Skyrmions in twisted bilayer graphene: stability, pairing, and crystallization

Yves H. Kwan,1 Glenn Wagner,1,2 Nick Bultinck,1,3 Steven H. Simon,1 and S.A. Parameswaran1

1Rudolf Peierls Centre for Theoretical Physics, Parks Road, Oxford, OX1 3PU, UK
2Department of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
3Department of Physics and Astronomy, Ghent University, Krijgslaan 281, 9000 Gent, Belgium

We study the excitations that emerge upon doping the translationally-invariant correlated insulating states in magic angle twisted bilayer graphene at various integer filling factors ν. We identify parameter regimes where these are associated with skyrmion textures in the spin or pseudospin degrees of freedom, and explore both short-distance pairing effects and the formation of long-range ordered skyrmion crystals. We perform a comprehensive analysis of the pseudospin skyrmions that emerge upon doping insulators at even ν, delineating the regime in parameter space where these are the lowest-energy charged excitations by means of self-consistent Hartree-Fock calculations on the interacting Bistritzer-MacDonald model. We explicitly demonstrate the purely electron-mediated pairing of skyrmions, a key ingredient behind a recent proposal of skyrmion superconductivity. Building upon this, we construct hopping models to extract the effective mass of paired skyrmions, and discuss our findings and their implications for skyrmion superconductivity in relation to experiments, focusing on the dome-shaped dependence of the transition temperature on the twist angle. We also investigate the properties of spin skyrmions about the quantized anomalous Hall insulator at ν = +3. In both cases, we demonstrate the formation of robust spin/pseudospin skyrmion crystals upon doping to a finite density away from integer filling.

I. INTRODUCTION

The experimental observation of an array of interaction-driven electronic phases in magic-angle twisted bilayer graphene (TBG)1–4 has stimulated intense theoretical efforts to formulate a unifying physical description of their emergence from the interplay of topology, symmetry, and correlations. One promising route to a solution has close ties to the now-classic paradigm of strongly interacting electrons in Landau levels (LLs). As shown in Ref. 5, this so-called ‘strong-coupling’ picture is rooted in the identification of an idealized limit with a $U(4) \times U(4)$ global symmetry that operates within the subspace of the eight central bands: the two $U(4)$ symmetries rotate within quartets of topological Chern bands with Chern numbers $C = \pm 1$ respectively (see also Refs. 6–9). Correlated insulating phases arise as a result of symmetry breaking within this (almost) degenerate subspace, in a manner analogous to the formation of quantum Hall ferromagnetism in integrally-filled LLs10. This perspective receives support from the detection of a quantized anomalous Hall insulator (QAHI) at ν = +311–13 and various magnetic-field stabilized Chern insulators at other fillings4,14–19. While a purely strong-coupling treatment does not completely resolve the experimental puzzles surrounding TBG, it serves as a natural starting point for considering how various physically realistic corrections influence the competition between ordered states. In this manner, it can aid in the identification of new and distinct broken-symmetry phases that are not accessible in any obvious way from the weakly-interacting limit, but can be viewed as descending from strong-coupling ferromagnetic orders in a natural, albeit non-trivial, fashion as the system is tuned towards the physically relevant intermediate-coupling regime20,21.

The underlying topology of TBG not only influences the selection of correlated states, but also leaves its fingerprint on the nature of their excitations — suggesting a tantalizing link between insulating behaviour at integer filling and superconducting behaviour away from it22–26. In this vein, neutral excitations of the QAHI have been predicted to inherit striking characteristics from the symmetry-breaking and topological properties of the parent insulator27–30. Given the intimate connection to QHFM, the strong-coupling insulators of TBG can also be expected to host charged excitations in the form of spin or pseudospin textures known as skyrmions or merons31–38. For example, under certain conditions additional charge carriers can enter the QAHI as skyrmions rather than as single spin-flip quasiparticles. The situation is richer at even integer fillings since the intricate structure afforded by the approximate $U(4) \times U(4)$ symmetry and the associated anisotropies is more apparent here. Besides spin, the relevant textures now also have access to the valley and sublattice degrees of freedom (additional `flavours’), leading to the emergence of pseudospin skyrmions. The importance of a detailed understanding of such topological objects is underscored by a recent proposal22 of purely electronic superconductivity arising from bosonic ‘paired skyrmions’39,40 of charge 2e (see Fig. 1), whose stabilization arises from the particular features of the strong-coupling Hamiltonian. Such unconventional Cooper pairs may be relevant for understanding the various superconducting domes in TBG2–4,41–45, whose properties and origins remain the subject of intense debate.

To date, however, the investigation of such flavour textures has been limited to either analyses of the effective non-linear sigma model (NLSM)22 or numerical studies of a simplified model mimicking the gross features of the central bands23. While such studies provide qualitative
FIG. 1. Caricature of 1e-skyrmions and paired 2e-skyrmions at $\nu = 0$. Arrows depict the local orientation of pseudospins $n_{\pm}(\mathbf{r})$ in the two Chern sectors ($C = \pm 1$) as defined in Eq. 2 for skyrmions about the KIVC insulator at charge neutrality. Green shaded regions show where the doped electric charge is localized. Blue arrows indicate the center of the skyrmions, and red arrows denote where pseudospins are aligned along the $z$-axis.

insight, a systematic exploration of these charged textures within a realistic microscopic model of TBG remains an outstanding challenge. There is no guarantee that topological textures are relevant once unavoidable complications such as momentum space structure, finite kinetic energy and the presence of a moiré lattice are accounted for. (A recent study of domain-wall textures in the $\nu = 3$ QAHI, where microscopic energetics significantly alter properties of the topologically mandated electronic boundary modes, suggests that it is indeed important to consider such details.) This is the lacuna which we address in this paper, by directly probing the nature of (pseudo)spin skyrmions using unbiased self-consistent Hartree-Fock (HF) calculations within the interacting Bistritzer-MacDonald (BM) model. This requires us to consider completely unrestricted Slater determinants, allowing for full breaking of moiré translation symmetry. This is a challenging task even within the constrained variational space of the HF mean-field ansatz, but one that we pursue successfully below.

We have three main objectives in this work: first, to provide a proof-of-principle demonstration of skyrmion pairing in an appropriately-chosen limit of a microscopic model of TBG; second, to investigate whether such pairing can result in superconducting phenomenology consistent with that seen in experiment (most notably, the dome-shaped dependence of the superconducting transition temperature $T_c$); third, to establish the regime in parameter space over which pairing survives away from the idealized limit. We also consider a subsidiary set of issues pertaining to the emergence of skyrmions at other fillings, and their formation of long-range ordered structures when doped into TBG at a finite density.

Within a HF treatment of the interacting BM model, we find that pseudospin skyrmion pairing does indeed occur about even integer filling and at sufficiently small chiral ratio, and in qualitative agreement with the predictions of the NLSM. Paired skyrmions are lower in energy than particle excitations for small chiral ratio, with relatively better energetics at charge neutrality than at $|\nu| = 2$. For chiral ratios approaching the realistic values, they become energetically unfavorable and harder to stabilize within mean-field theory. External perturbations such as strain and substrate, as well as increased screening of interactions, tend to also have the same effect. We devise a effective skyrmion hopping model to estimate the effective masses, and find that the corresponding superconducting critical temperature $T_c$ can be larger than typical scales observed experimentally. However, more detailed comparison of the twist-angle dependence is complicated by the sensitivity of the results to how electron interactions are incorporated (i.e. the ‘subtraction scheme’, to be explained in Sec. II A). We critically discuss our findings in the context of experiments, and conclude that, at least for some of the samples, skyrmion pairing is unlikely to be mechanism for superconductivity. For the QAHI at $|\nu| = 3$, we demonstrate the formation of spin skyrmions for doping of single charges, and various spin texture lattices at larger dopings.

The balance of this paper is organized as follows. After reviewing the basics of the NLSM and strong-coupling picture of TBG in Sec. II, we focus on the pseudospin skyrmions about the spinless insulator at charge neutrality ($\nu = 0$) in Sec. III. Via an effective skyrmion hopping model, we further estimate the dispersion and effective mass of delocalized skyrmions. We also briefly address the situation at $|\nu| = 2$ in the spinful case. In Sec. IV, we discuss the properties of spin skyrmions and spin texture lattices at $|\nu| = 3$ about the QAHI. We end with a discussion in Sec. V.
II. THEORETICAL CONSIDERATIONS

In this section, we review the key concepts that underpin the existence and properties of skyrmions about the strong-coupling insulators in TBG\textsuperscript{5,8,53}.

A. Interacting BM model

The starting point is the continuum BM model $\hat{H}_{BM}$ which generates the single-particle moiré bands of TBG in the moiré Brillouin zone (mBZ)\textsuperscript{47}. Details of the model are relegated to App. A, but we mention here the important parameters which will influence the skyrmions. The interlayer coupling is parameterized through two sublattice-dependent hopping constants $w_{AA}$ and $w_{AB}$, reflecting corrugation effects in the superlattice. We fix $w_{AB} = 110$ meV but consider variable $w_{AA}$, thereby tuning the value of the chiral ratio $\kappa \equiv w_{AA}/w_{AB}$. Estimated values for $\kappa$ in the literature range from 0.5 to 0.8\textsuperscript{54-56}. The kinetic bandwidth is minimal when the twist angle $\theta$ is close to the magic angle, which is weakly dependent on $\kappa$. While the physical relevance of $\theta$ as a tuning parameter is obvious, it will prove theoretically useful to vary $w_{AA}$ to control deviations from the fully symmetric theory, to be described below. (Note that it has been argued that the effective value of $w_{AA}$ may be susceptible to renormalization towards the chiral limit\textsuperscript{57}, $\kappa \rightarrow 0$, due to the effects of remote bands; thus, tuning $w_{AA}$ may be viewed as phenomenologically modeling this downward renormalization\textsuperscript{9}.)

Sometimes, we will also incorporate heterostrain\textsuperscript{20,58,59} with strength $\epsilon$ and a sublattice potential of strength $\Delta$\textsuperscript{60}. In the absence of such single-particle perturbations, the point group is $D_6$, which includes $\hat{C}_{2z}, \hat{C}_{3z}$, and $M$ (im-plane rotation about the $x$-axis). Accounting for both valley and spin degrees of freedom, the BM model contains a $U(2)_K \times U(2)_{K'}$ flavour symmetry which includes charge/valley conservation and independent $SU(2)$ spin-rotations within each valley. Naturally, a spinless time-reversal symmetry (TRS) $\hat{T} = \tau_x \hat{K}$ is also present, where $\hat{K}$ is the complex conjugation operator. Upon neglecting the weak relative rotation of the Dirac cones in the two layers, there is also a particle-hole symmetry (PHS) $\hat{P}$\textsuperscript{61-63}. Throughout this work we will focus solely on the central eight bands, and project out the remote bands which are separated by a large band gap. (Although, see the preceding comments on their possible role in adjusting $w_{AA}$.)

