlayer coupling of the adjoint mass appears to be (weakly) relevant (i.e., $\Delta_M \lesssim 1.5$), but the range given in Ref. [39] also allows for $\Delta_M \gtrsim 1.5$, potentially rendering the interlayer term irrelevant. In the following, we thus primarily focus on the effect of the interlayer monopole tunneling terms, and revisit the effect of mass couplings in the discussion in Sec. VIII.

IV. CONFORMAL PERTURBATION THEORY

We first consider the homogeneously stacked system (absent of any twisting). In this case, the interlayer interaction J is relevant and symmetry allowed, and moreover the only dimensionful scale of the theory. This implies that the bilayer system is unstable in the thermodynamic limit for any infinitesimal J. Dimensional analysis further provides the finite-size scaling law

$$
J \sim L^{-(3-2\Delta_\Phi)},
\tag{29}
$$

where L is the linear size of the system.

To study the *twisted* system's propensity towards an instability induced by the interlayer interaction, we perturbatively compute corrections to the interlayer correlator $\langle \Phi_1^\dagger(x)\Phi_2(y)\rangle$. At long distances $\lim_{|x-y|\to\infty}\langle\Phi_1^\dagger(x)\Phi_2(y)\rangle \neq 0$ can be taken to imply long-range order. Note that in the decoupled theory we have $\langle\Phi_1^\dagger(x)\Phi_2(y)\rangle_{\text{QED}_3} = \langle\Phi_1^\dagger(x)\rangle_{\text{QED}_3}\langle\Phi_2(y)\rangle_{\text{QED}_3} \equiv 0$ by conformal invariance. Upon switching on a finite (but small) J, this interlayer correlator no longer vanishes but will receive finite perturbative corrections which can be organized order by order. If these corrections are small, the physics is still controlled by the fixed point of two decoupled layers of QED$_3$. On the other hand, the breakdown of perturbation theory (i.e., corrections are no longer small) signals an instability to some other phase/fixed point.

For explicit calculations, we find it convenient to employ a path-integral based formulation, where we denote the full action of the interacting bilayer system by $\mathcal{S} = \mathcal{S}_{\text{QED}_3}^{(1)} + \mathcal{S}_{\text{QED}_3}^{(2)} + \mathcal{S}_{12}$, with the associated partition function $\mathcal{Z} = \int \mathcal{D}[\{\mathcal{O}\}]\exp(-\mathcal{S})$, defining a free energy $F = -\log\mathcal{Z}$. We will not attempt a rigorous definition of above path-integral measure $\mathcal{D}[\{\mathcal{O}\}]$ which implies integrating over all operators of the conformal field theory describing the IR fixed point. In the large-N limit, the measure can be rewritten in terms of the gauge fields a_l^μ and fermions ψ_l in the background of monopole configurations (which are also integrated over). Instead, for our purposes it is sufficient to think of the monopole operators as independent fields.

In the following, we focus on the instability due to the interlayer tunneling J_s of the spin monopoles Φ_a with $a = 4, 5, 6$, but an almost identical calculation holds for the VBS monopoles Φ_a with $a = 1, 2, 3$. We expand the Boltzmann weight in the path integral $e^{-\mathcal{S}_{12}} = 1 - \mathcal{S}_{12} + \mathcal{S}_{12}^2/2 + \cdots$ and denote expectation values with respect to the two copies of QED$_3$, corresponding to the two decoupled layers, by $\langle\cdot\rangle_0$. We can then obtain the leading contribution to the interlayer correlator of the spin monopoles at large distances,

$$
\begin{aligned}
\langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\rangle &\approx J\int d^3x\, [f_s^*(x) \\
&\times \langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\Phi_{2,b}^\dagger(x)\Phi_{1,b}(x)\rangle_0] \\
&\stackrel{|d|\to\infty}{\approx} \frac{J}{|d|^{2\Delta_\Phi}}\int d^3x \frac{1}{|x|^{2\Delta_\Phi}}f_s^*(x),
\end{aligned}
\tag{30}
$$

where we have used the leading-order term of the monopole-monopole OPE [which corresponds to the two-point function (11)] in the two layers. Recalling $f_s(x) = \sum_{i=1}^{3} e^{i\vec{q}_i \cdot \vec{x}}$, the integral $\int d^3x\, |x|^{-2\Delta_\Phi}f_s^*(x) \sim |q|^{2\Delta_\Phi - 3}$ converges, where the

wave vector $|q|$ is proportional to the inverse moiré lattice constant $|q| \sim a_m^{-1}$. Hence, at at leading order we have

$$\langle \Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\rangle \overset{|d|\to\infty}{\sim} \frac{J|q|^{-(3-2\Delta_\Phi)}}{|d|^{2\Delta_\Phi}}. \tag{31}$$

Evaluating the next term in the perturbative expansion, quadratic in $\mathcal{S}_{12}$, necessitates evaluating three-point functions of the form $\langle \Phi^\dagger \Phi \Phi \rangle_{\text{QED}_3}$ (and Hermitian conjugates) in each layer. Given that monopole operators are charged under the (emergent) $U(1)_{\text{top}}$ symmetry of the QED_3 fixed-point theory, it becomes clear that any correlation function of an *odd* num-

ber of monopole operators needs to vanish identically. The first (subleading) correction to (31) therefore occurs at third order and involves four-point correlation functions in each layer. Four-point correlation functions can be decomposed by successively applying the OPE. In principle, any choice of order is equivalent as the OPE is convergent. However, as we are truncating the OPE beyond leading order, these different choices become inequivalent. Each OPE channel corresponds to the most divergent contributions from respective regions of configuration space where the operators are "close."

In order to investigate the stability of the system at hand, we should therefore find the OPE channel with the strongest IR divergence. We find[1]

$$\delta\langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\rangle \sim J^3 \int d^3x\, d^3y\, d^3z [f_s^*(x)f_s(y)f_s^*(z)\langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\Phi_{1,b}(x)\Phi_{2,b}^\dagger(x)\Phi_{1,c}^\dagger(y)\Phi_{2,c}(y)\Phi_{1,d}(z)\Phi_{2,d}^\dagger(z)\rangle_0]$$

$$\sim J^3 \int d^3x\, d^3y\, d^3z\, \frac{\sum_{i,j,k} e^{-i(q_i\cdot x - q_j\cdot y + q_k\cdot z)}}{|d-x|^{2\Delta_\Phi}|z|^{2\Delta_\Phi}|y-z|^{2\Delta_\Phi}|x-y|^{2\Delta_\Phi}}$$

$$\overset{|d|\to\infty}{\sim} \frac{J^3|q|^{2\Delta_\Phi-3}}{|d|^{2\Delta_\Phi}} \int d^3y'\, d^3z \frac{\sum_{i,j,k} e^{-i(q_i-q_j)\cdot y'} e^{-i(q_i-q_j+q_k)\cdot z}}{|y'|^{2\Delta_\Phi}|z|^{2\Delta_\Phi}}. \tag{32}$$

In the last step, we have substituted $x' = x - y$ and performed the x' integration and similarly substituted $y' = y - z$. We note that for the summands with $i = j$ the y' integral reduces to $\int d^3y'|y'|^{-2\Delta_\Phi} \sim L^{3-2\Delta_\Phi}$, which has an IR divergence that we regularize with the system size L. On the other hand, the z integration is regular (the potential IR divergence is cut off by the oscillatory exponential for all index combinations i, j, k), and we therefore arrive at

$$\delta\langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\rangle \overset{|d|\to\infty}{\sim} \frac{J^3|q|^{2\cdot(2\Delta_\Phi-3)}L^{3-2\Delta_\Phi}}{|d|^{2\Delta_\Phi}}. \tag{33}$$

As discussed above, if this subleading correction is no longer small compared to the leading-order result (31), perturbation theory breaks down. We can therefore obtain the scaling relation which determines the critical point by asking when the ratio of subleading correction to the leading-order term is of order unity,

$$\frac{\delta\langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\rangle}{\langle\Phi_{1,a}^\dagger(d)\Phi_{2,a}(0)\rangle} \overset{!}{\sim} 1 \quad\Rightarrow\quad J^2|q|^{2\Delta_\Phi-3}L^{3-2\Delta_\Phi} \sim 1. \tag{34}$$

Importantly, we find that this condition for the instability still depends on the system size L, which was required to regularize IR divergence. Rewriting above result as

$$J \sim (a_m L)^{-\left(\frac{3}{2}-\Delta_\Phi\right)} \tag{35}$$

makes clear that in the thermodynamic limit (i.e., $L \to \infty$ with a_m fixed) an infinitesimal J is sufficient to induce an instability, implying that the bilayer system is unstable upon

twisting, as in the trivially stacked case (recall $3/2 > \Delta_\Phi \approx 1.05$ as discussed in Sec. II B).

Crucially, however, comparing with the finite-size scaling law for the homogeneously stacked system in (29), we find that the critical J in (35) scales slower as a function of L, which implies that there is a parametrically large region (as a function of L, compared to the homogeneously stacked case) in which the twisted system remains stable.

Consequently, our perturbative analysis implies that while an instability remains at finite twist angles, we find that for any $\theta \neq 0$ the instability is *softened* in the sense that the interlayer interaction has effectively become less relevant.

V. VARIATIONAL CONFORMAL MEAN-FIELD THEORY

While the breakdown of perturbation theory at the (finite-size) critical point (35) suggests that an instability occurs, the perturbative approach is not capable of describing the nature of the resulting phase. Given that monopole operators transform as order parameters for Néel/VBS order on the triangular lattice, it appears plausible that the interlayer interaction (23) leads to some ordered phase. We therefore employ mean-field theory, which consists in approximating the monopole-antimonopole interaction by monopole operators coupling to a mean field which is determined self-consistently. Physically, the mean fields being finite signals that monopole operators have condensed and thus the onset of VBS or magnetic order.

Commonly, the virtue of mean-field theory is that it replaces the task of solving an interacting problem with the solution of a noninteracting problem (which can be done exactly) and finding appropriate self-consistent parameters. In the model at hand, we emphasize that the mean-field approximation still requires us to solve QED_3 with monopole operators coupled to some classical background field. As

---

[1]Note that there is a more divergent term which corresponds to using the OPE $\Phi_1(d)^\dagger\Phi_1(x) \sim |d-x|^{-2\Delta_\Phi}$ *and* $\Phi_2(x)^\dagger\Phi_2(0) \sim |x|^{-2\Delta_\Phi}$, implying that x is simultaneously close to 0 and d, which contradicts our assumption of $|d| \to \infty$.

mentioned above, even in the absence of such background field, QED$_3$ is not exactly solvable and believed to be described by a strongly interacting fixed point. However, as we show below, the conformal structure of said fixed point places strong constraints on correlation functions and scaling behaviors, which in turn allow us to evaluate certain observables in QED$_3$ as functions of the classical background field in controlled limits.

As defined in the previous section, the system's free energy is given by $F = -\log \mathcal{Z}$ with the partition function $\mathcal{Z} = \int \mathcal{D}[\{\mathcal{O}\}] \exp(-\mathcal{S})$ and the action $\mathcal{S} = \mathcal{S}_{\text{QED}_3}^{(1)} + \mathcal{S}_{\text{QED}_3}^{(2)} + \mathcal{S}_{12}$. Using Jensen's inequality [53], it follows that F obeys the Bogoliubov-Gibbs-Feynman inequality

$$F \leqslant F_{\text{mf}} + \langle \mathcal{S} - \mathcal{S}_{\text{mf}} \rangle_{\text{mf}} \equiv F_{\text{var}}, \tag{36}$$

where $\mathcal{S}_{\text{mf}}[h_1, h_2] = \sum_{l=1,2} \mathcal{S}_{\text{QED}_3}^{(l)} + \mathcal{S}_h^{(l)}[h_l]$ is the mean-field action corresponding to two decoupled DSL with classical fields h_l ("mean fields") coupling to the monopole operators, $\mathcal{S}_h^{(l)}[h_l] = \int d^3x [-h_l^{a*}(x) \Phi_{l,a}(x) + \text{H.c.}]$. F_{mf} is the corresponding mean-field free energy. The h_l are subject to self-consistency equations upon minimizing F_{var}, as shown below.

We use brackets to indicate that the mean-field partition functions and expectation values are (via the action $\mathcal{S}_{\text{mf}}$ in the Boltzmann factor) functionals of the fields $h_l^a(x)$, with $l = 1, 2$

denoting the layer, and $a = 1, \ldots, 6$ a SO(6) index. Defining the single-layer partition function

$$\mathcal{Z}_{\text{mf}}[h_l] = \int \mathcal{D}[\{\mathcal{O}\}] e^{-\mathcal{S}_{\text{QED}_3}^{(l)} - \mathcal{S}_h[h_l]}, \tag{37}$$

the partition function associated with $\mathcal{S}_{\text{mf}}$ factorizes and thus the mean-field free energy $F_{\text{mf}} = -\log \mathcal{Z}_{\text{mf}} \equiv -\log \mathcal{Z}_{\text{mf}}[h_1] - \log \mathcal{Z}_{\text{mf}}[h_2]$. We thus rewrite (36) as

$$F \leqslant -\log \mathcal{Z}_{\text{mf}}[h_1] - \log \mathcal{Z}_{\text{mf}}[h_2] + \langle \mathcal{S}_{12} \rangle_{\text{mf}}[h_1, h_2]$$
$$- \sum_{l=1,2} \langle S_h^{(l)} \rangle_{\text{mf}}[h_l]. \tag{38}$$

Note that the expectation values of the interlayer interactions factorize due to the linearity of $\mathcal{S}_{\text{var}}$ such that $\langle \Phi_{1,a}^\dagger \Phi_{2,b} \rangle_{\text{mf}}[h_1, h_2] = \langle \Phi_{1,a}^\dagger \rangle_{\text{mf}}[h_1] \langle \Phi_{2,b} \rangle_{\text{mf}}[h_2]$.