Repsulsive interactions are implemented by augmenting the BM model with dual-gate screened density-density Coulomb interactions $V(q) = \frac{2\epsilon_r}{2\epsilon_{sr}+q} \tanh q d_{sc}$, where the relative permittivity $\epsilon_r = 10$ and the screening length $d_{sc} = 25$ nm\textsuperscript{5}. By neglecting subleading ‘intervalley-Hund’s couplings’ which scatter electrons between valleys, we retain the global $U(2)_K \times U(2)_{K'}$ symmetry.

Note that interactions may alter the one-body terms somewhat, leading to a total effective single-particle dispersion $H^{SP}$\textsuperscript{5,7,8,63-65}. The precise correction depends on the choice of ‘subtraction scheme’ which arises as follows (a more detailed exposition can be found in App. A). Including only normal-ordered interactions within the central band subspace neglects contributions from the remote bands. Furthermore, some interactions are double-counted because the hopping parameters of graphene (e.g., from density functional theory calculations) are obtained self-consistently with filled valence bands. A convenient parameterization to remedy this and restore PHS is in terms of a reference density matrix $P^0$. There is no unique prescription for fixing $P^0$ and we will consider three choices that have been used in the literature: (i) the ‘average’, (ii) ‘charge neutrality’ (CN), and (iii) ‘graphene’ schemes. This nomenclature is motivated by how $P^0$ is constructed: (i) $P^0 \propto I$, (ii) $P^0$ consists of the filled valence BM bands, and (iii) $P^0$ consists of the filled valence bands of decoupled layers of graphene. Most of the initial discussion on skyrmions will be based on the average scheme, which leads to no corrections when interactions are recast in the strong-coupling form that highlights the strong-coupling hierarchy of scales (see Sec. II B and App. A). The other two schemes will be analyzed in more detail in Sec. III D.

B. Strong-coupling hierarchy

Since the typical Coulomb scale exceeds the kinetic bandwidth, the BM band basis may not be the most natural one in which to understand interaction physics. A more appropriate choice, that reveals the underlying topological character of the bands, is furnished by the Chern basis obtained by diagonalizing the sublattice operator $\sigma_z$. Each band is predominantly polarized on one sublattice, which allows us to label the central bands with sublattice $\sigma$, valley $\tau$, and spin $s$. Combining these into a single index $\alpha$, we can write the Bloch functions as $|\psi_{\alpha}(k)\rangle = e^{i k \cdot \vec{r}} |u_{\alpha}(k)\rangle$ with associated creation operators $\hat{d}_{\alpha}^\dagger(k)$. Crucially, these bands have Chern number $C = \sigma_z \tau_z$ (Fig. 2a).

The link to QHFM is sharpened in the well-studied chiral limit $\kappa = 0$\textsuperscript{57,66}, where the kinetic bandwidth becomes exactly zero at the magic angle and the Chern bands become completely sublattice polarized. In addition, the new chiral symmetry $\{\sigma_z, \hat{H}_{BM}\}$, in concert with $\hat{C}_{2z} \hat{T}$ and $\hat{P}$, heavily constrains the form factors $\Lambda_{\alpha,\beta}(k, q) \equiv \langle u_{\alpha}(k) | u_{\beta}(k + q) \rangle$ to be diagonal and to depend only the Chern number: this allows us to write $\Lambda_{\alpha,\beta}(k, q) = \Lambda^S(k, q) \delta_{\alpha,\beta}$, with\textsuperscript{5}

$$\Lambda^S(k, q) = F^S(k, q)e^{i\phi^S(k, q)\sigma_z \tau_z},$$

(1)

where the superscript ‘S’ denotes ‘symmetric’. In the chiral-flat limit where the effective dispersion $\hat{H}^{SP}$ is neglected, this implies a huge $U(4)_{C=1} \times U(4)_{C=-1}$ symmetry, as can be seen by checking the invariance of the
FIG. 2. Strong-coupling hierarchy and Chern pseudospins. a) Schematic of Chern sectors with $C = \pm 1$. In the isotropic limit, there is a $U(2) \times U(2)$ symmetry ($U(4) \times U(4)$ with spin). At charge neutrality for a generalized ferromagnet with net $C = 0$, one can define pseudospins $n_{\pm}$. The dominant corrections can be captured as anisotropies $J, \lambda$ (wiggly lines) which couple the Chern sectors and partially lift the symmetry. b) NLSM parameters $E_J = J_A U_{UC}$ (squares), $E_\lambda = \lambda_A U_{UC}$ (crosses) as a function of twist angle $\theta$ and interlayer hopping parameter $w_{AA}$, where $A_{UC}$ is the moiré unit cell area. The pseudospin stiffness $\rho_{ps} \simeq 1.5$ meV, calculated in the isotropic limit, does not depend on $w_{AA}$ and subtraction scheme (see App. A), and is largely insensitive to $\theta$. c) Strong-coupling energy scales $U_S, U_A, t_S, t_A$ as a function of $\theta$ and $w_{AA}$. All quantities calculated using the average subtraction scheme.

density operator $\hat{\rho}^{S}(q) = \sum_k \hat{d}^\dagger(k) A^S(k, q) \hat{d}(k + q)$ under independent Chern-number preserving rotations. Remarkably, at $\nu = 0$ one can prove the existence of exact Slater determinant ground states, which simply involve uniformly polarizing any four orthogonal directions within the Chern quartets (‘ferromagnets’)$^{5,8}$. Hence, the chiral-flat limit will also be referred to as the isotropic limit.

The utility of the chiral-flat limit lies in the existence of a hierarchy of scales which permits corrections to be treated perturbatively within the manifold of strong-coupling ferromagnets$^{5,8}$. The energy scale of the full symmetry group is $U_S$, and the main competing scales are the single-particle inter-Chern tunneling ($t_S$) and deviation from chirality ($U_A$), with intra-Chern dispersion ($t_A$) being subleading [see Fig. 2d and App. A]. Focusing for simplicity on the spinless insulator at neutrality such that the parent symmetry group is $U(2) \times U(2)$, the effects of the above perturbations can be captured via anisotropies with positive strengths $J$ and $\lambda$ respectively (i.e. energy scales $E_J = A_{UC} J$ and $E_\lambda = A_{UC} \lambda$ where $A_{UC}$ is the moiré unit cell area). It turns out that the states that get energetically selected by these terms have total $C = 0$. Defining an intra-Chern Pauli triplet $\eta = (\sigma_x \tau_x, \sigma_y \tau_y, \tau_z)$, we can parameterize the strong-coupling ferromagnets residing in the Chern-neutral sector with two Chern-filtered pseudospins

$$n_{\pm}(k) = \langle \Psi | \hat{d}^\dagger(k) \eta \frac{1 \pm \sigma_z \tau_z}{2} \hat{d}(k) | \Psi \rangle$$

which are independent of $k$ (Fig. 2a). Inter-Chern tunneling generates superexchange, inducing an antiferromagnetic (AFM) coupling between the two pseudospins, while a finite $\kappa$ manifests as an AFM coupling in-plane, and an FM coupling out-of-plane. Both break the chiral-flat symmetry to distinct $U(2)$ subgroups. The resulting ground state with $n_{+} = -n_{-}$ (and hence valley-$U(1)_V$ degeneracy) is the Kramers intervalley-coherent insulator (KIVC), so called because it preserves a modified TRS $\hat{T}' = \tau_y \hat{K}$. An equivalent description of the spinless KIVC is via its density matrix $P = \frac{1}{2} (1 + Q_{KIVC})$, where

$$Q_{KIVC} = \sigma_y (\tau_x \cos \phi_{IVC} + \tau_y \sin \phi_{IVC})$$

is parametrized through the IVC angle $\phi_{IVC}$.

C. Non-linear sigma model and pseudospin textures

To understand low-energy excitations and spatially non-uniform configurations, it is useful to consider a con-
Therefore we will mostly be interested in the case of \(\lambda \approx \pm 1\) and \(\rho \approx \pm 1\). The pseudospin stiffness in the isotropic limit, and \(\delta \rho = \rho_+ + \rho_-\) consists of the topological charge (Pontryagin index) densities of the two Chern sectors

\[
\delta \rho(r) = \pm e\rho_{\text{top}, \pm}(r) = \pm \frac{e}{4\pi} n_\pm \cdot \nabla n_\pm \times \nabla n_\pm.
\]

The skyrmion number is given by

\[
N_{\text{top}, \pm} = \int d^2r \rho_{\text{top}, \pm}(r),
\]

where skyrmions and antiskyrmions have \(N_{\text{top}} = +1\) and \(N_{\text{top}} = -1\) respectively.

The NLSPM parameters \(J, \lambda, \rho_{ps}\) are plotted in Fig. 2, with explicit expressions supplied in App. A. In the “averaged scheme” described there, the superexchange has a prominent minimum and nearly vanishes around \(\theta \approx 1.08^\circ\), coincident with where the bare BM bandwidth is smallest. \(\lambda\) is a monotonically increasing function of \(w_{AA}\) and is largely insensitive to \(\theta\). The pseudospin stiffness \(\approx 1.3\text{ meV}\) is a property of the maximally-symmetric manifold and has a very weak dependence on twist angle. While there are sizable regions where \(\lambda > J\), this only occurs for the average scheme where \(J\) is directly tied to the bare BM scale, which is substantially smaller than the bandwidth obtained via STM measurements of the van Hove singularities in the density of states 67–71. Therefore we will mostly be interested in the case \(\lambda > J\), that is conducive to skyrmion pairing as explained below.

Consider first the isotropic limit \(J = \lambda = 0\) where the Chern sectors decouple. The ground state at \(\nu = 0\) is specified by free choices of uniform pseudospins \(n_{\pm}^0\). An additional charge enters as a 1e-skyrmion (in one of the Chern sectors) rather than a particle-like excitation if the energy for a well separated skyrmion-antiskyrmion pair \(8\pi\rho_{ps}\) exceeds the particle-hole gap (\(\Delta_{ph}\)). We can use the infinite-size Polyakov solution 31 which minimizes the gradient energy because of the lack of any Zeeman-like term, meaning that the texture prefers to expand to minimize the Coulomb term.

Reintroducing dispersion leads to an effective AFM coupling \(J\), constraining the parent ground state \(n_{\pm}^0 = -n_{\mp}^0\) which now belongs to an \(SO(3)\) manifold. A skyrmion of radius \(R_s\) in say \(n_+\) now experiences a Zeeman penalty \(\propto R_s^2\) arising from misalignment with \(n_{\mp}^0\). The skyrmion becomes finite with a size determined from the competition between \(J\) and the interaction term. (Note that in order to avoid divergences with system size, the tails of the skyrmion profile must decay faster in space than the Polyakov ansatz 22,32,72.) For large enough \(J\), the skyrmion shrinks and crosses over to an ordinary particle-like excitation. Having instead a finite \(\lambda\) leads to similar conclusions, except with a different residual \(SU(2)\) manifold. Including both perturbations restricts the insulator to a KIVC parameterized by a \(U(1)\) valley angle.

Adding instead a net charge of \(|Q| = 2e\) to the system leads to substantially different conclusions. In the scenario with non-zero \(J\), the AFM interaction leads to the pairing of a skyrmion in \(n_+\) with an antiskyrmion in \(n_-\) (the total charge is 2e, a crucial consequence of the opposite assignment of Chern numbers). To see this, note that the AFM inter-Chern coupling is completely satisfied if the two textures are centered at the same position with exactly the same profile such that \(n_{\text{pair}}(r) \equiv n_+(r) = -n_-(r)\). In this case the resulting ‘paired 2e-skyrmion’ dilates without limit to avoid the Coulomb self-energy. Hence an \(n_+\)-skyrmion and an \(n_-\)-antiskyrmion will bind for any positive value of \(J\), even though the underlying electron-interaction is purely repulsive. A paired skyrmion configuration preserves \(T\)-symmetry, which is a useful numerical diagnostic.

In the presence of an additional \(\lambda\) term, the paired skyrmion experiences an energy penalty \(\propto R_s^2\) from regions where \(n_{\text{pair}}(r)\) is not lying in-plane. This not only leads to a finite size, but also elongates the texture somewhat to reduce the area spent pointing out-of-plane. When \(\lambda\) is comparable to \(\rho_{ps}\), the texture deforms to resemble the topologically equivalent configuration of two paired 1e-merons separated by a finite distance. This can be understood as the \(\lambda\)-term shrinking the costly \(n_{\text{pair}}\parallel \hat{z}\) regions such that they become the cores of the merons (see red arrows in Fig. 1). If \(\lambda/J\) is sufficiently large, pairing is no longer favorable.

### D. Skyrmion superconductivity

We now assume that the microscopic parameters are chosen such that charges enter as skyrmions rather than particles. 1e-skyrmions from opposite Chern sectors attract to form paired skyrmions. Even if particle excitations are slightly lower in energy, this typically occurs in a small region of the mBZ centered at \(\Gamma_M\). Hence above a critical doping, additional charges are expected to form skyrmions 22.

A non-zero superconducting \(T_c\) requires a finite boson effective mass \(M_{\text{pair}}\), which can be motivated within the phenomenological picture of a pair of coupled quantum Hall ferromagnets 22 in equal and opposite magnetic fields. A single 1e-skyrmion feels a net magnetic field and hence has a flat-band dispersion. On the other hand, the magnetic fields experienced by a paired skyrmion cancel out. To understand the generation of an effective mass in this case, imagine a paired skyrmion propagating at some velocity \(v\). The Lorentz forces in the two sectors act to push the constituent 1e-skyrmions in op-
posite directions, which is counteracted by a restoring force depending linearly on \( J \) and the separation (\( \lambda \) is assumed small). By balancing the two, one can deduce that \( M_{\text{pair}} \) is inversely proportional to \( J \). Therefore the paired skyrmions condense with a finite superfluid stiffness and associated Berezenskii-Kosterlitz-Thouless transition temperature \( T_c = \frac{\hbar^2 \nu^2}{2k_B A_{\text{UC}} M_{\text{pair}}} = \frac{\nu J A_{\text{UC}}}{2k_B} \). \hspace{1cm} (7)

We define a corresponding filling-independent energy scale \( E_c = k_B T_c/\nu \), which equals \( E_J/2 \) in this framework. Above, \( \nu \) should be taken to be the damping relative to the integer filling of interest.

Note the sensitive dependence on the superexchange \( J \), which potentially provides a litmus test for the applicability of this mechanism to realistic samples of TDBG. Namely, experiments observe a superconducting dome in \( T_c \) as a function of twist angle, which peaks around \( \theta = 1.08^\circ \). \hspace{1cm} (41)

The value of \( J \), and indeed the effective mass as computed independently in Sec. III D, are also affected by \( \theta \) through the evolution of the bandwidth and character of the BM bands. However as we discuss later, the precise relationship is subtle and differs based on the choice of interaction. In particular, while the nature of the ordered normal state at \( \nu = 0 \) is unambiguous, depending only on the signs of \( J \) and \( \lambda \), finer details such as skyrmion effective masses are susceptible to the specifics of the subtraction scheme (App. A).

### E. Numerical methods

Underpinning the discussion in the following sections are explicit mean-field solutions \( \langle \phi \rangle \) for the (pseudo)spin skyrmions, constructed using unbiased self-consistent HF calculations. Since such configurations break translation and flavor rotation symmetries, the HF density matrix takes the most general form

\[
P_{k,\tau\sigma;k',\tau'\sigma'} = \langle \hat{d}_{k,\tau\sigma}^\dagger \hat{d}_{k',\tau'\sigma'} \rangle. \hspace{1cm} (8)
\]

The mBZ momenta \( k \) take discrete values due to the finite system dimensions \( L_1, L_2 \) measured in units of the moiré lattice vectors \( a_1^M, a_2^M \). Throughout this work we take \( L = L_1 = L_2 \) so that a non-degenerate moiré band can accommodate \( L^2 \) electrons. The filling is fixed by the constraint \( \text{Tr} P = (4 + \nu) L^2 = (4 + \nu_0) L^2 + N \), where \( \nu_0 \) is the parent integer filling factor and \( N \) represents the number of additional doped electrons. We typically initialize the calculations with random initial projectors, and accelerate convergence using the optimal damping algorithm. \hspace{1cm} (73)

Near \( \nu = 0 \), we neglect spin for simplicity and hence omit the index \( s \) when calculating both skyrmion and particle excitations.

For spin textures about the QAH state, the local order parameter is simply the local spin density

\[
S(r) = \sum_{ss'\tau I} \langle \hat{d}_{s\tau I}^\dagger(r) \hat{d}_{s'\tau I}(r) \rangle. \hspace{1cm} (9)
\]

where \( \sigma \) is a Pauli triplet in spin space, and \( I = \{1A, 1B, 2A, 2B\} \) labels the types of electrons living on different layers/sublattices. Replacing \( \sigma \) by the identity gives the local charge density. The real-space fermion creation operator is defined in terms of the Chern basis as

\[
\hat{\psi}_{\tau\sigma}^\dagger(r) = \sum_{k\sigma} \langle \psi_{\tau\sigma}(k); I \rangle \hat{d}_{\tau\sigma}^\dagger(k). \hspace{1cm} (10)
\]

For pseudospin textures, the natural extension is the ‘diagonal’ Chern-filtered pseudospin density

\[
n_{\text{C}}^{\text{diag}}(r) = \sum_I \langle \hat{\psi}_{I\tau\sigma}^\dagger(r) \frac{1 + C \tau \sigma z}{2} \eta \hat{\psi}_I(r) \rangle \hspace{1cm} (11)
\]

where spin has been dropped and \( \eta \) is the triplet of pseudospin Paulis defined above Eq. 2. The net global Chern polarization is defined by the difference in fractional occupation in each Chern sector. Above, the Chern band-filtered real-space creation operator is

\[
\left[ \hat{\psi}_{I\tau\sigma}^\dagger(r) \right]_{\tau\sigma} = \sum_k \langle \psi_{\tau\sigma}(k); I \rangle \hat{d}_{\tau\sigma}^\dagger(k). \hspace{1cm} (12)
\]

where \( \tau \sigma \) labels the Chern band. However, the in-plane components of Eq. 11 identically vanish in the chiral limit \( \kappa = 0 \), where the Chern basis becomes completely sublattice polarized. For example, for the \( C = +1 \) Chern sector, \( n_{+,x} \) and \( n_{+,y} \) should measure coherence between the \( K' \) and \( K'B \) bands, but \( \langle \hat{\psi}_{I\tau\sigma}^\dagger(K\bar{A}) \rangle \) is non-zero only when \( I = 1A, 2A \) or \( 1B, 2B \).

To remedy this, we occasionally consider an alternative definition that includes off-diagonal terms in the layer/sublattice degrees of freedom

\[
n_{\text{C}}^{\text{off}}(r) = \sum_{II'} \langle \hat{\psi}_{I\tau\sigma}^\dagger(r) \frac{1 + C \tau \sigma z}{2} \eta \hat{\psi}_{I'}(r) \rangle. \hspace{1cm} (13)
\]

To address the question of the effective mass, we can perform variational calculations in the space of states obtained by translating a localized paired skyrmion \( \langle \phi \rangle \) by all possible moiré lattice vectors. This defines an effective skyrmion hopping model leading to delocalized ‘Bloch skyrmions’ and a Bloch dispersion, from which \( M_{\text{pair}} \) can be extracted. Note that this way of calculating the paired skyrmion mass is completely different from the classical calculation of \( M_{\text{pair}} \) used to obtain the second equality of Eq. 7. The technical details of the skyrmion-plane wave calculations, including generalizations involving symmetry-related skyrmions, are outlined in App. B.

### III. PSEUDOSPIN TEXTURES AT EVEN INTEGER FILLING

#### A. 1e-skyrmions

Fig. 3a shows the charge density of a 1e-skyrmion relative to the translation-invariant \( \nu = 0 \) ground state. The
FIG. 3. Properties of $1e$-skyrmions (spinless $\nu = 0$). a) Charge density $\delta \rho (r)$ of $1e$-skyrmions in the entire simulation cell relative to $\nu = 0$ ground state. Panels show different combinations of turning on/off dispersion and achirality, which determine $J$ and $\lambda$. Calculations performed on $18 \times 18$ systems at $\theta = 1.05^\circ$. Textures manually translated to the center for clarity. b) Relative energy $\Delta E = E_{1e,sky} - E_{1e}$ of $1e$-skyrmions (compared to particle excitations) as a function of system size. Apart from the isotropic limit (black triangles), $1e$-skyrmions are not favored. c) $\Delta E$ for different $\theta$ and $w_{AA}$. Dispersion is always included. System size is $11 \times 11$. Missing data points indicate parameters where HF was unable to find skyrmion solutions. d) Local alignment of Chern pseudospins corresponding to the rightmost panel of a). All calculations used the average scheme (see App. A).