Next, we seek to minimize the right-hand side of (38) with respect to the functions h_l. To this end, we first note that $\delta/\delta h_l^{a*}(x) \log \mathcal{Z}_0[h_{l'}] = \langle \Phi_{l,a}(x) \rangle_{\text{mf}} \delta_{l,l'}$, which embodies that $\langle \Phi \rangle$ and h^* are conjugate variables by construction, and further

$$\frac{\delta}{\delta h_{l'}^{b*}(y)} \langle \Phi_{l,a}^\dagger(x) \rangle_{\text{mf}} = \langle \Phi_{l,a}^\dagger(x) \Phi_{l,b}(y) \rangle_{\text{mf}} \delta_{l,l'}, \tag{39}$$

and similarly for $\langle \Phi_{l,a}(x) \rangle_{\text{mf}}$, which yields the anomalous monopole correlation function. After some manipulations, the saddle-point condition $\delta F_{\text{var}}/\delta h_1^{b*}(y) = 0$ is rewritten as

$$0 = \int d^3x \sum_{a=1}^{6} [(J_a f_a(x) \langle \Phi_{2,a}(x) \rangle_{\text{mf}} + h_1^a(x)) \langle \Phi_{1,a}^\dagger(x) \Phi_{1,b}(y) \rangle_{\text{mf}} + (J_a f_a^*(x) \langle \Phi_{2,a}(x) \rangle_{\text{mf}} + h_1^{a*}(x)) \langle \Phi_{1,a}(x) \Phi_{1,b}(y) \rangle_{\text{mf}}], \tag{40}$$

for all x, and $b = 1, \ldots, 6$. A similar equation is obtained from $\delta F_{\text{var}}/\delta h_2^{b*}(y) = 0$. Here, we have employed $J_{1,2,3} \equiv J_{\text{v}}$ and $J_{4,5,6} \equiv J_{\text{s}}$, as well as $f_{1,2,3} = f_{\text{v}}^{1,2,3}$ and $f_{4,5,6} \equiv f_{\text{s}}$. While (40) is hard to solve directly, a *sufficient* condition for the integral to vanish is given by

$$h_1^a(x) = -J_{\text{v}} f_{\text{v}}^a(x) \langle \Phi_{2,a} \rangle_{\text{mf}}[h_2], \tag{41a}$$

$$h_2^a(x) = -J_{\text{v}} f_{\text{v}}^{a*}(x) \langle \Phi_{1,a} \rangle_{\text{mf}}[h_1] \tag{41b}$$

for $a = 1, 2, 3$ *and*

$$h_1^a(x) = -J_{\text{s}} f_{\text{s}}(x) \langle \Phi_{2,a} \rangle_{\text{mf}}[h_2], \tag{42a}$$

$$h_2^a(x) = -J_{\text{s}} f_{\text{s}}^*(x) \langle \Phi_{1,a} \rangle_{\text{mf}}[h_1] \tag{42b}$$

for $a = 4, 5, 6$.

The expectation values $\langle \cdot \rangle_{\text{mf}}$ implicitly depend on h_1 and h_2 (note that the dependence on h_2 drops out when the expectation value is taken of operators in layer 1 only, and vice versa). Hence, Eqs. (41) and (42) constitute a set of self-consistency equations for the mean fields h_1 and h_2. We note that these become a *necessary condition* for (40) to vanish *iff* the normal and anomalous monopole correlation functions are translationally invariant, $\langle \Phi_{1,a}^{(\dagger)}(x) \Phi_{1,b}(y) \rangle \equiv \langle \Phi_{1,a}^{(\dagger)}(x-y) \Phi_{1,b}(0) \rangle$. Then (40) can be rewritten as a convolution which vanishes if either kernel or argument are zero.

VI. WEAK-COUPLING SOLUTION OF MEAN-FIELD THEORY

We consider a scaling transformation to a length scale set by the moiré lattice scale $a_{\text{m}} \sim 1/|q| \sim a_0/\theta$. Performing such transformation explicitly in (23), we find that the interlayer interaction is multiplied by a dimensionless parameter $J/|q|^{3-2\Delta_\Phi}$. This parameter being small corresponds to interlayer couplings which are small compared to (fast) modulations of the interlayer tunneling amplitude due to moiré modulations $J \ll a_{\text{m}}^{2\Delta_\Phi - 3}$. Using (41) and (42), this further implies that the mean fields can be taken to be small as $|h_{1,2}| \sim J q^{\Delta_\Phi} \ll a_{\text{m}}^{\Delta_\Phi - 3}$. In this limit, we can evaluate expectation values $\langle \cdot \rangle_{\text{mf}}[h_l]$ perturbatively order by order in the background field h_l, allowing us to write and solve *linearized* self-consistency equations.

A. Linearized self-consistency equations

We start by evaluating the expectation value $\langle \Phi_{la} \rangle_{\text{mf}}$. Expanding $e^{-\mathcal{S}_{\text{mf}}} = e^{-\mathcal{S}_{\text{QED}_3}}(1 - \mathcal{S}_h + \mathcal{S}_h^2/2 + \cdots)$, we have

$$\langle \Phi_{la}(x) \rangle_{\text{mf}} = \mathcal{Z}_{\text{mf}}^{-1} \int \mathcal{D}[\{\mathcal{O}\}] \Phi_{la}(x) e^{-\mathcal{S}_{\text{mf}}}$$
$$= \int d^3y \frac{h_l^a(y)}{|x-y|^{2\Delta_\Phi}} + O(h^3), \tag{43}$$

where we have used the OPE of the monopoles (12) at leading order. Note that this result is equivalently obtained by integrating out all QED$_3$ degrees of freedom perturbatively to obtain $F_{\mathrm{mf}}[\boldsymbol{h}_1, \boldsymbol{h}_2]$ at quadratic order in $\boldsymbol{h}_1, \boldsymbol{h}_2$, and then using $\langle \Phi_{la} \rangle_{\mathrm{mf}} = -\delta F_{\mathrm{mf}} / \delta h_l^{a*}$.

Plugging this into the mean-field equations (41) and (42), we arrive at the relationship between the effective fields in the two layers

$$h_1^a(x) = -J_a f_a(x) \int d^3 y \, \frac{h_2^a(y)}{|x-y|^{2\Delta_\Phi}},$$
$$h_2^a(x) = -J_a f_a^*(x) \int d^3 y \, \frac{h_1^a(y)}{|x-y|^{2\Delta_\Phi}}, \tag{44}$$

where again $J_a = J_{\mathrm{v}}$, $f_a = f_{\mathrm{v}}^a$ for $a \in \{1, 2, 3\}$ and $J_a = J_{\mathrm{s}}$, $f_a = f_{\mathrm{s}}$ for $a \in \{4, 5, 6\}$. There is no implicit summation over a. Eliminating h_2^a from above, we arrive at

$$h_1^a(x) = J_a^2 f_a(x) \int d^3 z \, h_1^a(z) \int d^3 y \, f_a^*(y)$$
$$\times \frac{1}{|y-z|^{2\Delta_\Phi}} \frac{1}{|x-y|^{2\Delta_\Phi}}. \tag{45}$$

There is also a similar equation for $h_2^a(x)$. In the next two subsections, we will discuss the cases of spin and VBS monopoles, separately. From (44) and (45) it is clear that the perturbative approach employed here does not fix the magnitude of the mean fields h_1^a and h_2^a. Rather, they determine for which critical $J_{\mathrm{s}}, J_{\mathrm{v}}$ the self-consistency equations admit nontrivial solutions with $h_1^a, h_2^a \neq 0$ and the symmetry of the mean fields.

B. Spin monopoles

1. Solution of self-consistency equations

Given the periodicity of the moiré pattern, we Fourier expand the mean fields as $\boldsymbol{h}_1(x) = \sum_Q \tilde{\boldsymbol{h}}_1(Q) e^{iQ \cdot x}$ on both sides of (45), with some to-be-determined spacetime momenta Q. Note that because of instantaneous nature of the interaction [i.e., $q_i = (0, \vec{q}_i)^\top$ in $f_{\mathrm{s}}(x)$ and $f_{\mathrm{v}}(x)$ as given in (23)], we can immediately write $q_i \cdot x = \vec{q}_i \cdot \vec{x}$. We further expect Q to be some linear combination of the q_i and thus $\tilde{\boldsymbol{h}}_l(Q) \equiv$

$\tilde{\boldsymbol{h}}_l(0, \vec{Q}) \equiv \tilde{\boldsymbol{h}}_l(\vec{Q})$ to denote the Fourier coefficients of $\boldsymbol{h}_l(x)$. We hence find

$$\sum_{\vec{Q}'} \tilde{\boldsymbol{h}}_1(\vec{Q}') e^{i\vec{Q}' \cdot \vec{x}} = J_{\mathrm{s}}^2 \, f_{\mathrm{s}}(\vec{x}) \sum_i \int d^3 y \, \frac{e^{-i\vec{q}_i \cdot \vec{y}}}{|y-x|^{2\Delta_\Phi}}$$
$$\times \sum_{\vec{Q}} \tilde{\boldsymbol{h}}_1(\vec{Q}) \int d^3 z \, \frac{e^{i\vec{Q} \cdot \vec{z}}}{|z-y|^{2\Delta_\Phi}}. \tag{46}$$

Upon changing of variables $y - x \to y$ in the first integral and $z - y \to z$ in the second integral, the equation simplifies to

$$\sum_{\vec{Q}'} \tilde{\boldsymbol{h}}_1(\vec{Q}') e^{i\vec{Q}' \cdot \vec{x}} = J_{\mathrm{s}}^2 \, f_{\mathrm{s}}(\vec{x}) \sum_{\vec{Q}} \tilde{\boldsymbol{h}}_1(\vec{Q}) \sum_i e^{i(\vec{Q} - \vec{q}_i) \cdot \vec{x}}$$
$$\times \int d^3 y \, \frac{e^{i(\vec{Q} - \vec{q}_i) \cdot \vec{y}}}{|y|^{2\Delta_\Phi}} \int d^3 z \, \frac{e^{i\vec{Q} \cdot \vec{z}}}{|z|^{2\Delta_\Phi}}. \tag{47}$$

Integrals of the form $\int d^3 x \, e^{i\vec{k} \cdot \vec{x}} / |x|^{2\Delta_\Phi}$ diverge when $\vec{k} = 0$. In the equation above, only one of the two integrals can possibly diverge for any given wave vector $\vec{Q}$, corresponding to cases (i) $\vec{Q} = \vec{q}_i$ or (ii) $\vec{Q} = 0$, respectively. When Δ_Φ is smaller than $3/2$, i.e., the monopole tunneling term is a relevant perturbation to the $J_{\mathrm{s}} = J_{\mathrm{v}} = 0$ fixed point, the divergence is in the infrared (IR) limit. As discussed in Sec. II B, large-N calculations and a recent conformal bootstrap study points $\Delta_\Phi \simeq 1.02 < 3/2$, such that the interlayer term can be assumed to be strongly relevant, and we have indeed an IR divergence at hand.

To regulate above divergence, we introduce an IR cutoff $L > |x - y|$ to bound the maximal separation of two monopole operators (at coordinates x and y) from above. We emphasize this is a cutoff for the Euclidean spacetime integrals. The length scale set by the cutoff L can therefore be interpreted the system's linear size (i.e., its spatial extent) and simultaneously as an inverse temperature $\beta \sim 1/L$ (for the imaginary-time direction). The limit $L \to \infty$, which is ultimately of interest to us, thus corresponds to the zero-temperature thermodynamic limit of an infinitely large system.

Having introduced above cutoff scheme, the two singular contributions are

$$\sum_{\vec{Q}'} \tilde{\boldsymbol{h}}_1(\vec{Q}') e^{i\vec{Q}' \cdot \vec{x}} = J_{\mathrm{s}}^2 f_{\mathrm{s}}(\vec{x}) \sum_i \left\{ e^{-i\vec{q}_i \cdot \vec{x}} \tilde{\boldsymbol{h}}_1(0) \int d^3 y \, \frac{e^{-i\vec{q}_i \cdot \vec{y}}}{|y|^{2\Delta_\Phi}} \int d^3 z \, \frac{1}{|z|^{2\Delta_\Phi}} + \tilde{\boldsymbol{h}}_1(\vec{q}_i) \int d^3 y \, \frac{1}{|y|^{2\Delta_\Phi}} \int d^3 z \, \frac{e^{i\vec{q}_i \cdot \vec{z}}}{|z|^{2\Delta_\Phi}} + \cdots \right\}$$

$$= c J_{\mathrm{s}}^2 \left| \frac{q}{L} \right|^{2\Delta_\Phi - 3} \sum_{i,j} [\tilde{\boldsymbol{h}}_1(0) e^{i(\vec{q}_j - \vec{q}_i) \cdot x} + \tilde{\boldsymbol{h}}_1(\vec{q}_i) e^{i\vec{q}_j \cdot x}] + \cdots, \tag{48}$$

with $c = 16\pi^2 \Gamma(2 - 2\Delta_\Phi) \sin(\Delta_\Phi \pi) / (3 - 2\Delta_\Phi)$ a constant, $|q| = 4\pi\theta / 3a_0$ from the discussions in Sec. III A, and $\cdots$ represents the contributions from other Fourier components that are regular. Comparing the spatial dependence on the two sides, $\vec{Q}'$ should belong to the set $\{\vec{q}_i - \vec{q}_j, \vec{q}_i\}$ for arbitrary i, j. In the following, we are interested in the long-wavelength modulations of the mean fields and hence drop all oscillatory

terms with wave vectors outside the first moiré Brillouin zone [note that $(\vec{q}_i - \vec{q}_j) \in$ 1st MBZ iff $i = j$].

Since the equation holds for arbitrary x, we can equate the Fourier components directly. For both $\tilde{\boldsymbol{h}}_1(\vec{Q} = 0)$ and $\tilde{\boldsymbol{h}}_1(\vec{Q} = \vec{q}_i)$, this leads to

$$\tilde{\boldsymbol{h}}_1(\vec{Q}) = 3c J_{\mathrm{s}}^2 \tilde{\boldsymbol{h}}_1(\vec{Q}) \left| \frac{q}{L} \right|^{2\Delta_\Phi - 3}. \tag{49}$$

Note that $\tilde{\boldsymbol{h}}_1(\vec{Q}) = 0$ trivially satisfies this equation. A nontrivial solution exists for $3cJ_s^2|q/L|^{2\Delta_\Phi - 3} \equiv 1$, such that (49) becomes a constraint for J_s leading to the critical interlayer tunneling strength with scaling

$$J_s \sim (a_m L)^{\Delta_\Phi - 3/2}, \tag{50}$$

with $a_m \sim a_0/\theta \ll L$ the moiré lattice constant. This is precisely the same critical scaling for the twisted system as obtained via perturbation theory in the interlayer coupling [*without* any additional (mean-field) approximation] in Eq. (35). The fact that our mean-field treatment recovers the correct critical scaling for the instability is a nontrivial crosscheck and provides confidence for the reliability of our mean-field theory.