Chern couplings $J$ and $\lambda$ can be controlled by tuning the effective one-body term and the chiral ratio respectively. Since the Hamiltonian obeys particle-hole symmetry, adding holes instead of electrons leads to analogous results.

In the isotropic limit with $J = \lambda = 0$, the ground state can be any member of the $U(2) \times U(2)$ manifold described by independent choices of $n_+, n_-$. (Strictly speaking the $|C| = 2$ ferromagnets are also ground states, but these do not allow for textures). As verified numerically, the extra charge enters as a delocalized skyrmion solely in one of the Chern sectors, while the other sector remains unchanged. Hence its properties are qualitatively the same as for spin skyrmions about the $\nu = 3$ QAH insulator (see Sec. IV). Consistent with the absence of anisotropies, the skyrmion expands as much as possible and is limited by the finite simulation cell. This explains the slow convergence in Fig. 3b of the $1e$-skyrmion energy (measured with respect to the $1e$-particle excitation) with the linear system dimension $L$, compared with other cases. Fitting to a power law $\Delta E(L) = \Delta E(\infty) + \alpha L^{-\gamma}$, we obtain $\Delta E(\infty) = -4.1 \text{ meV}$, $\alpha = 18.3 \text{ meV}$ and $\gamma = 0.68$. Note that the band gap at $\nu = 0$ depends only weakly on $L$ since there is a direct gap at $\Gamma_M$. Hence in the NLSM picture, the $L$ dependence is expected to predominantly arise from the system size constraint on $R_s$ which controls the interaction contribution to Eq. 4. Given that the gate screening length is comparable to the moiré length $d_{sc} \simeq a_M$, one expects $\gamma > 1$ in the continuum limit for $R_s \gg a_M$. The discrepancy suggests that lattice corrections (the dotted patterns in Fig. 3a denote the AA-stacking regions) and finite size effects have a quantitative impact here.

In the presence of anisotropies, the $1e$-skyrmions become finite-sized. For example in the chiral-nonflat limit with $J \neq 0, \lambda = 0$, the charge density is localized to within a few moiré lengths, leading to faster energy convergence with system size. This is consistent with the intuition from the NLSM, and the energetics of the skyrmion becomes significantly less favorable. The parent insulator is a member of the $SU(2)$ family containing the KIVC and the valley Hall state (e.g. polarizing into $KA$ and $K'A$ bands). The global net Chern polarization is less than $1/L^2$, consistent with the fact that the
superexchange tunnels between opposite Chern sectors.

A similar story holds for the nonchiral-flat limit with $J = 0, \lambda \neq 0$, except the $\nu = 0$ insulator now interpolates between the KIVC and the valley-polarized state. This time, the added skyrmion is perfectly Chern polarized.

Including both perturbations, which is the case for realistic TBG, the symmetry reduces to the $U(1)$ family of KIVC insulators, and the Chern polarization of the skyrmion is imperfect again. Fig. 3d illustrates that the localized violation of Chern anti-alignment occurs at the same position as the charge density modulation, confirming the skyrmionic nature of the added charge. In Fig. 3c, we chart the relative energy of the skyrmions as a function of $\theta$ for different chiral ratios. For the average scheme, the $1e$-skyrmions are actually most costly near the magic angle where $J$ is supressed (compare with 2b). This is despite the fact that artificially turning off $J$ while keeping other parameters fixed improves the skyrmion energy significantly (Fig. 3b). On the other hand, $\lambda$ is largely constant as a function of $\theta$. This suggests that the continuum description in terms of a small number of coupling parameters is not completely adequate.

The numerical results for $\Delta E$ in any plot such as Fig. 3c are generally expected to represent upper bounds for two reasons. First, our calculations are performed on finite system sizes $L$. In the thermodynamic limit $L \to \infty$, the skyrmions will have some ideal radius $R_s(\infty)$ set by the intrinsic properties of the BM model and interaction potential. Unless $L \gg R_s(\infty)$, the pseudospins in our calculation will experience some degree of frustration from the finite simulation cell, leading to an energy penalty. In addition, a larger $L$ introduces a greater number of basis Bloch states which allow for smoother pseudospin rotations. Second, the restriction to Slater determinants in HF likely impacts the continuum description in terms of a small number of coupling parameters.

B. 2e-skyrmions

Fig. 4a shows the textured configurations when two electrons are added to the translation-invariant $\nu = 0$ ground state. As before, the Chern couplings $J, \lambda$ can be controlled through the effective one-body term and the chiral ratio respectively.

For $J = \lambda = 0$, the texture breaks moiré translation symmetry completely, but the charge density appears to spread throughout the system without a clear identification of one or two well-defined objects. Interestingly, the lowest energy solution consists of both additional charges entering the same Chern sector, meaning that from a symmetry standpoint the situation will again be qualitatively similar to adding two charges to the QAHE state. As shown in Fig. 4b, the pseudospin alignment (note that one Chern sector has a constant pseudospin) rotates in a complicated fashion without reconstructing a new superlattice periodicity. However in analogy to the case of $SU(2)$-invariant spin textures in Sec. IV, the inclusion of more doped electrons can result in a so-called ‘double-tetarton lattice’ with emergent periodicity if they all Chern polarize. The resulting pseudospins form an ordered pattern as shown in Fig. 12a–c.

In the chiral non-flat limit with $J \neq 0, \lambda = 0$, the charge density consists of a single smooth modulation with approximate circular symmetry. This is precisely an explicit realization of a paired 2e-skyrmion: a skyrmion and an antiskyrmion with identical spatial profiles in the two Chern sectors exactly overlapping. This binding is induced by the superexchange $J$, as evidenced by the perfect anti-alignment of Chern pseudospins (and hence $T'$-symmetry). In the absence of other perturbations, the paired skyrmions spreads out and is limited only by the system size, leading to slow convergence in Fig. 4c. Hence the resulting physics controlling the texture is similar to that of the 1e-skyrmion in the isotropic limit.

In the nonchiral-flat limit with $J = 0, \lambda \neq 0$, one may naïvely expect a similar paired skyrmion where the pseudospins are locked instead as $n_{+,z} = n_{-,z}$ and $n_{+,xy} = -n_{-,xy}$. However this does not work as it leads to an electrically neutral object. The numerics reveal that both charges predominantly go into the same Chern sector. For small system sizes, the state closely resembles the double-tetarton lattice of the isotropic limit, which is reflected in the energetic trends in Fig. 4c for small $L$. For larger sizes, the lattice deforms such that the texture is better described by a nearby pair of 1e-skyrmions. While this may be considered pairing (actually since $\Delta E > 0$, it is a metastable bound state that is unstable towards decaying into two particles), we reserve the term ‘paired skyrmion’ for the $T'$-symmetric cases where a skyrmion and an antiskyrmion from opposite Chern sectors bind. It is noteworthy that $\Delta E$ increases as a function of $L$, bucking the trends of all other types of 1e- and 2e-skyrmions. This occurs because for $L \lesssim R_s(\infty)$, the constituent skyrmions are forced to strongly interact with each other in the confined system area, and may therefore form a delocalized configuration with better energetics.

It is clear from Fig. 4c that reintroducing dispersion to the achiral-flat limit is favourable to the textures, which are once again paired skyrmions. The faster convergence of the relative energy with $L$ is a hint that these paired skyrmions are now finite in size. This is clear from Fig. 5a, which illustrates that increasing the chiral ratio not only reduces the skyrmion area, but also leads to an elliptical shape. As discussed in Sec. II C, this could be anticipated from the NLSM analysis which predicts that the $\lambda$-term penalizes the pseudospins when they anti-align out-of-plane. This anisotropy is apparent in Fig. 5b,c, where $n_{+,z}(r)$ has a much tighter profile than $n_{+,xy}(r)$. At mean-field level, the orientation of the $n_z$ lobes is very soft, with distinct HF solutions differing by $\lesssim 10\mu eV$. 
FIG. 4. Properties of 2e-skyrmions (spinless \(\nu = 0\)). a) Charge density \(\delta \rho(\mathbf{r})\) of 2e-skyrmions in the entire simulation cell relative to \(\nu = 0\) ground state. Panels show different combinations of turning on/off dispersion and achirality, which determine \(J\) and \(\lambda\). Calculations performed on \(16 \times 16\) systems at \(\theta = 1.05^\circ\). Textures manually translated to the center for clarity. b) Local alignment of Chern pseudospins corresponding to the leftmost panel of a). The system cell has been duplicated three times for presentation. c) Relative energy \(\Delta E = E_{2e-skyr} - E_{2e}\) of 2e-skyrmions (compared to adding two particle excitations) as a function of system size. \(\Delta E < 0\) indicates that the 2e-skyrmion is favored. d) \(\Delta E\) for different \(\theta\) and \(w_{AA}\). Dispersion is always included. Note that most of the solutions (indicated with crosses rather than squares) for \(1.05^\circ < \theta < 1.13^\circ\) are not \(\hat{T}'\)-symmetric paired skyrmions. System size is \(11 \times 11\). All calculations used the average scheme (see App. A).

The energetic trends of the 2e-skyrmions as a function of \(\theta\) and \(w_{AA}\) are shown in Fig. 4d. Note that specific to the average scheme, \(\lambda\) significantly exceeds \(J\) for a finite region of \(\theta\) around the flat-band point. Consequently in the range \(\sim 1.05^\circ\) to \(\sim 1.13^\circ\), the best textured solution has finite Chern polarization and breaks \(\hat{T}'\), and hence is not a paired skyrmion.