While the perturbative approach does not fully determine the order parameter, we can extract its symmetry properties. Neglecting higher-wave-vector oscillations, Eqs. (48) and (49) suggest the minimal form of $h_1(x)$ as

$$\boldsymbol{h}_1(\vec{x}) \approx \tilde{\boldsymbol{h}}_1(0) + \tilde{\boldsymbol{h}}_1(q)f_s(\vec{x}), \tag{51}$$

where $q = |q_i|$ is independent of the index i, and we emphasize that the magnitude of the Fourier coefficients $\tilde{h}(0)$, $\tilde{h}(q)$ is arbitrary: Due to the linearity of the self-consistency equations at weak coupling, any linear superposition of the two degenerate solutions with finite $\tilde{h}_1(0) \neq 0$ and $\tilde{h}_1(q) \neq 0$ is also a solution to the self-consistency equations. Similarly, if we eliminate h_1 from Eq. (44), we would get

$$\boldsymbol{h}_2(\vec{x}) \approx \tilde{\boldsymbol{h}}_2(0) + \tilde{\boldsymbol{h}}_2(-q)f_s^*(\vec{x}). \tag{52}$$

These four Fourier components in the two layers are related to each other through (44):

$$\tilde{\boldsymbol{h}}_1(q) = -A\tilde{\boldsymbol{h}}_2(0), \quad \tilde{\boldsymbol{h}}_2(-q) = -A\tilde{\boldsymbol{h}}_1(0),$$
$$A \equiv \frac{4\pi J_s}{3 - 2\Delta_\Phi}L^{3-2\Delta_\Phi} \sim (L/a_m)^{-\Delta_\Phi + 3/2}, \tag{53}$$

where we have used the critical condition (50) in the second line. Therefore, when J_s is positive, generically we have $A \gg 1$, and the spatial modulating pieces in $\boldsymbol{h}_l(x)$ dominate.

The relation between $\tilde{\boldsymbol{h}}_1(0)$ and $\tilde{\boldsymbol{h}}_2(0)$ can be further obtained by going to quartic order in perturbation theory (see Appendix B for details), giving rise to $\tilde{\boldsymbol{h}}_1(0) = r\tilde{\boldsymbol{h}}_2(0)$ with r being a constant phase. This r can then be fixed by noticing that in the trivial stacking limit $|q| \to 0$, when $J_s < 0$ (or $J_s > 0$), we expect the Néel order parameters to align (or antialign). We can therefore determine $r = -\text{sign}(J_s)$.

Combining above results with the mean-field equations (42), we finally arrive at the following spatial dependence of the spin monopole expectation values:

$$\langle \boldsymbol{\Phi}_1(\vec{x}) \rangle_{\text{mf}} \approx -\frac{A}{|J_s|}\tilde{\boldsymbol{h}}(0) - \frac{f_s(x)}{3J_s}\tilde{\boldsymbol{h}}(0),$$
$$\langle \boldsymbol{\Phi}_2(\vec{x}) \rangle_{\text{mf}} \approx \frac{A}{J_s}\tilde{\boldsymbol{h}}(0) + \frac{f_s^*(x)}{3|J_s|}\tilde{\boldsymbol{h}}(0), \tag{54}$$

where, for consistency, we again only keep wave vectors in the first moiré Brillouin zone [this implies that $f_s^*(x)f_s(x) \approx 3$]. We have added back the flavor degrees of freedom and used boldface to label the three-dimensional vectors in the $SO(3)_s$ space. Since $A \gg 1$, we observe that the first terms in the

two equations above dominate, and there are some corrections with modulations of moiré scale.

2. Lifting of global U(3) degeneracy at quartic order

The above weak-coupling analysis determines the ordering wave vectors as well as relative phase factors of the mean fields (or, equivalently, of the order parameter) in the two layers. However, so far our analysis has not determined the form of the Fourier coefficients $\tilde{\boldsymbol{h}}_l(0)$ and $\tilde{\boldsymbol{h}}_l(\pm q)$ which are understood to be three-dimensional vectors. Indeed, the self-consistency equations at quadratic order (45), or equivalently (48), which lie at the heart of our analysis, show that there is a *global* U(3) invariance of rotating $\boldsymbol{h}_{1,2}(x) \mapsto g\boldsymbol{h}_{1,2}(x)$ with $g \in U(3)$ which does not correspond to a *physical* symmetry operation. This redundancy is understood to be an artifact of the quadratic approximation: For a single layer, the mean-field free energy, after perturbatively integrating out QED_3 degrees of freedom, is given by (note that we omit the layer index)

$$F_{\text{mf}}[\boldsymbol{h}] = -\log \mathcal{Z}_{\text{mf}}[0] - \int d^3x\, d^3y \frac{\boldsymbol{h}^*(x) \cdot \boldsymbol{h}(y)}{|x-y|^{2\Delta_\Phi}} + O(h^3), \tag{55}$$

which is readily seen to be invariant under global U(6) transformations which reduce to $U(3) \subset U(6)$ when considering the spin sector with $h^1 = h^2 = h^3 \equiv 0$, where the upper indices are the SO(6) flavor indices. While the form of the quadratic term (55) is unique and mandated by symmetry, an analysis of higher-order SO(6)-invariant tensor structures reveals that at quartic order, this accidental degeneracy is broken. Note that there are no odd-order terms in the perturbative expansion of F_{mf} by $U(1)_{\text{top}}$ and SO(6) symmetry.

Expanding perturbatively in h, the mean-field free energy (for a single layer) is written as $F_{\text{mf}}[\boldsymbol{h}] = -\log \mathcal{Z}_{\text{mf}} - \langle \mathcal{S}_h^2 \rangle_0/2 - (\langle \mathcal{S}_h^4 \rangle_0 - 3\langle \mathcal{S}_h^2 \rangle_0^2)/4! + \cdots$, where $\langle \cdot \rangle_0$ denotes evaluating expectation values in the unperturbed compact QED_3 theory. The quartic term necessitates the evaluation of

$$\langle \mathcal{S}_h^4 \rangle_0 = \int d^3x_1 \ldots d^3x_4 \left\langle \prod_{i=1}^{4}(h^{a_i*}(x_i)\Phi_{a_i}(x_i) + \text{H.c.}) \right\rangle_0, \tag{56}$$

which requires knowledge of the four-point function of monopole operators $\langle \Phi_a^{s_1}(x_1)\Phi_b^{s_2}(x_2)\Phi_c^{s_3}(x_3)\Phi_d^{s_4}(w) \rangle_0$ (here $s_i = \pm$ and $\Phi^+ \equiv \Phi$, $\Phi^- \equiv \Phi^\dagger$ should be understood as monopole and antimonopole, respectively). In contrast to two- and three-point functions, conformal symmetry *does not* fully determine four-point functions. Rather, they can be written in terms of so-called *conformal blocks*, which are functions of conformally invariant parameters and obey certain associativity relations which are used in the conformal bootstrap approach [38,39].

In 2+1 dimensions, however, no closed form for the conformal blocks exist, and we instead follow an approximate strategy. First, we use the fact that the finite-wave-vector components of $\boldsymbol{h}_l(x)$ can be related to the constant components in the respective other layer, $\tilde{\boldsymbol{h}}_l(0)$, such that it is sufficient to consider uniform $\boldsymbol{h}_l(x) = \text{const}$ in (56). By $U(1)_{\text{top}}$ symmetry, only expectation values with zero net topological charge (i.e., containing two monopoles and two antimonopole operators) can be finite. We then posit that a dominant contribution to

the integral in (56) is given by configurations where operator insertions are *close* to each other. Concretely, we consider $\langle \Phi_a^\dagger(x_1)\Phi_b(x_2)\Phi_c^\dagger(x_3)\Phi_d(x_4)\rangle_0$ (the other terms follow by index permutations) and take $x_1 \to x_2$ and $x_3 \to x_4$. The OPEs in (13a) then yield

$$\langle \Phi_a^\dagger(x_1)\Phi_b(x_2)\Phi_c^\dagger(x_3)\Phi_d(x_4)\rangle_0$$
$$\sim \frac{\delta_{ab}\delta_{cd}}{|x_1-x_2|^{2\Delta_\Phi}|x_3-x_4|^{2\Delta_\Phi}}$$
$$- \left(c_{\Phi\Phi}^M\right)^2 \frac{\mathcal{F}_{ab}^{\mu\nu}\mathcal{F}_{cd}^{\rho\lambda}\langle M_{\mu\nu}(x_2)M_{\rho\lambda}(x_4)\rangle_0}{|x_1-x_2|^{2\Delta_\Phi-\Delta_M}|x_3-x_4|^{2\Delta_\Phi-\Delta_M}}. \quad (57)$$

The first term is readily seen to yield a contribution of the $(\boldsymbol{h}^*\cdot\boldsymbol{h})^2$ in F_{mf}, similar to the $\langle \mathcal{S}_h^2\rangle_0^2$ term stemming from reexponentiating the expanded $e^{-\mathcal{S}_h}$.

For the second term, we note that using the two-point function $\langle M_{\mu\nu}(x_2)M_{\rho\lambda}(x_4)\rangle_0 = \delta_{\mu\rho}\delta_{\nu\lambda}|x_2-x_4|^{-2\Delta_M}$, we can perform the contraction $\mathcal{F}_{ab}^{\mu\nu}\mathcal{F}_{cd}^{\rho\lambda}\delta_{\mu\rho}\delta_{\nu\lambda} = \delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc}$. Since the remaining integrals have integrands which are strictly positive functions, the second term gives a contribution of the form

$$\langle \mathcal{S}_h^4\rangle_0 \sim -\left(c_{\Phi\Phi}^M\right)^2 C(L)[(\boldsymbol{h}\cdot\boldsymbol{h})(\boldsymbol{h}^*\cdot\boldsymbol{h}^*) - (\boldsymbol{h}^*\cdot\boldsymbol{h})^2], \quad (58)$$

where $C(L) > 0$ is an IR-divergent prefactor which depends on the cutoff length scale L introduced earlier. Importantly, this implies that the fourth-order contribution in F_{mf} for $\boldsymbol{h}(\vec{x}) \equiv \boldsymbol{h} = $ const can be written to be of the form

$$F_{\mathrm{mf}}[h] \sim \cdots + |h|^4(D_1 + D_2(\hat{\boldsymbol{h}}^*\cdot\hat{\boldsymbol{h}}^*)(\hat{\boldsymbol{h}}\cdot\hat{\boldsymbol{h}})) + \cdots, \quad (59)$$

where (58) implies that the prefactor $D_2 > 0$ (note that the sign of D_1 is left undetermined).

Importantly, the overall sign of the $|\hat{\boldsymbol{h}}\cdot\hat{\boldsymbol{h}}|^2$ term being positive implies that the free energy is minimized for configurations where the static components of the mean field satisfy $\hat{\boldsymbol{h}}\cdot\hat{\boldsymbol{h}} = 0$, i.e., they can be written as $\hat{\boldsymbol{h}} = (\hat{\boldsymbol{u}} + i\hat{\boldsymbol{v}})/\sqrt{2}$, with the two orthonormal vectors $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{v}}$ such that $\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{u}} = \hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{v}} = 1$ and $\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{v}} = 0$. Considering that $\hat{\boldsymbol{h}}$ ultimately determines the vectorial nature of the monopole expectation value $\langle\boldsymbol{\Phi}\rangle$ in the ordered phase (which can be identified with the Néel order parameter), we hence conclude that noncollinear intralayer spin order is energetically preferred [the generic form $\hat{\boldsymbol{h}} = \hat{\boldsymbol{u}} + i\hat{\boldsymbol{v}}$ is seen to give rise to spin spiral ordering with basis vectors $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{v}}$ using Eq. (10)].

C. VBS monopoles

The analyses for the VBS monopoles are in parallel. We again perform a Fourier expansion on both sides of (45), yielding

$$\sum_{\vec{Q}'} h_1^a(\vec{Q}')e^{i\vec{Q}'\cdot\vec{x}} = J_{\mathrm{v}}^2 f_{\mathrm{v}}^a(\vec{x}) \int d^3y \frac{e^{-i\vec{q}_a\cdot\vec{y}/2}}{|y-x|^{2\Delta_\Phi}}$$
$$\times \sum_{\vec{Q}} h_1^a(\vec{Q}) \int d^3z \frac{e^{i\vec{Q}\cdot\vec{z}}}{|z-y|^{2\Delta_\Phi}}. \quad (60)$$

Taking into account the two possible singular contributions to the integrals at $\vec{Q} = \vec{q}_a/2$ and $\vec{Q} = 0$, respectively, we

get

$$\sum_{\vec{Q}'} h_1^a(\vec{Q}')e^{i\vec{Q}'\cdot\vec{x}}$$
$$= cJ_{\mathrm{v}}^2 \left|\frac{q}{2L}\right|^{2\Delta_\Phi-3}\left[h_1^a(0) + h_1^a\left(\frac{\vec{q}_a}{2}\right)e^{i\vec{q}_a\cdot\vec{x}/2}\right] + \cdots. \quad (61)$$

For a nontrivial solution to exist, i.e., when $h_1^a(\vec{x}) \neq 0$, we must have $cJ_{\mathrm{v}}^2|q/2L|^{2\Delta_\Phi-3} = 1$. It immediately follows that J_{v} exhibits the same scaling

$$J_{\mathrm{v}} \sim (a_{\mathrm{m}}L)^{\Delta_\Phi-3/2} \quad (62)$$

as J_{s} in (50). Back to real space, the minimal description of the mean field is thus $h_1^a(\vec{x}) = \tilde{h}_1^a(0) + \tilde{h}_1^a(\frac{q}{2})e^{i\vec{q}_a\cdot\vec{x}/2}$ where $q = |q_a|$ is again independent of a. Similarly, the mean field for the other layer can be derived using (44):

$$h_2^a(\vec{x}) \approx \tilde{h}_2^a(0) + \tilde{h}_2\left(-\frac{q}{2}\right)e^{-i\vec{q}_a\cdot\vec{x}/2}. \quad (63)$$

These four Fourier components in the two layers are related to each other through (44):

$$\tilde{h}_1^a\left(\frac{q}{2}\right) = -A\tilde{h}_2^a(0), \quad \tilde{h}_2^a\left(-\frac{q}{2}\right) = -A\tilde{h}_1^a(0), \quad (64)$$

$A \gg 1$ is the same as defined in (53). Combining with the mean-field equations (42), we arrive at

$$\langle\Phi_1^a(x)\rangle_{\mathrm{mf}} = \left(A\tilde{h}_1^a(0) - \tilde{h}_2^a(0)e^{i\vec{q}_a\cdot\vec{x}/2}\right)/J_{\mathrm{s}},$$
$$\langle\Phi_2^a(x)\rangle_{\mathrm{mf}} = \left(A\tilde{h}_2^a(0) - \tilde{h}_1^a(0)e^{-i\vec{q}_a\cdot\vec{x}/2}\right)/J_{\mathrm{s}}. \quad (65)$$

Similar to the spin monopole case, we observe that the first terms in the two equations above dominate, and there are some corrections with modulations of moiré scale. The relation between $\tilde{h}_1^a(0)$ and $\tilde{h}_2^a(0)$ can also be obtained by going to quartic order perturbation, giving rise to $\tilde{h}_1^a(0) = s\tilde{h}_2^a(0)$ with s a constant phase factor. More details can be found in Appendix B.

We would like to comment that, within a microscopic theory, terms which gives rise to a coupling of the VBS order parameters involve four spin operators. In contrast, the spin monopole tunneling can be generated from an interlayer spin-spin interaction. In any setting where interlayer interactions are weak, we hence expect the interlayer tunneling of the spin monopoles to be dominant, and therefore instabilities with VBS order to be less likely.

VII. MEAN-FIELD THEORY AT STRONG COUPLING: LOCAL DENSITY APPROXIMATION

In the limit of strong interlayer couplings $J_a \gg |\nabla f_a/f_a|^{2\Delta_\Phi-3} \sim (1/a_{\mathrm{m}})^{3-2\Delta_\Phi}$ (recall that we take $2\Delta_\Phi < 3$ which amounts to assuming that interlayer monopole tunneling term is relevant), the spatial variations of the interlayer coupling due to twisting are on a much larger scale compared to the characteristic length scale set by the coupling strength J_α.

We note that the use of IR conformal field theory in the strong-coupling limit is justified: this low-energy theory is expected to apply at energy scales $\Lambda \ll 1/a$ much smaller

than the inverse lattice constant of the parent triangular lattice. Recall that the moiré lattice constant $a_m \gg a$ and thus $(1/a_m)^{3-2\Delta_\Phi} \ll (1/a)^{3-2\Delta_\Phi}$. We conclude that these inequalities show that there exists a parameter regime such that the strong-coupling limit holds, but the typical energy scale $\Lambda \sim J_a^{\frac{1}{3-2\Delta}}$ set by the interlayer coupling is small compared to the UV cutoff determined by the inverse parent lattice scale $1/a$, allowing us to use the low-energy conformal field theory to study the strong-coupling limit. We also note that there is also *another* parameter regime where $J_a \gg (1/a)^{3-2\Delta_\Phi}$, implying that J_a is the dominant scale in the problem and larger than the UV cutoff scale, necessitating the use of microscopic lattice models to study strong interlayer interactions. However, this scenario seems less relevant to van der Waals heterostructures, where the interlayer interactions are generically weaker than intralayer interactions, both being small compared to the UV cutoff provided by the inverse lattice constant (we emphasize that the interlayer interactions being "weak" compared to intralayer interactions does not contradict the *strong-coupling limit* defined earlier, where the interlayer interaction scale is large compared to moiré lattice modulations).

In this strong-coupling limit, we may perform a *local approximation* where we first obtain monopole expectation values as a function of the background fields from a *local* scaling form of the compact QED$_3$ free energy in the presence of (large) background fields. Within this local approximation, the self-consistency equations for distinct points in spacetime decouple and can be solved straightforwardly.

To this end, we first note from $\mathcal{S}_h[\boldsymbol{h}_l]$ that the mean fields $\boldsymbol{h}_l$ have the scaling dimension $[h_l] = 3 - \Delta_\Phi$, such that a symmetry-allowed ansatz for the mean-field free energy $F_{\rm mf}^{\rm loc}[\boldsymbol{h}_l]$ (for each layer l) in the local approximation reads as

$$F_{\rm mf}^{\rm loc}[\boldsymbol{h}_l] = -\int d^3x \, \frac{2c_{\rm loc}}{\chi} |\boldsymbol{h}_l(x)|^\chi, \qquad (66)$$

where $c_{\rm loc}$ is some real constant and $\chi = 3/(3 - \Delta_\Phi)$ from scaling arguments, with the form of the prefactor chosen for later convenience.

Using that $\langle \Phi_{la} \rangle_{\rm mf} = -\delta F_{\rm mf}/\delta h_l^{a*} = c_{\rm loc} \hat{h}_l^a |\boldsymbol{h}_l|^{\chi-1}$ (with $\hat{h}_l^a$ denoting the a component of the complex unit vector $\hat{\boldsymbol{h}}_l = \boldsymbol{h}_l/|\boldsymbol{h}_l|$), the self-consistency equations (42) can be written as

$$h_1^a(x) = -J_a f_a(x) c_{\rm loc} \hat{h}_2^a(x) |\boldsymbol{h}_2(x)|^{\chi-1}, \qquad (67a)$$

$$h_2^a(x) = -J_a f_a^*(x) c_{\rm loc} \hat{h}_1^a(x) |\boldsymbol{h}_1(x)|^{\chi-1}, \qquad (67b)$$

with no implicit summation over a, and recall that $J_{1,2,3} \equiv J_{\rm v}$, $J_{4,5,6} \equiv J_{\rm s}$, and $f_{1,2,3}(x) \equiv f_{\rm v}^{1,2,3}(x)$ and $f_{4,5,6} \equiv f_{\rm s}$.

A. Spin monopoles

We first focus on order in the spin sector (i.e., $J_{\rm s} \neq 0$, $J_{\rm v} = 0$) such that $\boldsymbol{h}$ in each layer can be replaced by a three-dimensional complex vector $(h^4, h^5, h^6)^\top$. First focusing on magnitudes $|\boldsymbol{h}_l|$ and substituting (67) into each other yields

$$|\boldsymbol{h}_l(x)| = |J_{\rm s} c_{\rm loc} f(x)|^{\frac{3-\Delta_\Phi}{3-2\Delta_\Phi}} \qquad (68)$$

for $l = 1, 2$. Using this in Eq. (67), we find that the *relative* phase factor of the two unit vectors $\hat{\boldsymbol{h}}_l$ is fixed by

$$\hat{\boldsymbol{h}}_1 = \exp\left[i(\pi + \arg J_{\rm s} + \varphi_{\rm s}(x) + \arg c_{\rm loc})\right]\hat{\boldsymbol{h}}_2, \qquad (69)$$

where we have introduced $\varphi_{\rm s,v}(x) \equiv \arg f_{\rm s,v}(x)$. Note that, considering the case of a homogeneous bilayer [i.e., $f_{\rm s}(x) \equiv 1$] with a ferromagnetic interlayer coupling $J_{\rm s} < 0$, we expect the two Néel vectors (and so the mean fields $\boldsymbol{h}$) in the two layers to be aligned, which allows us to *a posteriori* fix the sign of the constant $c_{\rm loc} > 0$. Above result is also readily understood when considering the expectation value of the interlayer tunneling term and rewriting using (42),

$$\langle \mathcal{L}_1 \rangle_{\rm mf} \sim J_{\rm s} f(x) \langle \boldsymbol{\Phi}_1^\dagger \rangle_{\rm mf} \cdot \langle \boldsymbol{\Phi}_2 \rangle_{\rm mf}$$

$$\sim \frac{1}{J_{\rm s} f(x)} |h(x)|^2 \hat{\boldsymbol{h}}_1 \cdot \hat{\boldsymbol{h}}_2^* + \text{H.c.}, \qquad (70)$$

such that all phase fluctuations cancel out upon substituting (69).

While we have thus fixed the magnitude of the mean fields (and thus of the magnetic order parameters, identified with the monopole operators) and their relative phase factor, we note that the self-consistency equations in the local approximation have a *local* U(3) redundancy: Given a configuration $\{\boldsymbol{h}_1(x), \boldsymbol{h}_2(x)\}$ which satisfies Eqs. (67), the configuration $\{G(x)\boldsymbol{h}_1(x), G(x)\boldsymbol{h}_2(x)\}$ with an arbitrary matrix field $G(x) \in$ U(3) is a solution as well. This local redundancy is unphysical as it does not correspond to a symmetry of the system and thus will be lifted by corrections to the purely local approximation made above.

In particular, we posit that there is an intrinsic stiffness to the system which energetically favors order parameter textures with small gradients. Using SO(6) symmetry and the above-derived scaling of $\boldsymbol{h}_l$, a corresponding mean-field stiffness term in the free energy may be written as a sum over the two individual stiffness terms in each layer,

$$F_{\rm mf}^{\nabla^2}[\boldsymbol{h}_1, \boldsymbol{h}_2] = \sum_{l=1,2} \rho \int d^3x |\partial_\mu h_l^{a*} \, \partial^\mu h_l^a|^\gamma, \qquad (71)$$

where $\rho > 0$ is some dimensionless constant and $\gamma = 3/(8 - 2\Delta_\Phi)$ from scaling (note summation over $a = 1, \ldots, 6$ is implied). Clearly, (71) does not support the local U(3) redundancy found earlier, and is minimized *in a single layer* through a uniform $\boldsymbol{h}(x) = $ const; however, choosing *both* $\boldsymbol{h}_1(x)$ and $\boldsymbol{h}_2(x)$ constant is no longer a self-consistent solution as is readily verified using Eqs. (67). We here assume that the stiffness is parametrically small compared to the density term $\rho \ll c_{\rm loc}$, such that the gradient terms can be treated as a perturbation which selects a favorable configuration out of the U(3) locally degenerate manifold, without qualitatively altering the nature of the thus selected configuration. Hence, we plug in the solution to the self-consistency equations, given by Eqs. (68) and (69), to rewrite the gradient term as a functional $F_{\rm mf}^{\nabla^2}[\boldsymbol{h}_1, \boldsymbol{h}_2] \equiv F_{\rm mf}^{\nabla^2}[\hat{\boldsymbol{h}}_2]$ of the to-be-determined unit-vector-valued field $\hat{\boldsymbol{h}}_2(x)$. We find the resulting form difficult to extremize analytically due to the non-analyticity of the gradient term.

Instead, we note that $\gamma = 3/(8 - 2\Delta_\Phi) < 1$ and thus the modulus $|\ldots|^\gamma$ is concave, such that Jensen's inequality

$|X|^\gamma + |Y|^\gamma \leqslant 2^{1-\gamma}|X + Y|^\gamma$ holds. Hence, we resort to extremizing an *upper bound* for $F_{\mathrm{mf}}^{\nabla^2}$,

$$\frac{F_{\mathrm{mf}}^{\nabla^2}}{2^{1-\gamma}\rho} \leqslant \int d^3x [||h(x)|^2 (2\partial_\mu \hat{\boldsymbol{h}}_2^* \cdot \partial^\mu \hat{\boldsymbol{h}}_2 - i(\partial_\mu \varphi_{\mathrm{s}})(\hat{\boldsymbol{h}}_2^* \cdot \partial^\mu \hat{\boldsymbol{h}}_2$$
$$- \mathrm{H.c.}) + \partial_\mu \varphi_{\mathrm{s}} \partial^\mu \varphi_{\mathrm{s}}) + 2\partial_\mu |h| \partial^\mu |h||^\chi$$
$$+ \lambda(x)(\hat{\boldsymbol{h}}_2^* \cdot \hat{\boldsymbol{h}}_2 - 1)], \tag{72}$$

where the short form $|h| \equiv |\boldsymbol{h}_l|$ as in (68), and we use "the center dot" to denote the dot product of $\mathrm{SO(3)_{spin}}$ vectors. Note that $\hat{\boldsymbol{h}}_2^* \cdot \partial_\mu \hat{\boldsymbol{h}}_2 = -\hat{\boldsymbol{h}}_2 \cdot \partial_\mu \hat{\boldsymbol{h}}_2^*$ holds due to normalization, the latter being enforced by introducing a real Lagrange multiplier field $\lambda(x)$. Here we have chosen to write (72) in a manifestly Hermitian form.

Considering the right-hand side of (72), we note that due to the chain rule, it is sufficient to only consider the term in the square brackets when extremizing (by varying with respect to $\hat{\boldsymbol{h}}_2$, $\hat{\boldsymbol{h}}_2^*$ independently). Using a redefined (real) Lagrange multiplier field $\tilde{\lambda}$, one thus obtains the differential equation

$$0 = -2\partial_\mu \partial^\mu \hat{\boldsymbol{h}}_2 - 2i\partial_\mu \varphi_{\mathrm{s}} \partial^\mu \hat{\boldsymbol{h}}_2 - i\partial_\mu \partial^\mu \varphi_{\mathrm{s}} \hat{\boldsymbol{h}}_2 + \tilde{\lambda}\hat{\boldsymbol{h}}_2, \tag{73}$$

and similarly for $\hat{\boldsymbol{h}}_2^*$. Making the ansatz $\hat{\boldsymbol{h}}_2(x) = \hat{\boldsymbol{u}}e^{ig(x)}$ for some constant complex unit vector $\hat{\boldsymbol{u}}$ such that the normalization constraint (retrieved by varying with respect to the Lagrange multiplier) is satisfied, the real function $g(x)$ is determined via

$$0 = -2i\partial_\mu \partial^\mu g + 2\partial_\mu g \partial^\mu g + 2\partial_\mu \varphi_{\mathrm{s}} \partial^\mu g - i\partial_\mu \partial^\mu \varphi_{\mathrm{s}} + \tilde{\lambda}. \tag{74}$$

Adding and subtracting (74) and its complex conjugate, we find the partial differential equations

$$0 = 2\partial_\mu g \partial^\mu g + 2\partial_\mu \varphi_{\mathrm{s}} \partial^\mu g + \tilde{\lambda}, \tag{75a}$$

$$0 = 2\partial_\mu \partial^\mu g + \partial_\mu \partial^\mu \varphi_{\mathrm{s}}. \tag{75b}$$

While the first equation can be solved to determine the Lagrange multiplier $\tilde{\lambda}$, the latter can be used to determine $g(x)$. *One particular solution* consists in choosing $g_{\mathrm{p}}(x) = -\varphi_{\mathrm{s}}(x)/2$ and $\tilde{\lambda} = -\partial_\mu \varphi_s \partial^\mu \varphi_s/2$.[2]

However, we note that $\varphi_{\mathrm{s}}(x) = \arg f_{\mathrm{s}}(x)$ has branch cuts that connect pairs of zeros $x^\pm$ of f_{s} at which $f_{\mathrm{s}}(x^\pm) = 0$ and upon encircling f_{s} has a positive (negative) winding number $\mathrm{Ind}_{f_{\mathrm{s}}}(x^\pm) = \pm 1$. Note that here, we use conventions such that $\arg f_{\mathrm{s}} \in [-\pi, \pi]$. Here, we may take $x^+ = (1/\sqrt{3}, 0)a_{\mathrm{m}}$ and $x^- = (2/\sqrt{3}, 0)a_{\mathrm{m}}$, as marked by a red cross in Fig. 3, and by spatial periodicity of $f_{\mathrm{s}}(x)$, all further zeros of positive (negative) winding number are obtained through translations by lattice vectors of the magnetic Brillouin zone

While the branch cut above leaves $e^{i\varphi_{\mathrm{s}}(x)}$ single valued, the function $g_{\mathrm{p}}(x) = -\varphi_{\mathrm{s}}(x)/2$ features a discontinuity across these branch cuts and hence the $e^{ig_{\mathrm{p}}(x)}$ is multivalued at these branch cuts (i.e., there is a jump $\pi/2 \to -\pi/2$ when

encircling a zero of f_{s} with positive chirality, x^+). Since the phase of $\hat{\boldsymbol{h}}_l(x) = \hat{\boldsymbol{u}}e^{\pm ig(x)}$ determines the phase of $\langle \boldsymbol{\Phi}_l \rangle$, which in turn determines the local orientation of the ordered spins via Eq. (10). This is an *unphysical* discontinuity, and hence the choice of $g_{\mathrm{p}}(x) = -\varphi_{\mathrm{s}}(x)/2$ is not admissible. To remedy this, we recall that above choice corresponds to a *particular* solution, and in fact any $g(x) = g_{\mathrm{h}}(x) + g_{\mathrm{p}}(x)$ is a solution to (75b), where g_{h} is a solution to Laplace's equation $\partial_\mu \partial^\mu g_{\mathrm{h}}(x) = 0$. A boundary condition for g_{h} consists in demanding that it has the same spatial symmetries as $g_{\mathrm{p}}(x)$.