Strain and substrate alignment represent two single-particle perturbations that are deleterious to the paired skyrmions, making them less energetically favorable and harder to find within HF. Strain significantly increases the kinetic bandwidth and is known to substantially degrade the gap of the KIVC at charge neutrality, eventually leading to a symmetric semimetal\(^{59}\). As shown in Fig. 6a, strain of strength \(\epsilon \lesssim 0.1\%\) is enough to induce a positive relative energy \(\Delta E\) for paired skyrmions. This should be interpreted in the context of STM studies which typically measure strains of \(0.1 - 0.7\%\)\(^{67,68,70}\). A staggered sublattice potential \(\sim \Delta \sigma_z\) breaks \(C_{2z}\) and acts as a constant easy-axis Zeeman field with opposite direction in the two Chern sectors, which deters smooth rotations in pseudospin space (Fig. 6b). For \(\Delta\) beyond a fraction of an meV, HF fails to stabilize paired skyrmions at all. In fact if the coupling is strong enough, the mean-field ground state becomes a valley Hall state with no intervalley coherence\(^{20}\). On the other hand, increasing the strength of interactions tends to favor the paired skyrmions. This can be achieved by increasing the gate screening distance \(d_{sc}\) (Fig. 6c) or reducing the relative permittivity \(\epsilon_r\) (Fig. 6d).

C. Skyrmion composites and crystallization

By going beyond one or two particles, we can generate (metastable) higher skyrmion composites and skyrmion crystals. Fig. 7a shows a typical configuration obtained by adding 3 electrons on top of \(\nu = 0\), resulting in a 1e-skyrmion (in the \(C = +1\) sector) and a paired 2e-skyrmion that position themselves apart due to Coulomb...
repulsion. This identification can be straightforwardly made by looking at the pseudospin orientation. On the other hand, adding 4 electrons can lead to a $\mathbf{T}'$-symmetric clump of two paired skyrmions which arrange their $n_z$ lobes in order to minimize the gradient cost (Fig. 7b). For larger numbers of particles, the random initialization of the self-consistent HF loop tends to drive the system into local minima with complicated translation symmetry-breaking patterns (Fig. 7f). Such states are often well-characterized by clumps or ‘trains’ of paired skyrmions with suitably arranged pseudospin lobes, punctuated by 1e-skyrmions within the gaps. The prevalence of such motifs, even for large $w_{AA}$ where particle-hole excitations are favorable, points to the robustness of skyrmions as well-defined localized excitations far away from the perturbative strong-coupling regime.

For intermediate dopings, the skyrmions can lose their individual identities and order into textured crystals. This was already shown in the isotropic limit for two added particles in Fig. 4a,b, but can occur even with anisotropies if the inter-skyrmion spacing becomes comparable with $R_s(\infty)$. For 6 extra particles, we can find a meron crystal where the lobes of the paired skyrmions lie on the honeycomb sites (Fig. 7c, compare with Fig. 12d). Fig. 7d,e are examples of double-tetarton lattices with different emergent supercells.

D. Effective mass of paired skyrmions

A key prediction of the NLSM analysis is that paired skyrmions have a finite dispersion, which is crucial for generating a finite BKT energy scale $E_c$ for superconductivity. In the NLSM framework, this generation of a finite mass is non-trivial, arising from the interplay between the superexchange $J$ and the contrasting magnetic fields. An important question is how this picture holds up in the periodic moiré setting, where magnetic fields and flat bands are replaced by inhomogeneous Berry curvatures and interaction-renormalized dispersions. We address these issues using the effective hopping model described in App. B, paying close attention to the dependence on the twist angle $\theta$ and subtraction scheme (the role of the subtraction is explained in Sec. II A and App. A). Because of moiré translation invariance under $\mathbf{T}_R$, our method is in essence a variational calculation of $\langle \psi' | \mathbf{H} | \psi'\rangle = \frac{1}{\cal Z} \sum_{\cal R} e^{\mathbf{q} \cdot \mathbf{R}} T_R | \phi \rangle$ based on a starting localized skyrmion $| \phi \rangle$ with an ideal pseudospin structure at the mean-field level. The required inputs are the matrix elements $\langle \psi' | T_R | \phi \rangle$, $\langle \phi | H T_R | \phi \rangle$ of $| \phi \rangle$ and its translated images.

In Fig. 8a,b, we show the resulting bandstructure of 1e-skyrmions and paired skyrmions, measured relative to the energy of the starting texture. The 1e-skyrmion is characterized by a sharp peak at $\Gamma_M$ and shallow minima near the zone boundaries (the positions of these features may change for larger $w_{AA}$, but the overall structure re-
FIG. 7. Skyrmion composites and crystallization (spinless $\nu = 0$). Charge density $\delta \rho$, inter-Chern pseudospin alignment $n_+ \cdot n_-$, and pseudospin orientation $n_+$ for different numbers of electrons above charge neutrality $\nu = 0$ in the spinless model. Emergent supercells for the double-tetarton lattice are highlighted in grey. All pseudospin densities are diagonal in microscopic layer/sublattice. System size is $16 \times 16$ with average scheme.

mains the same). Usually a large number of overlaps and matrix elements beyond nearest neighbours need to be computed to converge for a fixed set of parameters.

On the other hand, the paired skyrmions are robustly associated with a broad energy minimum at $\Gamma_M$ and peaks at the zone corners. Typically only matrix elements for distances $\lesssim 2a_M$ need to be computed to accurately describe the properties of the full hopping model. The skyrmion mass $M_{\text{pair}}$, and hence the BKT energy scale $E_c$, can be estimated by fitting a parabola to the band minimum. This typically leads to $T_c$ of order 1 K. Fig. 8c plots the convergence of $E_c$ with system size $L$, which implies that the finite-size calculations here will typically underestimate the $L \to \infty$ results. As can be verified by checking nearest neighbor overlaps, there is no ‘orthogonality catastrophe’ for moiré translations—the spatially inhomogeneous part of a paired skyrmion is finite in size, and most regions of the HF state remain translation invariant.

Having established the basic properties of the hopping model for paired skyrmions, we now turn to details of the dependence of $E_c$ on various parameters. The main motivation is to touch base with experiments which have observed a $T_c$ dome as a function of twist angle\(^\text{41}\). However as noted in Sec. III B, there is a large window of $\theta$ centred at the magic angle where paired skyrmions cannot be found in the average scheme. The utility of this scheme is that $J$ can be easily toggled by artificially turning off the kinetic Hamiltonian, but the drawback is that $J$ therefore becomes substantially smaller than $\lambda$ when the BM bands are flat, leading to a different type of $2e$-skyrmion. To address this issue, we consider now two alternative schemes which are not fine-tuned to have a vanishing superexchange.

Fig. 8d charts the relative energy of the paired skyrmion as a function of $\theta$ and $w_{AA}$ for the CN scheme. In contrast to Fig. 4d, the paired skyrmions are energetically favored in an energy window centered around the magic angle. Again, $\Delta E$ has an inverted relation compared to $J$, which is plotted with dotted lines in the right panel. Note that paired skyrmions can be found up to $w_{AA} = 60$ meV for $L = 11$, which should improve with increasing system size. Looking at the hopping model results, $E_c$ appears to qualitatively follow the same trends as $J$ for the whole range of twist angles investigated, suggesting that the physical intuition from the NLSM maintains some level of validity in the lattice case. Notably, the CN scheme suggests that a skyrmion superconductor would have a $T_c$ dome around the magic angle.

The graphene scheme in Fig. 8e paints a somewhat different picture—$E_c$ monotonically increases in concert with $J$ for decreasing twist angle. For smaller angles the applicability of this calculation will be cut off by the fact that the parent insulator is no longer the ground state.
Similarly the paired skyrmions, at least for small chiral ratios, become less favorable compared to particle excitations. Hence there are still possibilities for the hopping model results in the graphene scheme to be consistent with a $T_c$ dome.

The stability of paired skyrmions in the CN, graphene and average schemes is summarized in Fig. 9. Energetically favorable skyrmions can be found up to $w_{AA} \simeq 50$ meV for the right twist angles in the CN and graphene schemes (this is likely an underestimate when accounting for skyrmion delocalization and finite system size). Beyond this, metastable solutions are obtained for chiral ratios as large as $w_{AA} \simeq 65$ meV. Closer to realistic values of $w_{AA} \simeq 80$ meV, we find that HF is not able to converge to any paired skyrmion states. In the average scheme, the paired skyrmions are energetically unfavourable close to the magic angle.

It is notable that different subtraction schemes lead to radically different behaviors in $E_c$. Indeed we believe that this is one of the few cases where such a choice impacts a physical quantity in a qualitative way. Many theoretical studies focus primarily on the type of symmetry-breaking order—$5,6,8,20,64,65,74–83$—all the schemes studied here lead to the KIVC which only requires that $J, \lambda > 0$. On the other hand, we are interested in the $\theta$-dependence of $E_c$ which depends sensitively on the value of $J$ itself.

The origin of this discrepancy can be understood by examining the effective dispersion $H^{SP}$ in more detail. As explained in App. A, its matrix elements take the form

$$h^{SP}(k) = h^{BM}(k) + \frac{1}{2A} \sum_q V_q \Lambda_q(k)Q^0(k+q)\Lambda_q^\dagger(k).$$

(14)

$J$ depends quadratically on the overall magnitude of $h^{SP}$. In the expression above, $P^0 = \frac{1}{2}(1+Q^0)$ is the reference projector of the particular scheme. For the average scheme, $Q^0 = 0$ so that the only single-particle contributions (when the interaction terms are recast in strong-coupling form) arise from the kinetic piece $h^{BM}$ which gets heavily suppressed at the magic angle.

$P^0$ for the CN scheme is constructed by occupying the valence bands of the BM Hamiltonian, while $P^0$ for the graphene scheme is built by filling the decoupled Dirac cones of each layer to the charge neutrality point and projecting to the central bands. An important clue is that both schemes qualitatively agree for $\theta$ above the magic angle. We focus on the states in the vicinity of the Dirac points for simplicity. For large angles, the Dirac cones of TBG are simply renormalized versions of those in decou-
The crucial insight is that the filling of the CN projector tracks this band reversal, while the graphene projector is oblivious to this physics. Note that the second term of Eq. 14 represents the subtraction of the exchange gain for the filled bands of \( \mathcal{P}^0 \). For the CN scheme, this is a positive contribution for the valence band of the BM model, which counteracts the negative contribution from the BM kinetic term itself. Therefore the two parts of Eq. 14 tend to cancel each other out, leading to a dome at the magic angle when \( h_{\text{BM}}(k) \) becomes suppressed. For the graphene scheme, this destructive interference occurs above the magic angle. However when the BM bands swap roles, the graphene projector does not follow, and hence the two parts add constructively for smaller \( \theta \).