One can make progress by noting that $\arg(x, y) \equiv \arg(x + iy) = \mathrm{Im} \log(x + iy)$ is a solution to Laplace's equation and features a branch cut on the negative half of the x axis (note that other conventions may be chosen), such that $\lim_{y \to \pm 0} \arg(x, y) = \pm\pi$ for $x < 0$. Hence, by superposing two such solutions to define a vortex-pair function

$$G_{\mathrm{vp}}^{(3)}(x, y) \equiv \arg\left(-x + \frac{a_{\mathrm{m}}}{\sqrt{3}}, y\right) - \arg\left(-x + \frac{2a_{\mathrm{m}}}{\sqrt{3}}, y\right), \tag{76}$$

one obtains a branch cut between the points $(1/\sqrt{3}, 0)a_{\mathrm{m}}$ and $(2/\sqrt{3}, 0)a_{\mathrm{m}}$ where $G_{\mathrm{vp}}^{(3)}(x, y)$ winds by -2π (clockwise direction) around the former point and by $+2\pi$ around the latter. One may form similar functions $G_{\mathrm{vp}}^{(1)}(x, y)$ and $G_{\mathrm{vp}}^{(2)}(x, y)$ with branch cuts connecting the singular points $a_{\mathrm{m}}(1/(2\sqrt{3}), 1/2)$ and $a_{\mathrm{m}}(1/\sqrt{3}, 1)$ [$a_{\mathrm{m}}(-1/(2\sqrt{3}), 1/2)$ and $a_{\mathrm{m}}(-1/\sqrt{3}, 1)$, respectively].

We may then take the homogeneous solution as a linear superposition of the vortex-pair functions translated by the lattice vectors of the magnetic moiré Brillouin as given in (22),

$$g_{\mathrm{h}}(\vec{x}) = \frac{1}{2} \sum_{m,n \in \mathbb{Z}} q_{mn}^{(\alpha)} G_{\mathrm{vp}}^{(\alpha)}(\vec{x} + n\vec{b}_1^{(\mathrm{m})} + m\vec{b}_2^{(\mathrm{m})}), \tag{77}$$

where the coefficients $q_{mn}^{(\alpha)} = \pm 1$ and should not be confused with the moiré reciprocal lattice vectors $\vec{q}_i$. The prefactor of $1/2$ is chosen such that at the discontinuity $g_{\mathrm{h}}(\vec{x})$ changes $\pm\pi/2 \to \mp\pi/2$.

Considering the full function $g(x) = g_{\mathrm{h}}(x) + g_{\mathrm{p}}(x)$, one thus finds that by choosing the value of $q_{mn}^{(\alpha)} = \pm 1$, one may remove the half branch cut due to the $-\varphi_{\mathrm{s}}/2$ in one layer (such that the phase of $e^{ig(x)}$ does not wind around a vortex pair) and obtain a full branch cut (associated with $\pm 2\pi$ winding around two paired zeros) in the other layer such that $e^{i\varphi_{\mathrm{s}}(x) + g(x)}$ is single valued and thus physical. Note that reversing the sign of $q_{mn}^{(\alpha)}$ then corresponds to placing the (physical) vortex pair in the opposite layer. Within this approach, each choice of $q_{mn}^{(\alpha)}$ is an allowed solution, i.e., for each pair of zeros of $\arg f_{\mathrm{s}}(x)$ we have the choice of placing a vortex pair in the upper or lower layer. While these are all *local* extrema of the free energy [as they solve Eqs. (75a) and (75b)], the total energy of these configurations may differ.

Considering (71), it is reasonable to expect that the energetically favored configurations consist in placing all vortex pairs in either the top or bottom layer such that the gradient term for the other layer is minimized (i.e., $q_{mn}^{(\alpha)} \equiv +1$ or $q_{mn}^{(\alpha)} \equiv -1$ for all m, n, α). Consequently, we expect the interlayer exchange symmetry to be spontaneously broken in this scenario.

To summarize this section, when the tunnelings of spin monopoles are present, in this strong-coupling, local

---

[2]One may worry that the choice of $p = -1/2$ implicitly depends on using the manifestly Hermitian form of (72), however, writing $\hat{h}_2^* \cdot \partial^\mu \hat{h}_2 - \mathrm{H.c.} \equiv (1 + \alpha)\hat{h}_2^* \cdot \partial^\mu \hat{h}_2 - (1 - \alpha)\hat{h}_2 \cdot \partial^\mu \hat{h}_2^*$ and repeating the calculation for a generic α shows that $p = -1/2$ independent of α.

approximation, we have found a magnetic vortex lattice of tunable moiré scale. The expectation values of the spin monopoles in the two layers differ by a spatially modulating phase $\langle \Phi_2 \rangle_{\mathrm{mf}} / \langle \Phi_1 \rangle_{\mathrm{mf}} = e^{i \arg f_s}$, which winds by $\pm 2\pi$ around the moiré downward and upward triangles, as depicted in Fig. 1.

B. VBS monopoles

In the VBS sector ($J_v \neq 0$), we can proceed similarly to the spin sector. Now, we consider the first three components of the (vectorial) mean fields $\boldsymbol{h}_1$, $\boldsymbol{h}_2$. Recalling that all f_a^a's are different, we take the magnitude of each component of h_l^a in Eqs. (67) and (42). Using that $|f_a(x)| = 1$ we hence find that the magnitude of the mean fields is constant,

$$|h_l^a(x)| = |J_v c_{\mathrm{loc}}|^{\frac{3 - \Delta_\Phi}{3 - 2\Delta_\Phi}} \qquad (78)$$

for $l = 1, 2$ and $a = 1, 2, 3$, and the constant $c_{\mathrm{loc}} > 0$ as argued earlier. Performing a polar decomposition of each mean-field component, $h_l^a = |h_l^a| e^{i \vartheta_l^a}$, we find that the phases in the two layers must satisfy

$$\vartheta_1^a(x) = \pi + \arg J_v + \arg c_{\mathrm{loc}} + \varphi^a(x) + \vartheta_2^a(x), \qquad (79)$$

where $\varphi^a(x) \equiv \arg f_v^a(x) = \vec{k}_a \cdot \vec{x}$. Similar to the case of spin monopoles discussed in the previous section, the self-consistency equations admit a *local* redundancy in the VBS sector. However, due to flavor dependence of the interlayer monopole tunneling function, the space of locally degenerate solutions is smaller compared to spin case: In the VBS sector, the self-consistency equations (and their solutions) remain invariant under three independent U(1) rotations $h_l^a \mapsto e^{i \xi^a} h_l^a$ with some real field $\xi^a(x)$. This unphysical local redundancy is again lifted in first order by a gradient-type term as given in (71). We again consider an upper bound for the gradient term using Jensen's inequality (also taking $h^{4,5,6} = 0$). Inserting $h_l^a = |h_l^a| e^{i \vartheta_l^a}$ the problem reduces to minimizing the right-hand side of

$$\frac{F_{\mathrm{mf}}^{\nabla^2}}{2^{1 - \gamma} \rho} \leqslant \int d^3 x \left| \sum_{a = 1, 2, 3} |h|^2 \left(2 \partial_\mu \vartheta_2^a \partial^\mu \vartheta_2^a + 2 \partial_\mu \varphi^a \partial^\mu \vartheta_2^a \right. \right.$$
$$\left. \left. + \partial_\mu \varphi^a \partial^\mu \varphi^a \right) \right|^\gamma \qquad (80)$$

with respect to ϑ_2^a, and $|h| = |h_l^a(x)|$ given in (78). Note that no Lagrange multiplier is required as the polar decomposition ensures normalization of $h_l^a / |h_l^a|$. Again, by the chain rule, it is sufficient to consider the expression in parentheses. Varying with respect to ϑ_2^a, one obtains

$$0 = -2 \partial_\mu \partial^\mu \varphi_a(x) - 4 \partial_\mu \partial^\mu \vartheta_2^a(x). \qquad (81)$$

Note that since $\partial_\mu \partial^\mu \varphi_a(x) = \partial_\mu \partial^\mu (\vec{k}_a \cdot \vec{x}) \equiv 0$, Eq. (81) reduces to Laplace's equation $\partial^\mu \partial_\mu \vartheta_2^a(x) = 0$. Aiming to find a *global* minimum of the right-hand side in (80), we find that choosing $\vartheta_2^a(x) = -\varphi_a(x)/2$ gives a lower bound than the constant solution $\theta_2^a = \mathrm{const}$.

We stress that, in contrast to the case of spin monopoles discussed in the previous subsection, the phase $\varphi^a(x) \equiv \vec{k}_a \cdot \vec{x}$ does not lead to zeros in $e^{\pm i \varphi^a(x)/2}$, but rather doubles the wavelength of the spatial periodicity of the modulation of the phase of the VBS order parameter.

Note that there is a freedom to pick a global U(1) phase ϕ_0^a per flavor (analogous to the choice of $\hat{u}$ in the spin case). The mean fields can hence be written as

$$h_1^a(x) = |J_v c_{\mathrm{loc}}|^{\frac{3 - \Delta_\Phi}{3 - 2\Delta_\Phi}} e^{i[\pi + \arg J_v + \varphi_a(x)/2 + \phi_0^a]} \qquad (82a)$$

$$h_2^a(x) = |J_v c_{\mathrm{loc}}|^{\frac{3 - \Delta_\Phi}{3 - 2\Delta_\Phi}} e^{i[-\varphi_a(x)/2 + \phi_0^a]}. \qquad (82b)$$

Consequently, the VBS order parameters, obtained from inserting above result in (41), remain constant in magnitude throughout the moiré lattice, but feature oscillating phases with wave vectors $\vec{k}_a = -\vec{K}_a/2$ corresponding to half of the moiré lattice's reciprocal lattice vectors. This further leads to modulating phase difference between the monopole expectation values in the two layers $\langle \Phi_{1a} \rangle_{\mathrm{mf}} / \langle \Phi_{2a} \rangle_{\mathrm{mf}} = \exp\{i(-\pi - \arg J_v + \varphi_a(x))\}$. Namely, on the one hand, in certain regions where the phase difference is small, the singlets in the two layers almost lie on top of each other; on the other hand, where the phase difference is near π, the singlets in the two layers avoid each other.

VIII. DISCUSSION AND CONCLUSION

We conclude the paper with a summary of our results, a discussion pertinent to underlying assumptions of our study, and an outlook.

A. Summary

In this work, we have shown that in bilayer systems of U(1) Dirac spin liquids, interlayer monopole-monopole and mass-mass interactions are present in the low-energy field theory. These terms constitute as *relevant* perturbations to the strongly coupled conformal fixed point described by two copies of QED$_3$. Employing a perturbative calculation informed by the low-energy CFT data, we find that an instability due to interlayer interactions persists in the twisted system, but its relevance is effectively reduced [i.e., $J_c \sim (a_m L)^{(\frac{3}{2} - \Delta_\Phi)}$, compared with $J_c \sim L^{-(3 - 2\Delta_\Phi)}$ in the homogeneous case].

Motivated by the identification of monopole operators with antiferromagnetic Néel and VBS order parameters on the triangular lattice, we have focused on monopole-monopole interactions (likely the most relevant interlayer term) and developed a conformal mean-field theory framework which allows us to study ordered phases described by the proliferation of corresponding monopoles. Solving the mean-field equations at weak coupling, we recover the modified critical scaling for finite twist angles, and find that the induced spin/VBS order are mostly uniform, with small twist-induced modulations. We expect this weak-coupling approach to be justified when the interlayer coupling is weak compared to the moiré modulations due to the twisting (red regime in Fig. 4). In the limit of strong interlayer couplings (or, equivalently, *large* moiré unit cells achieved at small twist angles), $J \gg a_m^{2\Delta_\Phi - 3}$, we employ a local approximation to show that the magnetic order parameter forms a moiré vortex lattice.
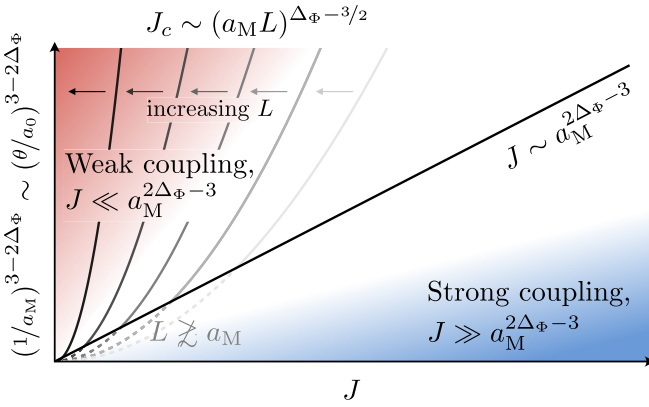
FIG. 4. Illustration of the relation between the weak-coupling (red shading) and the strong-coupling limits (blue shading), and scaling of the critical J_c, as a function of $(1/a_m)^{3-2\Delta_\Phi}$ and $J \equiv J_\alpha$. Note that increasing values on the ordinate corresponds to *decreasing* the size of the moiré unit cell. The dashed lines indicate the parameter regime for which the finite-size weak-coupling instability at $J_c \sim (a_m L)^{\Delta_\Phi - 3/2}$ implies that $L \ll a_m$, which is deemed unphysical and vanishes in the thermodynamic $L \to \infty$ limit. Here, we use a_0 to denote the bare lattice constant of single-layer triangular lattice.

A qualitative overview of the two limits and resulting scaling laws is presented in Fig. 4. We remark that the mean-field solutions in the two regimes cannot be continuously connected to each other: in the weak-coupling case, the solutions (4) exhibit a $\mathbb{Z}_2$ symmetry of $\Phi_1 \leftrightarrow -\Phi_2^\dagger$, while at strong coupling, this symmetry is broken. This suggests the presence of an additional transition, or even intervening phases, in the intermediate regime, constituting an interesting (albeit challenging) task for further study.

B. Other studies of twisted spin liquids

There have been a few prior studies of twisted spin liquids. Reference [54] studied the twisted bilayer of the staggered flux state on the square lattice, which is also a candidate mean-field state of a U(1) Dirac spin liquid, but the effect of monopoles was not taken into account. In a similar vein, Ref. [55] explored mean-field spinon band structures of twisted van der Waals magnets hosting U(1) Dirac quantum spin liquids, finding a twist-induced gap opening and arguing that resulting band structures can be tuned upon magnetic encapsulation and applied magnetic fields. Moving away from U(1) Dirac spin liquids, we further mention that in Refs. [56,57] twisted versions of the bilayer Kitaev honeycomb $\mathbb{Z}_2$ spin liquid [58] were studied. While all these prior works are interesting, they differ fundamentally from our results as in our case we account fully for the nature of the Dirac spin liquid as a nontrivial conformal field theory.

C. Conformal renormalization group flow for homogeneous interlayer couplings

In the study at hand, we have primarily focused on instabilities towards ordered phases due to a (simultaneous)

proliferation of monopoles in both layers driven by the monopole interlayer tunneling terms. However, as discussed in Sec. III B, interlayer mass-mass couplings are also symmetry allowed and thought to be relevant. A key question thus pertains if (additional) instabilities can occur due to these additional couplings. To this end, we have performed a perturbative renormalization group calculation to quadratic order, utilizing the operator product expansions admitted by the conformal nature of the QED$_3$ fixed point [59]. Here, we consider the case of $\vec{q} \equiv 0$, i.e., homogeneous interlayer couplings which occur for trivial (AA) stacking, such that combining Eqs. (23) and (27) leads to the full Lagrangian, given by

$$\mathcal{L} = \mathcal{L}_{\mathrm{QED}_3}^{(1)} + \mathcal{L}_{\mathrm{QED}_3}^{(2)} + \mathcal{L}_{\mathrm{int}}^{(12)}, \tag{83a}$$

$$\mathcal{L}_{\mathrm{int}}^{(12)} = \sum_{a=1}^{3} J_{\mathrm{v}}(\Phi_{1,a}^\dagger \Phi_{2,a} + \mathrm{H.c.}) + \sum_{a=4}^{6} J_{\mathrm{s}}(\Phi_{1,a}^\dagger \Phi_{2,a} + \mathrm{H.c.})$$
$$+ g_1 \sum_{i=1}^{3} M_{1,i0} M_{2,i0} + g_2 \sum_{j=1}^{3} M_{1,0j} M_{2,0j}$$
$$+ g_3 \sum_{i,j=1}^{3} M_{1,ij} M_{2,ij}. \tag{83b}$$

As mentioned in Sec. II, the low-energy theory $\mathcal{S} = \int d^3 x \, \mathcal{L}$ is naturally endowed with a UV cutoff (the lattice spacing a_0 which bounds the separation of two (low-energy) operator insertions from below. The conformal renormalization group (RG) is most naturally formulated in real space and proceeds by changing the the UV cutoff $a_0 \to b a_0$ with $b = (1 + \delta l) > 1$ (with δl being infinitesimal) and subsequently integrating out operator insertion pairs (at coordinates x, y) with relative separations in the spherical shell defined by $b a_0 \geqslant |x - y| \geqslant a_0$. We relegate a more detailed discussion to Appendix C and give the RG flow equations in terms of the dimensionless variables $\tilde{J}_{\mathrm{s,v}} = J_{\mathrm{s,v}} a_0^{3-2\Delta_\Phi}$ and $\tilde{g}_i = g_i a_0^{3-2\Delta_M}$.

We thus find the perturbative RG equations (up to quadratic order in $\tilde{J}_{\mathrm{s,v}}$ and $\tilde{g}_{1,2,3}$) to be of the form

$$\frac{d\tilde{J}_{\mathrm{s}}}{dl} = (3 - 2\Delta_\Phi)\tilde{J}_{\mathrm{s}} - c_1(2\tilde{J}_{\mathrm{s}}\tilde{g}_1 + 3\tilde{J}_{\mathrm{s}}\tilde{g}_3),$$

$$\frac{d\tilde{J}_{\mathrm{v}}}{dl} = (3 - 2\Delta_\Phi)\tilde{J}_{\mathrm{v}} - c_1(2\tilde{J}_{\mathrm{s}}\tilde{g}_2 + 3\tilde{J}_{\mathrm{s}}\tilde{g}_3),$$

$$\frac{d\tilde{g}_1}{dl} = (3 - 2\Delta_M)\tilde{g}_1 - 2c_2|\tilde{J}_{\mathrm{s}}|^2,$$

$$\frac{d\tilde{g}_2}{dl} = (3 - 2\Delta_M)\tilde{g}_2 - 2c_2|\tilde{J}_{\mathrm{v}}|^2,$$

$$\frac{d\tilde{g}_3}{dl} = (3 - 2\Delta_M)\tilde{g}_3 - c_2(\tilde{J}_{\mathrm{s}}\tilde{J}_{\mathrm{v}}^* + \tilde{J}_{\mathrm{v}}\tilde{J}_{\mathrm{s}}^*). \tag{84}$$

Here, the constants $c_1 = \pi(c_{\Phi M}^\Phi)^2 > 0$ and $c_2 = 4\pi (c_{\Phi\Phi}^M)^2 > 0$ are given in terms of the OPE coefficients as defined in (13a) and (13b) and are left undetermined at this point. We find several nontrivial fixed points to the RG
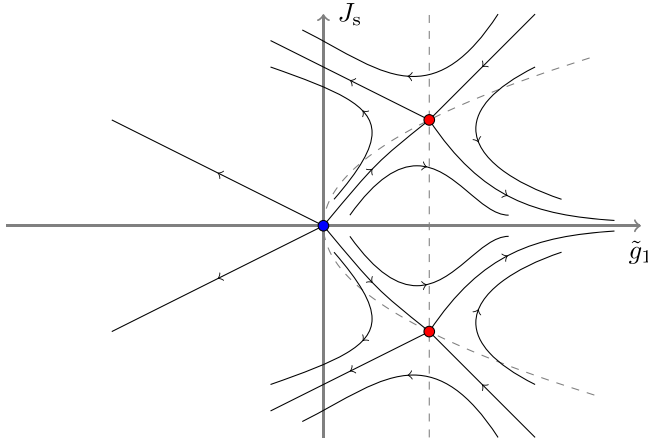
FIG. 5. RG diagram for $\tilde{J}_v = 0 = \tilde{g}_2 = \tilde{g}_3$. See main text for discussions.

equations,

$$
\text{(i)} \ \tilde{J}_s = \pm \frac{1}{\sqrt{4\tilde{c}_1 \tilde{c}_2}}, \quad \tilde{J}_v = 0, \quad \tilde{g}_1 = \frac{1}{2\tilde{c}_1}, \quad \tilde{g}_2 = \tilde{g}_3 = 0,
$$

$$
\text{(ii)} \ \tilde{J}_v = \pm \frac{1}{\sqrt{4\tilde{c}_1 \tilde{c}_2}}, \quad \tilde{J}_s = 0, \quad \tilde{g}_2 = \frac{1}{2\tilde{c}_1}, \quad \tilde{g}_1 = \tilde{g}_3 = 0,
$$

$$
\text{(iii)} \ \tilde{J}_s = \tilde{J}_v = \pm \frac{1}{\sqrt{10\tilde{c}_1 \tilde{c}_2}}, \quad \tilde{g}_1 = \tilde{g}_2 = \frac{1}{5\tilde{c}_1} = \tilde{g}_3,
$$

$$
\text{(iv)} \ \tilde{J}_s = -\tilde{J}_v = \pm \frac{1}{\sqrt{10\tilde{c}_1 \tilde{c}_2}}, \quad \tilde{g}_1 = \tilde{g}_2 = \frac{1}{5\tilde{c}_1} = -\tilde{g}_3, \quad (85)
$$

where we have defined $\tilde{c}_1 = c_1/(3 - 2\Delta_\Phi)$ and $\tilde{c}_2 = c_2/(3 - 2\Delta_M)$. Motivated by the expectation that on a microscopic (lattice) level, spin-spin interactions lead to bare interlayer couplings in the spin sector which are dominant compared to those in VBS sectors, we consider the example of $\tilde{J}_v = 0 = \tilde{g}_2 = \tilde{g}_3$. A schematic resulting RG diagram is shown in Fig. 5. The blue dot corresponds to the *unstable* fixed point of two decoupled copies of QED$_3$, while the two red dots indicate the two *critical* fixed points (i) as given in (85). We observe that there are four distinct regimes of strong coupling, corresponding to (a1) $g_1 \to -\infty, J_s \to -\infty$; (a2) $g_1 \to -\infty$, $J_s \to \infty$; (b1) $g_1 \to \infty, J_s \to 0^-$; (b2) $g_1 \to \infty, J_s \to 0^+$.

In the cases (b1) and (b2), the interlayer mass-mass couplings flow to strong coupling. In this regime, it can therefore be expected that the two masses are spontaneously generated (this simultaneous mass condensation in each layer would correspond to the fermion adjoint operators in each layer acquiring a finite expectation value, $\langle M_{1,i0} \rangle, \langle M_{2,i0} \rangle \neq 0$), with $\tilde{g}_1 > 0$ indicating that the signs of the generated masses are opposite. As discussed in Ref. [34] using a Gross-Neveu type model, in a single layer, the generation of a spontaneous mass opens up a channel for a corresponding monopole to be proliferated, thus constituting a mechanism for a transition from the DSL to ordered phases. In particular, the mass term $\pm M_{c0}$ will proliferate different $\Phi_a \pm i\Phi_b$ monopoles, where $(a - 3, b - 3, c)$ form an even permutation of (1,2,3), and $\Phi_a \pm i\Phi_b$ implies noncollinear, but coplanar, magnetic order. Note that while the perturbative RG flow with respect to the decoupled DSL fixed point appears to indicate that $J_s \to 0^\pm$

in (b1) and (b2), it is to be expected that J_s (and potentially further monopole-monopole interactions) become relevant at the strong-coupling fixed point (i.e., they are *dangerously irrelevant* couplings) and are crucial to the description of the confinement transition in the strong-coupling regimes (b1) and (b2) without referencing a Gross-Neveu theory as in Ref. [34].

For the cases (a1) and (a2), a similar scenario as above (spontaneous mass generation which leads to the proliferation of the two monopoles) is conceivable. However, the fact that $J_s \to \pm \infty$ could also be taken to suggest that the physics in this regime is dominantly controlled by the monopole tunneling J_s growing strong, implying a *direct* instability due to monopole tunneling. As to the nature of the phase that is induced, two scenarios appear likely.

On the one hand, the interlayer monopole-monopole interaction can induce simultaneous monopole proliferation in both layers, thus realizing a (confined) long-range ordered phase. This is the underlying assumption of our work at hand, justifying the use of mean-field theory (see Sec. V) in which the monopole operators acquire finite expectation values $\langle \boldsymbol{\Phi}_1 \rangle, \langle \boldsymbol{\Phi}_2 \rangle \neq 0$ which are in one-to-one correspondence with the Néel/VBS order parameters of the thus obtained ordered phases.

On the other hand, we note that the monopole-monopole tunneling term still preserves the total U(1)$_{top}$ magnetic flux of the system. Thinking about parton constructions [or the low-energy field theory (5)], this could be taken to suggest that excitations coupled to the *relative* gauge field $a_-^\mu \equiv a_1^\mu - a_2^\mu$ between the two layers (the flux of which is not conserved by the monopole tunneling term) become confined (the associated flux is no longer conserved), but the *total* gauge field $a_+^\mu \equiv a_1^\mu + a_2^\mu$ remains in a deconfined state as the total U(1)$_{top}$ flux conservation prohibits monopole proliferation. The resulting state, obtained after such a "partial confinement" transition, would still support deconfined excitations. However, as of now, it is unclear if (1) such a phase, which one might refer to as a "bilayer spin liquid," is *stable* [single-fermion interlayer tunneling may lead to a gap opening in the spinon dispersion, and the resulting pure deconfined U(1) gauge theory is unstable by Polyakov's mechanism [60]] *and* if (2) this phase is *energetically competitive* compared with conventional (fully confined) ordered states. A study of this intriguing scenario, necessary and plausible intermediate transitions involved, and the nature of resulting phases is left for future study.

The fixed points (2), (3), and (4) in Eq. (85) can be analyzed analogously with identical RG diagrams looking the same upon appropriate reparametrizations. Finally, a conformal RG study involving moiré modulations in the interlayer couplings is more involved and will be left for future research.

D. Applications and outlook

Potential material realizations include the van der Waals material 1T-TaS$_2$, which has been proposed to be either a U(1) Dirac or $\mathbb{Z}_2$ spin liquid [33].

We remark that generally in mean-field treatments, the tendency to order might be overestimated. However, experience with the analogous calculations in one lower dimension has

shown them to be often very accurate. In the future, a study of fluctuations above mean-field level is nevertheless important to consider, though it is beyond the scope of this paper. We do want to mention one possible way to study new disordered phases: Consider the couplings of bilayer masses as discussed in Sec. III B and take an appropriate mean-field ansatz such that the moiré low-energy spinon bands are flat (similar to what was found in the twisted bilayer staggered flux state on square lattice [54]). The high density of low-energy states might help to suppress the monopoles and favor a spin liquid phase again. As alluded in the previous subsection, one can conceive spin liquid phases that are intrinsic to the bilayer system, but it is unclear how stable those might be and if they can be energetically favored.

Further, we note that the CFT as a low-energy theory for the coupled U(1) Dirac spin liquids is applicable in the regime of small twist angles and thus large moiré lattice constants $a_\mathrm{m} \gg a$. If the twist angle is large (in the most extreme case $\theta = \pi/6$ since a twist by $0 \leqslant \theta \leqslant 2\pi/6$ is equivalent to a twist by $2\pi/6 - \theta$ due to the C_6 lattice symmetry), there is no clear separation of scales $a_\mathrm{m} \sim a$ which would justify the use of the low-energy (continuum) theory and our results become uncontrolled. Instead, microscopic (lattice) details will become important, an analysis of which we leave for further study.

On the methodological side, we have established explicit operator-product expansions for the low-energy QED$_3$ theory of the Dirac spin liquid on the triangular lattice (similar expansions are readily obtained for, e.g., the kagome lattice), and developed a mean-field framework exploiting the conformal symmetry of the SU(4) DSL. Offering up a unique avenue for analytical treatments beyond large-N calculations, our framework can be applied to the study of a wide array of perturbations to the Dirac spin liquid.