We reserve judgment on the matter of which scheme is most appropriate for capturing the physics in experimental TBG. Each choice has its own merits and justifications. The average scheme is the simplest and puts the strong coupling hierarchy front and center. The graphene scheme aims to prevent additional renormalizations of the Dirac cones that have already been accounted for in the input value for the bare Dirac velocity. The CN scheme provides a basepoint (i.e., charge neutrality of the BM model) at which the BM kinetic energy is precisely the mean-field band structure. However it is known to lead to incommensurate Kekulé spiral (IKS) order at extremely small strains for non-zero integer \( \nu \).

The microscopic correction answer is likely complicated and may differ based on the twist angle itself.

Since paired skyrmions also break TRS and all point group symmetries, a natural extension of the effective mass calculation is to include symmetry-related partners of the starting HF solution \( |\phi\rangle \) in a ‘hybridization + hopping’ model (see App. B for details). By restoring the symmetries, this would shed light on the internal structure of the quantum skyrmion, including the angular momentum of pairing. However we find that an orthogonality catastrophe prevents feasibility of this method. Partners that are constructed from symmetries that do not leave the \( \nu = 0 \) KIVC invariant, such as \( \mathcal{T} \), have effectively vanishing overlaps with \( |\phi\rangle \). This is not surprising since these operators have a non-trivial action on the pseudospins even in the bulk of \( |\phi\rangle \) away from the localized paired skyrmion. However partners related by certain symmetries of the KIVC, such as \( C_3 \), also have suppressed matrix elements and overlaps with \( |\phi\rangle \). Fundamentally the reason is that the skyrmion texture is really a many-body pseudospin rotation of the parent insulator involving many particles, as evidenced by Fig. 5b,c. Hence the Hamiltonian, which only involves few-body terms, cannot effectively couple the different partners.

**E. Spinful KIVC at \(|\nu| = 2\)**

We now reintroduce spin and turn to filling factors near \(|\nu| = 2\) where the mean-field ground state is also a

---

**FIG. 9. Stability of paired skyrmions (spinless \( \nu = 0 \)).**

a) Phase diagram in the \( \theta \) vs \( w_{\text{AA}} \) plane showing where paired skyrmions are energetically favorable (blue) or unfavorable (orange) compared to 2 particle-like excitations about the skyrmions are energetically favorable (blue) or unfavorable (orange) for both TBG and decoupled layers. As \( \theta \) is reduced, the BM bands distort until we reach the magic angle regime where the Dirac velocity vanishes and the bandwidth becomes tiny. Continuing past this point, the valence and conduction bands of TBG actually swap roles, the graphene projector does not follow, and hence the two parts add constructively for smaller \( \theta \).

b,c) Same as a) but with graphene and average scheme.

**FIG. 10. Paired skyrmions (spinful \( \nu = +2 \)).**

a) Hartree-Fock dispersion of spin-polarized KIVC insulator at \( \nu = +2 \). Red (black) indicates minority (majority) spin carrier. System size is \( 24 \times 24 \) with \( \theta = 1.06^\circ \). b) Relative energy \( \Delta E = E_{\text{2e}}(b)-\text{skyrv} - E_{\text{2e}}(b) \) of \( \mathcal{T}^\prime \)-symmetric paired 2e- and 2h-skyrmions about \( \nu = +2 \). Note sharp minimum of unfilled band at \( \Gamma_M \). System size is \( 11 \times 11 \) with CN scheme.
KIVC insulator (we only show results for \( \nu = +2 \), but the situation at \( \nu = -2 \) can be inferred from PHS). A representative member of the spin-degenerate manifold of states, arising from the approximate \( SU(2)_K \times SU(2)_{K'} \) symmetry, consists of fully filling the spin-up flat bands while forming a spinless KIVC in the (minority) spin-down subspace. We concentrate on pseudospin skyrmions by enforcing collinear spins.

Paired skyrmions can be constructed as before by treating the majority spin bands as spectator bands and performing non-trivial pseudospin rotations in the half-filled minority spin bands, but there are complications compared with the spinless neutrality case. First, the NLSM parameters \( J, \lambda \) reflect the relevant energy scales at the neutrality point of the strong-coupling Hamiltonian. However the starting insulator now contains additional majority carriers, which impacts the effective energetics of the minority subspace. Second, this additional interaction renormalization enters in a particle-hole asymmetric way. Generally the bands away from neutrality have a significantly enhanced bandwidth with a prominent extremum at \( \Gamma \) (Fig. 10a), which disfavors skyrmions because they are built from momentum states throughout the mBZ. Third, the edge of the majority spin bands can lie inside the minority band gap which reduces the gap to particle-hole excitations. On the other hand, a pseudospin skyrmion cannot be formed by adding holes to a completely filled spin sector.

These considerations are reflected in the HF calculations in Fig. 10b at \( \nu = +2 \), showing that paired 2e-skyrmions on the electron-doped side are more expensive than paired 2h-skyrmions on the hole-doped side. Overall the relative energies \( \Delta E \) are less favorable than the results at charge neutrality.

**IV. SPIN TEXTURES AT \( |\nu| = 3 \)**

The physical intuition and considerations behind skyrmions in Sec. III can also be applied to odd integer fillings. In this section, we briefly discuss textures near \( \nu = +3 \) that arise from doping the Chern, spin and valley-polarized QAHI (the results at \( \nu = -3 \) can be found from particle-hole symmetry about neutrality). The situation here is significantly simpler because of the \( SU(2)_S \) spin-rotation symmetry that holds independent of the presence of dispersion or deviation from the chiral limit. Furthermore the starting insulator is easy-axis in valley space. The low-energy charged topological excitations are then expected to be spin skyrmions \( ^{36,37,51} \) in the partially filled Chern sector without any texturing in the fully filled Chern sector.

In Fig. 11, we compute the properties of 1h- and 1e-skyrmions by relaxing the spin collinearity constraint in our HF calculations. In a similar fashion to isotropic 1e-skyrmions in Fig. 3a, the change in charge density is roughly circularly symmetric with moiré-periodic modulations on the AA-stacking regions. The local spin orientation is consistent with the hallmark features of an \( SU(2) \)-symmetric skyrmion. Indeed the calculation of the topological density \( \rho_{\text{top}} \) (in spin space) confirms the topological origin of the excitation. Unlike in QHFM, there is no explicit Zeeman field, meaning that the skyrmion expands to fill the available system size to minimize the Coulomb energy. This is reflected in the slow convergence of \( \Delta E \) with \( L \). There is a clear asymmetry between adding holes and electrons because the interaction-renormalized band structure \( ^{35-90} \) at finite integer fillings strongly breaks particle-hole symmetry (see Fig. 10a).

Just as in Sec. III C, doping additional charges leads to the formation of various skyrmion crystals. The unit cell of the double-tetarton lattice \(^{51}\) contains two charges with an intricate spin pattern illustrated in Fig. 12a,c. An alternative configuration is the meron crystal whose unit cell is associated with a charge of 3e (Fig. 12d). Computations on system sizes and dopings where both options can be stabilized are required to numerically determine the energetically preferred lattice. It has also been proposed that applying a Zeeman field will enrich
the possible set of phases\textsuperscript{51}.

V. DISCUSSION

In closing, we consider various generalizations and extensions of the work presented above, comment on the possibility of experimentally observing skyrmion physics, and critically assess the implications of our work for the proposed skyrmion-driven mechanism for superconductivity in TBG.

A. More general skyrmions

In this work, motivated by the particular physics and questions relevant to TBG, we have imposed restrictions on the directions in flavor space that the skyrmions are allowed to rotate in. The strong-coupling insulators at different integer $\nu$ are all Chern ferromagnets, and starting from the fully symmetric limit, one can consider more general skyrmions where the only constraint is that locally the state is polarized within the Chern sectors. For instance, starting from the $|\nu| = 3$ QAHI, it is possible to form a texture in both spin and pseudospin to create an ‘entangled’ skyrmion, akin to what happens at $|\nu| = 1$ in the zeroth LL of monolayer graphene\textsuperscript{91}. With the full Hamiltonian, pseudospin rotations will be gapped since the QAHI is easy-axis, while spin rotations remain low-energy due to the $SU(2)_{S}$-symmetry.

These considerations can be generalized to any integer filling. Starting from a generalized ferromagnetic insulator with $\tilde{\nu}_{C}$ filled bands in Chern sector $C$, the local flavor configuration is parameterized by two matrix spinors $f_{C}(r)$ living in Grassmannian projective spaces\textsuperscript{92}

$$f_{C}(r) \in \frac{SU(4)}{SU(\tilde{\nu}_{C}) \times SU(4 - \tilde{\nu}_{C}) \times U(1)}$$

where $\tilde{\nu}_{+} + \tilde{\nu}_{-} = \nu$. The factors in the normal subgroup represent unitary rotations in within the filled bands, unfilled bands, and the phase difference between the filled and unfilled bands respectively. The presence of four flavors and two spinors leads to a large manifold of skyrmions. From the discussion in Sec. II B, the leading corrections from the $U(4) \times U(4)$ limit will be anisotropic couplings between the spinors, while additional single-particle perturbations such as substrate coupling may manifest as anisotropic fields. The relevant space of textures will be dictated by the energetics at a given filling.

The ground state in the spinful model at charge neutrality is the spin-impolarized KIVC\textsuperscript{5}

$$Q_{KIVC} = \sigma_{y}(\tau_{+}V + \tau_{-}V^{\dagger})$$

where $V = e^{i\phi}e^{i\frac{s}{2}\tau_{z}}$ and $\tau_{\pm} = \frac{1}{2} (\tau_{x} \pm i\tau_{y})$. This represents two copies of the spinless version in Eq. 3 with spin projections $\pm m$ and IVC phases $\phi \pm \frac{\phi}{2}$ (the relative phase $\varphi$ is set by the intervalley Hund’s terms). Equivalently we have four Chern pseudospins $n_{C,s}$ where $s$ indicates the projection of the spin along $m$. Consider doping two electrons. Forming a paired skyrmion in one of the spin sectors in the usual way can be understood using the analysis of Sec. III. Attempting to rotate in spin space will either lead to Pauli blocking or the loss of antiferromagnetic exchange $J$, which justifies our restriction to pseudospin textures in Sec. III. Similar conclusions can be reached for the spin-polarized KIVC insulator at $|\nu| = 2$, though we cannot rule out the majority spin being non-trivially involved if it is the closest filled band to the chemical potential.