APPENDIX A: DERIVATION OF OPERATOR PRODUCT EXPANSIONS

In order to constrain terms in the operator product expansion (12), we first consider the allowed U(1)$_\mathrm{top}$ quantum numbers (i.e., the topological charges). In particular, the operator product $\Phi_a^\dagger \Phi_b$ has 0 net topological charge so that it is expanded in U(1)$_\mathrm{top}$ singlets. Next, we consider the SO(6) tensorial structure. A first step to decomposing the product of two SO(6) tensor representations is to consider correspond-

ing Young tableaux. However, we note that Young diagrams for the orthogonal groups are in general not irreducible, as (partial) traces can be subtracted.

1. Monopole OPE

It is easily seen that the monopole-monopole operator product $\Phi_a \Phi_b$ possesses a topological charge of 2, such that its expansion may only contain higher-order monopole operators which are notably less relevant and thus excluded from our study.

For the monopole-antimonopole OPE, both transforming as SO(6) vectors, we find

$$\square \otimes \square = \square\square \oplus \begin{array}{c}\square\\\square\end{array}, \tag{A1}$$

where $\square\square$ further splits into the rank-2 traceless symmetric tensor (20-dimensional) and trivial irreps. The 15-dimensional antisymmetric $\begin{array}{c}\square\\\square\end{array}$ representation is isomorphic to the 15 SU(4) adjoint masses. An explicit isomorphism is constructed by writing

$$\Phi_a^\dagger \Phi_b \sim \frac{i}{2} \Phi_\alpha^\dagger L_{\alpha\beta}^{ab} \Phi_\beta + \cdots, \tag{A2}$$

where $L_{\alpha\beta}^{ab} = -i(\delta_\alpha^a \delta_\beta^b - \delta_\beta^a \delta_\alpha^b)$ are SO(6) generators in the defining representation. We can identify two mutually commuting subalgebras $\{L^{23}, L^{31}, L^{12}\}$ and $\{L^{56}, L^{64}, L^{45}\}$ which generate two copies of SO(3). Explicitly comparing the symmetry transformations of $\mathbf{\Phi}^\dagger L^{ab} \mathbf{\Phi}$ (recall that the bold font denotes the six-component vector such that L^{ab} acts as a matrix, with summation over α, β is implied) and the adjoint masses $M_{\mu\nu}$, we find that the first set may be identified with the SO(3)$_\mathrm{valley}$ generators M_{0i} (where $i = 1, 2, 3$) and the latter with M_{i0} which generate SO(3)$_\mathrm{spin}$. We further find that each set $\{L^{i,4}\}_{i=1,2,3}$, $\{L^{i,5}\}_{i=1,2,3}$, $\{L^{i,6}\}_{i=1,2,3}$ furnish a vector representation of SO(3)$_\mathrm{valley}$ and $\{L^{1,j}\}_{j=4,5,6}$, etc., transform as vectors under SO(3)$_\mathrm{spin}$, such that the mixed generators $\mathbf{\Phi}^\dagger L^{ij} \mathbf{\Phi}$ transform as $M_{j-3,i}$ where $i = 1, 2, 3$ and $j = 4, 5, 6$. Note that the C_6 and $\mathcal{T}$ symmetries which reverse the U(1)$_\mathrm{top}$ charges of monopole operators further imply that this identification holds up to a *real* constant, implying $c_{\Phi\Phi}^M \in \mathbb{R}$ in (13a).

This mapping of SO(6) fundamental generators to SU(4) adjoints can be conveniently expressed in terms of a tensor $\mathcal{F}$ (and its inverse $\bar{\mathcal{F}}$ defined via $\mathcal{F}_{\mu\nu}^{ab} \bar{\mathcal{F}}_{ab}^{\rho\lambda} = \delta_\mu^\rho \delta_\nu^\lambda$) which allows us to write $L^{ab} = \mathcal{F}_{\mu\nu}^{ab} M^{\mu\nu}$ and $M^{\mu\nu} = \bar{\mathcal{F}}_{ab}^{\mu\nu} L^{ab}$ [note that all sums over SU(4) adjoint indices (μ, ν) are taken to imply $\mu = \nu = 0$ is excluded]. Matrix elements of $\mathcal{F}$ can be read off explicitly given the above mapping, and we find that the nonzero components are of the form

$$\mathcal{F}_{0i}^{ab} = \epsilon^{abj} \qquad \text{for } a, b \leqslant 3,$$
$$\mathcal{F}_{i0}^{ab} = \epsilon^{a-3,b-3,i} \quad \text{for } a, b \geqslant 4,$$
$$\text{and} \quad \mathcal{F}_{ij}^{ab} = \delta_j^a \delta_{i+3}^b - \delta_{i+3}^a \delta_j^b, \tag{A3}$$

where the latin indices take values $i, j = 1, 2, 3$. An explicit calculation shows that one may take $\bar{\mathcal{F}}_{ab}^{\mu\nu} = \mathcal{F}_{\mu\nu}^{ab}/2$.

2. Monopole-mass OPE

Next, we consider the monopole-mass operator product $\Phi_a M_{\mu\nu}$ which has unit topological charge, suggesting that unit-flux monopole operators appear in the expansion. In order to decompose the product under $SO(6) \simeq SU(4)$, it is convenient to use the previously established mapping $M^{\mu\nu} = \bar{\mathcal{F}}_{ab}^{\mu\nu} L^{ab}$ and decompose the product $\Phi_a L^{bc}$ instead, with Young tableaux

$$\square \otimes \square\!\square = \square\!\square \oplus \square\!\square\!\square , \qquad (A4)$$

where the 70-dim. $\square\!\square$ is further decomposed by subtracting a trace over the horizontal boxes in the second diagram, such that $70 \to 64 + 6$. Note that this result is corroborated by the $SU(4)$ Young tableaux (which are irreducible)

$$\square\!\square \otimes \square = \square\!\square\!\square \oplus \square\!\square\!\square \oplus \square\!\square\!\square \oplus \square , \qquad (A5)$$

with the last diagram on the right-hand side corresponding to the six-dimensional rank-2 antisymmetric representation of $SU(4)$. Explicitly, one can decompose

$$\Phi_a L^{bc} = \tfrac{1}{3}\left(\Phi_a L^{bc} + \text{cyclic}\right)$$
$$+ \tfrac{1}{3}\left(2\Phi^a L^{bc} + \Phi_b L^{ac} - \Phi_c L^{ab} - \tfrac{3}{5} T^{abc}\right)$$
$$+ \tfrac{1}{5}\underbrace{\left(\delta^{ac}\Phi_m L^{bm} - \delta^{ab}\Phi_k L^{ck}\right)}_{=: T^{abc}} . \qquad (A6)$$

Using the explicit transformations in Table I, we find that $\Phi_m L^{bm}$ transforms as $\sim i\Phi_b$ with a *real* constant of proportionality (note that the complex phase is constrained by C_6 and $\mathcal{T}$ symmetries). The additional Kronecker deltas in (A6) keep the antisymmetry in b, c manifest. Using above introduced $\mathcal{F}$ symbols, we thus find $\Phi^a M^{\mu\nu} = \bar{\mathcal{F}}_{bc}^{\mu\nu}\Phi^a L^{bc} \sim \bar{\mathcal{F}}_{ba}^{\mu\nu} i\Phi^b - \bar{\mathcal{F}}_{ac}^{\mu\nu} i\Phi^c$. Using the antisymmetry of $\bar{\mathcal{F}}_{ac}^{\mu\nu} = -\bar{\mathcal{F}}_{ca}^{\mu\nu}$, Eq. (13b) with some $c_{\Phi M}^{\Phi} \in \mathbb{R}$ follows.

APPENDIX B: QUARTIC TERMS IN PERTURBATION THEORY

In this Appendix, we extend the line of thought sketched in Sec. V and determine how the perturbative effective action at quartic order lifts the degeneracy of linearly combining $h_1^a(0)$ and $h_2^a(0)$ at the mean-field saddle point (found at quadratic order). Note that while the effective free energy $F_{\text{mf}}[\boldsymbol{h}_1, \boldsymbol{h}_2] = F_{\text{mf}}[\boldsymbol{h}_1] + F_{\text{mf}}[\boldsymbol{h}_2]$ is linear in the layer index, $F_{\text{mf}}[\boldsymbol{h}_1]$ depends on $\boldsymbol{h}_2$ (or vice versa) once we plug in (53) which holds as a solution of the self-consistency equations at the mean-field saddle point.

The free energy up to quartic order is

$$F_{\text{mf}}[h] = -\log \mathcal{Z}_0 - \frac{\langle \mathcal{S}_h^2 \rangle_0}{2} + \frac{\langle \mathcal{S}_h^2 \rangle_0^2}{8} - \frac{\langle \mathcal{S}_h^4 \rangle_0}{24} + \cdots , \quad (B1)$$

where as usual, $\langle \cdot \rangle_0$ denotes evaluating expectation values in the unperturbed compact QED$_3$. The first two terms on the right-hand side have no preference on the relationship between $h_1^a(0)$ and $h_2^a(0)$ because of the linearity of the self-consistency equations at weak coupling.

The expression for $\langle \mathcal{S}_h^2 \rangle_0$ is relatively straightforward, with

$$\langle \mathcal{S}_h^2 \rangle_0 = 2 \sum_{l,a} \int d^3x \, d^3y \, \frac{h_l^a(x) h_l^{a*}(y)}{|x - y|^{2\Delta_\Phi}} , \qquad (B2)$$

where we have omitted the OPE coefficients for convenience. Plugging in the expressions that we found,

$$a \leqslant 3 : \begin{cases} h_1^a(\vec{x}) = \tilde{h}_1^a(0) - A\tilde{h}_2^a(0) f_s(\vec{x}), \\ h_2^a(\vec{x}) = \tilde{h}_2^a(0) - A\tilde{h}_1^a(0) f_s^*(\vec{x}), \end{cases}$$

$$a \geqslant 4 : \begin{cases} h_1^a(\vec{x}) = \tilde{h}_1^a(0) - A\tilde{h}_2^a(0) f_v^a(\vec{x}), \\ h_2^a(\vec{x}) = \tilde{h}_2^a(0) - A\tilde{h}_1^a(0) f_v^{a*}(\vec{x}), \end{cases} \qquad (B3)$$

and upon evaluating the integrals, one easily observes that the leading terms in (B2) scale as

$$\langle \mathcal{S}_h^2 \rangle_0 \sim L^{6-2\Delta_\Phi}[|\tilde{\boldsymbol{h}}_1(0)|^2 + |\tilde{\boldsymbol{h}}_2(0)|^2], \qquad (B4)$$

where we define the norm of the (complex) $SO(6)$ vectors $|\tilde{\boldsymbol{h}}_l(0)|^2 = \sum_a \tilde{h}_l^a(0)\tilde{h}_l^{a*}(0)$. From (B4), it is clear that the $\langle \mathcal{S}_h^2 \rangle_0$ and $\langle \mathcal{S}_h^2 \rangle_0^2$ (and any higher powers of $\langle \mathcal{S}_h^2 \rangle_0$) together pin down the total magnitude $[|\tilde{\boldsymbol{h}}_1(0)|^2 + |\tilde{\boldsymbol{h}}_2(0)|^2]$, but will not determine the relative ratio $|\tilde{\boldsymbol{h}}_1(0)|^2/|\tilde{\boldsymbol{h}}_2(0)|^2$.

Hence, we now turn to the contribution $\langle \mathcal{S}_h^4 \rangle_0$. Using the OPEs (13a) and (13b), the leading terms come from the fusion channels $(\Phi^\dagger \times \Phi) \times (\Phi^\dagger \times \Phi) \to M \times M \to 1$:

$$\langle \mathcal{S}_h^4 \rangle_0 = -\sum_l \sum_{a \neq b} \int d^3x \, d^3y \, d^3z \, d^3w$$
$$\times \frac{1}{|x - y|^{2\Delta_\Phi - \Delta_M}} \frac{1}{|y - z|^{2\Delta_M}} \frac{1}{|z - w|^{2\Delta_\Phi - \Delta_M}}$$
$$\times \left[h_l^{a*}(x) h_l^b(y) h_l^{a*}(z) h_l^b(w) \right.$$
$$\left. - h_l^{a*}(x) h_l^b(y) h_l^{b*}(z) h_l^a(w) \right] + \text{H.c.} \qquad (B5)$$

The expression above contains various combinations of quartic terms $\tilde{h}_l^*(Q_1)\tilde{h}_l(Q_2)\tilde{h}_l^*(Q_3)\tilde{h}_l(Q_4)$, with Q_i being either zero or some finite momentum wave vector appearing in $f_s(x)$ or $f_v^a(x)$. Naively, one would expect the contributions from the uniform piece $\tilde{h}_l^*(0)\tilde{h}_l(0)\tilde{h}_l^*(0)\tilde{h}_l(0)$ to dominate, as the corresponding integrals contain the biggest power of the IR cutoff L, and scales as $J_\alpha^4 L^{12-4\Delta_\Phi}$. However, the factor A which appears in (B3) defined in (53) contributes nontrivial scaling $A \sim (L/a_m)^{3/2-\Delta_\Phi}$. Carefully taking this into consideration, we find the dominating terms to be ones of the form $\tilde{h}_l^*(0)\tilde{h}_l(Q_1)\tilde{h}_l^*(Q_2)\tilde{h}_l(0)$ with $Q_1, Q_2 \neq 0$. They scale like $A^2 L^{6-4\Delta_\Phi+2\Delta_M}|q|^{2\Delta_M-3} \sim L^{12-4\Delta_\Phi}(L/a_M)^{2\Delta_M-\Delta_\Phi}$, which is more divergent than $L^{12-4\Delta_\Phi}$ as $2\Delta_M > \Delta_\Phi$, such that we expect this term to dominate in the thermodynamic limit $L \to \infty$. Combining with (B3), the leading terms can thus be derived:

$$-\langle \mathcal{S}_h^4 \rangle_0 \sim L^{12-4\Delta} \left(\frac{L}{a_M}\right)^{2\Delta_M - 2\Delta_\Phi}$$
$$\times [|\boldsymbol{h}_1(0) \cdot \boldsymbol{h}_2(0)|^2 - |\boldsymbol{h}_1(0)|^2 \cdot |\boldsymbol{h}_2(0)|^2]. \quad (B6)$$

The first term can be minimized by taking $\tilde{\boldsymbol{h}}_1(0) = s\tilde{\boldsymbol{h}}_2(0)$, with $s \in \mathbb{C}$ being a complex number. In order to fix the relative magnitudes, it is convenient to employ a more symmetric parametrization that keeps the total magnitude fixed, $\tilde{\boldsymbol{h}}_1(0) = \boldsymbol{h}\cos\alpha$ and $\tilde{\boldsymbol{h}}_2(0) = r\boldsymbol{h}\sin\alpha$, where $r \in \mathbb{C}$ is a complex phase, $|r|^2 = 1$, and $\boldsymbol{h}$ is some complex vector whose

magnitude is renormalized by higher-order terms. Above parametrization in terms of α hence keeps the U(1) degeneracy in $[|\tilde{\boldsymbol{h}}_1(0)|^2 + |\tilde{\boldsymbol{h}}_2(0)|^2] = \text{const}$ manifest. Maximizing the second term (B6) amounts to taking $\cos^2\alpha \sin^2\alpha = 1/4$, implying that $|\tilde{\boldsymbol{h}}_1(0)|^2 = |\tilde{\boldsymbol{h}}_2(0)|^2$. Consequently, we find $\tilde{\boldsymbol{h}}_1(0) = r\tilde{\boldsymbol{h}}_2(0)$ where r is a complex phase.