B. Disorder

The effects of disorder on skyrmions have been treated previously in the quantum Hall context\textsuperscript{93–97}. A strong and smooth disorder potential can drive the system into a spin glass. We note that all the strong-coupling insulators of interest in this work violate TRS, and hence ordinary impurities cannot directly couple to the symmetry-breaking order as a random field. Isolated charged impurities are expected to pin individual skyrmions as well.
as spin texture lattices, which may aid in their detection as discussed in the following subsection. TBG can also harbor more subtle forms of disorder. For the QAHI, modulations in the magnitude and sign of the sublattice coupling $\Delta$ can lead to the enforced nucleation of domain walls, including ones which separate regions with opposite Chern number. These would interrupt the propagation and crystallization of spin skyrmions. Twist angle inhomogeneity, which is pervasive in experimental samples, is notoriously difficult to model, and its theoretical impact on the correlated insulators at integer fillings is still beyond quantitative characterization. For sufficiently long-wavelength fluctuations, the properties of the skyrmions (e.g., the effective masses) are likely to depend on the local twist angle.

C. Detecting skyrmions

The presence of spin textures about the QAHI at $\nu = +3$ could be detected by measuring the degradation of magnetization for small dopings, e.g., through NMR measurements, NV center magnetometry, or using a superconducting quantum interference device (SQUID). Individual spin skyrmions involve a large number of spin flips, and the spin texture lattices of Fig. 12 are close to completely spin unpolarized. However there is an orbital contribution from the spontaneous Chern polarization which is likely larger. A more direct probe would be spin-resolved STM near impurity sites that could pin a localized skyrmion.

Measuring pseudospin textures at even integer filling is trickier since experimental techniques are not able to directly couple to the valley degree of freedom. IVC generally leads to $\sqrt{3} \times \sqrt{3}$ spatial order at the microscopic graphene scale, but the $T'$-symmetry of the KIVC insulator means that it does not exhibit a Kekulé density distortion (KD). Paired skyrmions preserve $T'$ and hence do not give rise to KD. They may still leave a dipole-shaped fingerprint in sublattice polarization within regions where the state is locally in the valley Hall configuration, i.e., pseudospins anti-aligned and pointing along the $z$-axis (Fig. 5c), but this is likely a faint signature since away from the chiral limit the Chern basis is only partially sublattice polarized. On the other hand, 1$c$-skyrmions are tightly localized within a few moiré lengths and give rise to a spatially varying KD when pinned by charged impurities. KD has been observed in the related context of QHFM within the lowest LL of monolayer graphene, including the imaging of an individual valley skyrmion. We caution that KD in TBG has also been predicted for IKS order in the presence of a small amount of strain.

D. Skyrmion superconductivity

We finally turn to the implications of our work for skyrmion superconductivity. First, note that two important factors controlling the feasibility of this proposed mechanism are the stability of paired skyrmions (i.e., relative energy $\Delta E$ compared to particle excitations) and their effective masses. In general, we can conclude that paired skyrmions are especially favored close to the strongly-interacting isotropic limit. The realistic value of $w_{AA}$ lies in the range $55 - 90$ meV (i.e., $\kappa \simeq 0.5 - 0.8$) and we only find paired skyrmions in the lower range of these values (Fig. 9). We note that a mechanism has been proposed that might drive a downward renormalization of the chiral ratio $\kappa$. We note also that large skyrmions, which are relatively classical and relevant for small $w_{AA}$, are likely to be well captured in our mean-field treatment. However, for larger $w_{AA}$, the paired skyrmions become smaller and quantization effects are more important, and fluctuations can be more significant. In this regime the mean-field result is really only an upper bound on the skyrmion energy, which could be lowered by fluctuations, which are not expected to substantially affect the single-particle excitations.

Enhancing interactions by suppressing screening, either through increasing the gate distance $d_{sc}$ or decreasing the permittivity $\epsilon_r$, also favors skyrmions (Fig. 6). However, this observation makes the superconducting domes that persist in Refs. 42–45 upon reducing the interaction strength difficult to reconcile with a topological mechanism. In these experiments, the insulators at integer $\nu$, from which the skyrmions would be seeded, disappear with increasing screening. The superconducting region can also straddle the integers where the BKT transition temperature from Eq. 7 seemingly vanishes. Substrate coupling rapidly destroys pseudospin skyrmions, consistent with the absence of superconductivity in aligned samples. This topological mechanism would not be effective in other moiré platforms that lack $C_{2z}$-symmetry. Strain takes the system away from the strong-coupling regime and similarly disfavors skyrmions: for instance, the parent insulator has been predicted to give way to a symmetric semimetal at charge neutrality or an IKS at non-zero integer fillings. Hence the general expectation from our work and from these experimental observations is that skyrmion superconductivity is most likely to emerge in ‘pristine’ samples with minimal screening.

The question of effective masses (and hence $T_c$ via Eq. 7) is more subtle. We have demonstrated that the dependence of the BKT transition scale $E_c$ on $\theta$ is rather sensitive to precise details of how electron interactions are incorporated. Without further external inputs, e.g., from detailed ab initio studies or spectroscopic probes of the band structure over a range of twist angles, it is difficult to make quantitative contact with experiments such as Ref. 41 which show a dome in $T_c$ near the magic angle. Any comparison would also inevitably be complicated...
by the presence of confounding variables such as twist angle disorder which are difficult to fully characterize, let alone control. However, what we can reliably distill from our numerical study is that paired skyrmions can in principle emerge with a sufficiently mass to support an estimate of $T_c \lesssim 5 \text{ K}$ that is comparable to experimentally observed values. Both the CN and graphene schemes are able to support a non-vanishing superfluid velocity at the magic angle, and the fragility of the parent correlated insulators to deviations of $\theta$ will tend to also reduce the strength of skyrmion superconductivity away from this regime.

The discussion above focuses primarily on a “BEC-like” limit of skyrmion superconductivity, where the binding is present already in the dilute limit; this is the regime primarily accessed in this work. An intriguing alternative possibility is a more “BCS-like” picture, where skyrmions while unstable at low density nevertheless become favored and paired at finite density. We note that Ref. 23 which studies skyrmion superconductivity in the Landau level limit found superconductivity in the doped case, even in the absence of a pairing gap at zero doping.

Another potential challenge to the applicability of the skyrmion mechanism to TBG lies in the fact that superconductivity is most frequently observed near $\nu = -2$ on the side away from charge neutrality. On the other hand, our numerics show that paired skyrmions are relatively harder to stabilize at $|\nu| = 2$ compared to $\nu = 0$ (Fig. 10b). Furthermore, the skyrmions are more dis-favored when doping in the direction away from charge neutrality. This latter observation can be explained by the increased dispersion due to the interaction renormalization from the extra filled bands (Fig. 10a), and has been argued to be consistent with the asymmetry of the Landau fans.

We cannot rule out a scenario where skyrmion superconductivity is operative in only a subset of samples, for instance the device studied in Ref. 4 which is nominally non-aligned with hBN and exhibits an remarkably large number of correlated insulators and superconducting domes, including near neutrality. Another moiré material where the skyrmion mechanism may be a plausible explanation of superconductivity is mirror-symmetric magic-angle twisted trilayer graphene (TTG), which is closely related to TBG but has a somewhat larger value for the magic-angle. Interestingly, the superconductor in TTG is observed to have a very short coherence length, and an associated pseudogap regime. Partly because of this, Refs. 111 and 112 suggested that part of the TTG superconducting dome is in the BEC regime. Skyrmion pairing is a natural way to get preformed charge-2e bosons, and is at least known to give rise to superconductivity in the chiral limit of TBG. It is therefore worthwhile to investigate whether this mechanism can explain at least a subset of the experimental observations in TTG.

In summary, our work provides clear evidence that the formation and pairing of skyrmions can indeed occur in microscopically faithful treatments of TBG, thereby illustrating that a purely electronic “topological” Cooper pairing mechanism can operate away from the exactly-solvable limit without leveraging any approximate sigma-model description. However, despite this in-principle demonstration of the feasibility of a novel pairing mechanism, we cannot on the basis of present evidence definitively attribute superconductivity in TBG to this mechanism. This is highlighted by the difficulty in reconciling the deleterious effect of variations in strain and interaction strength and deviation from the chiral limit on the stability of skyrmions with the apparent robustness of superconductivity to such effects. Despite this, the uncertainty of various microscopic parameters and the mean-field nature of our study leaves open a real possibility that a skyrmionic mechanism may ultimately survive these challenges. Further work, especially involving numerical approaches that can capture fluctuations beyond the mean-field level, is clearly warranted, as are the exploration of new regimes and systems where skyrmion pairing may be more favored and the identification of new probes that can directly interrogate the nature of the pairing ‘glue’ in TBG.

Note Added. A preprint by Eslam Khalaf and Ashvin Vishwanath appearing in the same arXiv posting performs a complementary analysis of 1e-skyrmion formation in the limit of small skyrmion size. In this regime, skyrmions may be viewed as “spin polarons”, consisting of a charged electron or hole bound to a single spin/pseudospin flip particle-hole pair excitiation. Our results are broadly in agreement where they overlap.

ACKNOWLEDGMENTS

We thank Andrei Bernevig, Shubhayu Chatterjee, Eslam Khalaf, Ashvin Vishwanath, and Michael Zaletel for discussions. We are grateful to Eslam Khalaf and Ashvin Vishwanath for sharing an advance copy of their preprint (see Note Added) and to both of them and Mike Zaletel for comments on this manuscript. GW acknowledges NCCR MARVEL funding from the Swiss National Science Foundation. NB is supported by a senior postdoctoral research fellowship of the Flanders Research Foundation (FWO). We acknowledge support from the European Research Council under the European Union Horizon 2020 Research and Innovation Programme, Grant Agreement No. 804213-TMCS and from EPSRC Grant EP/S020527/1. Statement of compliance with EPSRC policy framework on research data: This publication is theoretical work that does not require supporting research data.
1. Yuan Cao, Valla Fatemi, Ahmet Demir, Shiang Fang, Spencer L. Tomarken, Jason Y. Luo, Javier D. Sanchez-Yamagishi, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, Ray C. Ashoori, and Pablo Jarillo-Herrero, “Correlated insulator behaviour at half-filling in magic-angle graphene superlattices,” Nature 556, 80–84 (2018).

2. Yuan Cao, Valla Fatemi, Shiang Fang, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, and Pablo Jarillo-Herrero, “Unconventional superconductivity in magic-angle graphene superlattices,” Nature 556, 43–50 (2018).

3. Matthew Yankowitz, Shaowen Chen, Hryhoriy Polshyn, Yuxuan Zhang, K. Watanabe, T. Taniguchi, David Graf, Andrea F. Young, and Cory R. Dean, “Tuning superconductivity in twisted bilayer graphene,” Science 363, 1059–1064 (2019).