APPENDIX C: PERTURBATIVE CONFORMAL RENORMALIZATION GROUP

In this Appendix, we briefly review the conformal renormalization group [59] and describe its application to the homogeneous bilayer system, which serves as a preparation for the discussions in Sec. VIII C.

The procedure, formulated in real space, consists in (1) raising the UV cutoff $a \to ba$ which sets a minimal separation $a \leqslant |x - y|$ of two operator insertions at points x, y, where $b > 1$, and (2) a subsequent scale transformation which restores the action, but with modified couplings g'. Explicitly, we may consider the partition function $\mathcal{Z} = \int \mathcal{D}[\{\mathcal{O}\}] e^{-\mathcal{S}_0 - \mathcal{S}_g}$, where $\mathcal{S}_0$ is the action at the conformal fixed point, $\mathcal{S}_g = \sum_i a^{\Delta_i - 3} \int d^3x\, g_i \mathcal{O}_i(x)$ is the perturbing action with some operator $\mathcal{O}_i(x)$ with scaling dimension Δ_i. Expanding perturbatively in g, one may write $\mathcal{Z} = \mathcal{Z}_0 \langle 1 - \mathcal{S}_g + \mathcal{S}_g^2 + \cdots \rangle_0$, with the expectation value to be taken with respect to the fixed-point action and over configurations obeying the UV cutoff a.

We now consider the effective action $\mathcal{S}_g'$ obtained by introducing a new cutoff ba with $b > 1$, integrating over those configurations with operator insertion separations below the new cutoff, and rescaling coordinates (and operators) $x = bx'$ such that the effective action has again cutoff a. As the linear term $\langle \mathcal{S}_g \rangle_0$ only features single operator insertions, it is unaffected from raising the cutoff and contributes the "bare" scale transformation of the coupling at the fixed point,

$$a^{\Delta_i - 3} \int d^3x\, g_i \mathcal{O}_i(x) = a^{\Delta_i - 3} \int d^3x'\, g_i b^{3 - \Delta_i} \mathcal{O}_i(x'). \quad \text{(C1)}$$

At quadratic order, one has

$$\langle \mathcal{S}_g \rangle_0 = \sum_{ij} a^{\Delta_i + \Delta_j - 6} g_i g_j \Bigg[\int_{|x-y|>ba} d^3x\, d^3y \, \langle \mathcal{O}_i(x) \mathcal{O}_j(y) \rangle_0$$

$$+ \int_{ba>|x-y|>a} d^3x\, d^3y \, \langle \mathcal{O}_i(x) \mathcal{O}_j(y) \rangle_0 \Bigg]. \quad \text{(C2)}$$

Upon rescaling, the first term in (C2) reproduces the "bare" scaling behavior of the quadratic term. For the second term,

however, $\mathcal{O}_i(x)$ and $\mathcal{O}_j(y)$ are close (in the sense that their separation lies below the raised cutoff) and thus their product is replaced using their OPE (12), which contains (at leading order) a primary operator $\mathcal{O}_k$, and thus contributes to the scaling of the linear term in the expansion of $e^{-\mathcal{S}_g'}$. Considering $b = 1 + \delta l$ with infinitesimal δl, the occurring integral over an infinitesimal shell is evaluated as $\int_{a<|x-y|<(1+\delta l)a} d^3x\, |x - y|^{-\Delta_i - \Delta_j + \Delta_k} = 4\pi a^{3 + \Delta_k - \Delta_i - \Delta_j} \delta l$ and one arrives at the differential RG equations

$$\frac{dg_k}{dl} = (3 - \Delta_k)g_k - 2\pi \sum_{i,j} g_i g_j C_{ij}^k, \quad \text{(C3)}$$

where C_{ij}^k are the OPE coefficients as defined in (12). Turning to the homogeneous bilayer system with Lagrangian (83a), one proceeds analogously. In practice, we find it more convenient to explicitly expand $\exp[-\int d^3x\, \mathcal{L}_{\text{int}}^{(12)}]$ to quadratic order and then read off the respective contributions to the renormalized couplings. Here, we note the identities

$$\sum_{a,i=1}^{3} \mathcal{F}_{0i}^{ab} \mathcal{F}_{0i}^{ac} = 2\delta^{bc} \text{ for } b, c \leqslant 3, \quad \text{(C4)}$$

$$\sum_{a=4}^{6} \sum_{i=1}^{3} \mathcal{F}_{i0}^{ab} \mathcal{F}_{i0}^{ac} = 2\delta^{bc} \text{ for } b, c \geqslant 4, \quad \text{(C5)}$$

$$\sum_{a=1}^{3} \sum_{i,j=1}^{3} \mathcal{F}_{ij}^{ab} \mathcal{F}_{ij}^{ac} = 3\delta^{bc} \text{ for } b, c \geqslant 4, \quad \text{(C6)}$$

$$\sum_{a=4}^{6} \sum_{i,j=1}^{3} \mathcal{F}_{ij}^{ab} \mathcal{F}_{ij}^{ac} = 3\delta^{bc} \text{ for } b, c \leqslant 3, \quad \text{(C7)}$$

$$\sum_{a,b=1}^{3} \mathcal{F}_{\mu\nu}^{ab} \mathcal{F}_{\rho\lambda}^{ab} = 2\delta_{\mu,0}\delta_{\rho,0}\delta_{\nu,i}\delta_{\lambda,j}\delta_{i,j}, \quad \text{(C8)}$$

$$\sum_{a=1}^{3} \sum_{b=4}^{6} \mathcal{F}_{\mu\nu}^{ab} \mathcal{F}_{\rho\lambda}^{ba} = \sum_{a=4}^{6} \sum_{b=1}^{3} \mathcal{F}_{ij}^{ab} \mathcal{F}_{mn}^{ba} = -2\delta_{im}\delta_{jn}, \quad \text{(C9)}$$

where latin indices $i, j, \ldots \in \{1, 2, 3\}$ as usual, which follow straightforwardly from the explicit form of the $\mathcal{F}$ symbol given in (A3). With these results, the flow equations given in (84) follow.

[1] A. K. Geim and I. V. Grigorieva, Nature (London) **499**, 419 (2013).

[2] K. S. Novoselov, A. Mishchenko, A. Carvalho, and A. H. C. Neto, Science **353**, aac9439 (2016).

[3] S. Carr, D. Massatt, S. Fang, P. Cazeaux, M. Luskin, and E. Kaxiras, Phys. Rev. B **95**, 075420 (2017).

[4] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature (London) **556**, 80 (2018).

[5] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 43 (2018).

[6] K. Hejazi, Z.-X. Luo, and L. Balents, Proc. Natl. Acad. Sci. USA **117**, 10721 (2020).

[7] C. Wang, Y. Gao, H. Lv, X. Xu, and D. Xiao, Phys. Rev. Lett. **125**, 247201 (2020).

[8] Y. Xu, A. Ray, Y.-T. Shao, S. Jiang, K. Lee, D. Weber, J. E. Goldberger, K. Watanabe, T. Taniguchi, D. A. Muller, K. F. Mak, and J. Shan, Nat. Nanotechnol. **17**, 143 (2022).

[9] T. Song, Q.-C. Sun, E. Anderson, C. Wang, J. Qian, T. Taniguchi, K. Watanabe, M. A. McGuire, R. Stöhr, D. Xiao, T. Cao, J. Wrachtrup, and X. Xu, Science **374**, 1140 (2021).

[10] Y.-H. Li and R. Cheng, Phys. Rev. B **102**, 094404 (2020).

[11] Q. Tong, F. Liu, J. Xiao, and W. Yao, Nano Lett. **18**, 7194 (2018).

[12] M. Akram and O. Erten, Phys. Rev. B **103**, L140406 (2021).

[13] M. Akram, H. LaBollita, D. Dey, J. Kapeghian, O. Erten, and A. S. Botana, Nano Lett. **21**, 6633 (2021).

[14] F. Xiao, K. Chen, and Q. Tong, Phys. Rev. Res. **3**, 013027 (2021).

[15] K. Hejazi, Z.-X. Luo, and L. Balents, Phys. Rev. B **104**, L100406 (2021).

[16] L. Savary and L. Balents, Rep. Prog. Phys. **80**, 016502 (2017).

[17] N. Karthik and R. Narayanan, Phys. Rev. D **93**, 045020 (2016).

[18] N. Karthik and R. Narayanan, Phys. Rev. D **94**, 065026 (2016).

[19] S. J. Hands, J. B. Kogut, L. Scorzato, and C. G. Strouthos, Phys. Rev. B **70**, 104501 (2004).

[20] L. Di Pietro and E. Stamou, J. High Energy Phys. 12 (2017) 054.

[21] Z. Li, arXiv:1812.09281.

[22] Z. Li, Phys. Lett. B **831**, 137192 (2022).

[23] P. A. Lee, N. Nagaosa, and X.-G. Wen, Rev. Mod. Phys. **78**, 17 (2006).

[24] M. B. Hastings, Phys. Rev. B **63**, 014413 (2000).

[25] M. Hermele, T. Senthil, and M. P. A. Fisher, Phys. Rev. B **72**, 104404 (2005).

[26] Y. Ran, M. Hermele, P. A. Lee, and X.-G. Wen, Phys. Rev. Lett. **98**, 117205 (2007).

[27] M. Hermele, Y. Ran, P. A. Lee, and X.-G. Wen, Phys. Rev. B **77**, 224413 (2008).

[28] Z. Zhu and S. R. White, Phys. Rev. B **92**, 041105(R) (2015).

[29] Y. Iqbal, W.-J. Hu, R. Thomale, D. Poilblanc, and F. Becca, Phys. Rev. B **93**, 144411 (2016).

[30] W.-J. Hu, S.-S. Gong, W. Zhu, and D. N. Sheng, Phys. Rev. B **92**, 140403(R) (2015).

[31] A. Wietek and A. M. Läuchli, Phys. Rev. B **95**, 035141 (2017).

[32] Y. Cui, J. Dai, P. Zhou, P. S. Wang, T. R. Li, W. H. Song, J. C. Wang, L. Ma, Z. Zhang, S. Y. Li, G. M. Luke, B. Normand, T. Xiang, and W. Yu, Phys. Rev. Mater. **2**, 044403 (2018).

[33] K. T. Law and P. A. Lee, Proc. Natl. Acad. Sci. USA **114**, 6996 (2017).

[34] X.-Y. Song, C. Wang, A. Vishwanath, and Y.-C. He, Nat. Commun. **10**, 4254 (2019).

[35] X.-Y. Song, Y.-C. He, A. Vishwanath, and C. Wang, Phys. Rev. X **10**, 011033 (2020).

[36] E. Dyer, M. Mezei, and S. S. Pufu, arXiv:1309.1160.

[37] C.-M. Jian, A. Thomson, A. Rasmussen, Z. Bi, and C. Xu, Phys. Rev. B **97**, 195115 (2018).

[38] Y.-C. He, J. Rong, and N. Su, SciPost Phys. **13**, 014 (2022).

[39] S. Albayrak, R. S. Erramilli, Z. Li, D. Poland, and Y. Xin, Phys. Rev. D **105**, 085008 (2022).

[40] H. J. Schulz, Phys. Rev. Lett. **77**, 2790 (1996).

[41] O. A. Starykh and L. Balents, Phys. Rev. Lett. **93**, 127202 (2004).

[42] O. A. Starykh, H. Katsura, and L. Balents, Phys. Rev. B **82**, 014421 (2010).

[43] T. W. Appelquist, M. Bowick, D. Karabali, and L. C. R. Wijewardhana, Phys. Rev. D **33**, 3704 (1986).

[44] A. Bashir, A. Raya, I. C. Cloët, and C. D. Roberts, Phys. Rev. C **78**, 055201 (2008).

[45] T. Grover, Phys. Rev. Lett. **112**, 151601 (2014).

[46] J. Braun, H. Gies, L. Janssen, and D. Roscher, Phys. Rev. D **90**, 036002 (2014).

[47] I. F. Herbut, Phys. Rev. D **94**, 025036 (2016).

[48] V. Borokhov, A. Kapustin, and X. Wu, J. High Energy Phys. 11 (2002) 049.

[49] S. M. Chester and S. S. Pufu, J. High Energy Phys. 08 (2016) 019.

[50] D. Pappadopulo, S. Rychkov, J. Espin, and R. Rattazzi, Phys. Rev. D **86**, 105043 (2012).

[51] D. Poland, S. Rychkov, and A. Vichi, Rev. Mod. Phys. **91**, 015002 (2019).

[52] J. Jung, A. Raoux, Z. Qiao, and A. H. MacDonald, Phys. Rev. B **89**, 205414 (2014).

[53] R. Feynman, *Statistical Mechanics: A Set Of Lectures*, Advanced Books Classics (CRC Press, Boca Raton, FL, 1998).

[54] Z.-X. Luo, C. Xu, and C.-M. Jian, Phys. Rev. B **104**, 035136 (2021).

[55] G. Chen and J. L. Lado, Phys. Rev. Res. **3**, 033276 (2021).

[56] J. May-Mann and T. L. Hughes, Phys. Rev. B **101**, 245126 (2020).

[57] S. Haskell and A. Principi, Phys. Rev. B **106**, L161404 (2022).

[58] U. F. P. Seifert, J. Gritsch, E. Wagner, D. G. Joshi, W. Brenig, M. Vojta, and K. P. Schmidt, Phys. Rev. B **98**, 155101 (2018).

[59] J. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge Lecture Notes in Physics (Cambridge University Press, Cambridge, 1996).

[60] A. M. Polyakov, Phys. Lett. B **59**, 82 (1975).