4. Xiaobo Lu, Petr Stepanov, Wei Yang, Ming Xie, Mohammed Ali Aamir, Ipsita Das, Carles Urgell, Kenji Watanabe, Takashi Taniguchi, Guangyu Zhang, Adrian Bachtold, Allan H. MacDonald, and Dmitri K. Efetov, “Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene,” Nature 574, 653–657 (2019).

5. Nick Bultinck, Eslam Khalaf, Shang Liu, Shubhayan Chatterjee, Ashvin Vishwanath, and Michael P. Zaitel, “Ground state and hidden symmetry of magic-angle graphene at even integer filling,” Physical Review X 10 (2020), 10.1103/physrevx.10.031054.

6. Jian Kang and Oskar Vafek, “Strong coupling phases of partially filled twisted bilayer graphene narrow bands,” Phys. Rev. Lett. 122, 246401 (2019).

7. B. Andrei Bernevig, Zhi-Da Song, Nicolas Regnault, and Biao Lian, “Twisted bilayer graphene. iii. interacting hamiltonian and exact symmetries,” Phys. Rev. B 103, 205413 (2021).

8. Biao Lian, Zhi-Da Song, Nicolas Regnault, Dmitri K. Efetov, Ali Yazdani, and B. Andrei Bernevig, “Twisted bilayer graphene. iv. exact insulator ground states and phase diagram,” Phys. Rev. B 103, 205414 (2021).

9. Oskar Vafek and Jian Kang, “Renormalization group study of hidden symmetry in twisted bilayer graphene with coulomb interactions,” Phys. Rev. Lett. 125, 257602 (2020).

10. Steven M. Girvin, “The quantum hall effect: novel excitations and broken symmetries,” in Aspects topologiques de la physique en basse dimension. Topological aspects of low dimensional systems (Springer, 1999) pp. 53–175.

11. Aaron L. Sharpe, Eli J. Fox, Arthur W. Barnard, Joe Finney, Kenji Watanabe, Takashi Taniguchi, M. A. Kastner, and David Goldhaber-Gordon, “Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene,” Science 365, 605–608 (2019).

12. M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, and A. F. Young, “Intrinsic quantized anomalous hall effect in a moiré heterostructure,” Science 367, 900–903 (2020).

13. C. L. Tschirhart, M. Serlin, H. Polshyn, A. Shragai, Z. Xia, J. Zhu, Y. Zhang, K. Watanabe, T. Taniguchi, M. E. Huber, and A. F. Young, “Imaging orbital ferromagnetism in a moiré chern insulator,” Science 372, 1323–1327 (2021).

14. Yu Saito, Jingyuan Ge, Louk Rademaker, Kenji Watanabe, Takashi Taniguchi, Dmitry A. Abanin, and Andrea F. Young, “Hofstadter subband ferromagnetism and symmetry-broken chern insulators in twisted bilayer graphene,” Nature Physics 17, 478–481 (2021).

15. Jeong Min Park, Yuan Cao, Kenji Watanabe, Takashi Taniguchi, and Pablo Jarillo-Herrero, “Flavour hund’s coupling, chern gaps and charge diffusivity in moiré graphene,” Nature 592, 43–48 (2021).

16. Ipsita Das, Xiaobo Lu, Jonath Herzog-Arbeitman, Zhi-Da Song, Kenji Watanabe, Takashi Taniguchi, B Andrei Bernevig, and Dmitri K Efetov, “Symmetry-broken chern insulators and rashba-like landau-level crossings in magic-angle bilayer graphene,” Nature Physics 17, 710–714 (2021).

17. Petir Stepanov, Ming Xie, Takashi Taniguchi, Kenji Watanabe, Xiaobo Lu, Allan H. MacDonald, B. Andrei Bernevig, and Dmitri K Efetov, “Competing zero-field chern insulators in superconducting twisted bilayer graphene,” (2020), arXiv:2012.15126 [cond-mat.mes-hall].

18. Kevin P Nuckolls, Myungchul Oh, Dillon Wong, Biao Lian, Kenji Watanabe, Takashi Taniguchi, B Andrei Bernevig, and Ali Yazdani, “Strongly correlated chern insulators in magic-angle twisted bilayer graphene,” Nature 588, 610–615 (2020).

19. Shuang Wu, Zhenyuan Zhang, K. Watanabe, T. Taniguchi, and Eva Y. Andrei, “Chern insulators, van hove singularities and topological flat bands in magic-angle twisted bilayer graphene,” Nature Materials 20, 488–494 (2021).

20. Yves H. Kwan, Glenn Wagner, Tomohiro Soejima, Michael P. Zaitel, Steven H. Simon, Siddharth A. Parameswaran, and Nick Bultinck, “Kekulé spiral order at all nonzero integer fillings in twisted bilayer graphene,” (2021), arXiv:2105.05857 [cond-mat.str-el].

21. Glenn Wagner, Yves H. Kwan, Nick Bultinck, Steven H. Simon, and S. A. Parameswaran, “Global phase diagram of twisted bilayer graphene above $t_c$,” (2021), arXiv:2109.09749 [cond-mat.str-el].

22. Eslam Khalaf, Shubhayan Chatterjee, Nick Bultinck, Michael P Zaletel, and Ashvin Vishwanath, “Charged skyrmions and topological origin of superconductivity in magic-angle graphene,” Science advances 7, eabf5299 (2021).

23. Shubhayan Chatterjee, Matteo Ippoliti, and Michael P. Zaletel, “Skyrmion superconductivity: Dmrg evidence for a topological route to superconductivity,” (2020), arXiv:2010.01144 [cond-mat.str-el].

24. Maine Christos, Subir Sachdev, and Mathias S. Scheurer, “Superconductivity, correlated insulators, and wess-zumino–witten terms in twisted bilayer graphene,” Proceedings of the National Academy of Sciences 117, 29543–29554 (2020).

25. Eslam Khalaf, Patrick Ledwith, and Ashvin Vishwanath, “Symmetry constraints on superconductivity in twisted bilayer graphene: Fractional vortices, 4e condensates or non-unitary pairing,” (2020), arXiv:2012.05915 [cond-mat.supr-con].

26. Dian Jing, Alexander Conkey Tyner, and Pallab Goswami, “Mechanism of skyrmion condensation
and pairing for twisted bi-layer graphene,” 2021), arXiv:2107.00012 [cond-mat.supr-con].

Fengcheng Wu and Sanlar Das Sarma, “Collective excitations of quantum anomalous hall ferromagnets in twisted bilayer graphene,” Phys. Rev. Lett. 124, 046403 (2020).

Yves H. Kwan, Yichen Hu, Steven H. Simon, and S. A. Parameswaran, “Exciton band topology in spontaneous quantum anomalous hall insulators: Applications to twisted bilayer graphene,” Phys. Rev. Lett. 126, 137601 (2021).

Yves H Kwan, Yichen Hu, Steven H Simon, and SA Parameswaran, “Excitonic fractional quantum Hall hierarchy in moiré heterostructures,” arXiv preprint arXiv:2003.11559 (2020).

Nikolaos Stefanidis and Inti Sodemann, “Excitonic laughlin states in ideal topological insulator flat bands and their possible presence in moiré superlattice materials,” Phys. Rev. B 102, 035158 (2020).

AA Belavin and AM Polyakov, “Metastable states of two-dimensional isotropic ferromagnets,” JETP lett 22, 245–248 (1975).

S. L. Sondhi, A. Karlhede, S. A. Kivelson, and E. H. Rezayi, “Skyrmions and the crossover from the integer to fractional quantum Hall effect at small zeeman energies,” Phys. Rev. B 47, 16419–16426 (1993).

K. Moon, H. Mori, Kun Yang, S. M. Girvin, A. H. MacDonald, L. Zheng, D. Yoshioka, S. Zhang, “Spontaneous interlayer coherence in double-layer quantum Hall systems: Charged vortices and kosterlitz-thouless phase transitions,” Phys. Rev. B 51, 5138–5170 (1995).

H. A. Fertig, L. Brey, R. Côté, and A. H. MacDonald, “Charged spin-texture excitations and the hartree-fock approximation in the quantum Hall effect,” Phys. Rev. B 50, 11018–11021 (1994).

M. Abolfath, J. J. Palacios, H. A. Fertig, S. M. Girvin, and A. H. MacDonald, “Critical comparison of classical field theory and microscopic wave functions for skyrmions in quantum Hall ferromagnets,” Phys. Rev. B 56, 6795–6804 (1997).

Shubhayan Chatterjee, Nick Bultinck, and Michael P. Zaletel, “Symmetry breaking and skyrmionic transport in twisted bilayer graphene,” Phys. Rev. B 101, 165141 (2020).

Yu-Hui Zhang, Dan Mao, Yuan Cao, Pablo Jarillo-Herrero, and T. Senthil, “Nearly flat chern bands in moiré superlattices,” Phys. Rev. B 99, 075127 (2019).

Fengcheng Wu and S. Das Sarma, “Quantum geometry and stability of moiré flatband ferromagnetism,” Phys. Rev. B 102, 165118 (2020).

A. G. Abanov and P. B. Wiegmann, “Chiral nonlinear σ models as models for topological superconductivity,” Phys. Rev. Lett. 86, 1319–1322 (2001).

Tarun Grover and T. Senthil, “Topological spin hall states, charged skyrmions, and superconductivity in two dimensions,” Phys. Rev. Lett. 100, 156804 (2008).

Yuan Cao, Daniel Rodan-Legrain, Jeong Min Park, Noah F. Q. Yuan, Kenji Watanabe, Takashi Taniguchi, Rafael M. Fernandes, Liang Fu, and Pablo Jarillo-Herrero, “Nematicity and competing orders in superconducting magic-angle graphene,” Science 372, 264–271 (2021).

Yu Saito, Jingyuan Ge, Kenji Watanabe, Takashi Taniguchi, and Andrea F Young, “Independent superconductors and correlated insulators in twisted bilayer graphene,” Nature Physics 16, 926–930 (2020).

Petr Stepanov, Ipsita Das, Xiaoabo Lu, Ali Fahimniya, Kenji Watanabe, Takashi Taniguchi, Frank H. L. Koppens, Johannes Lischner, Leonid Levitov, and Dmitri K. Efetov, “Untying the insulating and superconducting orders in magic-angle graphene,” Nature 583, 375–378 (2020).

Xiaoxue Liu, Zhi Wang, Kenji Watanabe, Takashi Taniguchi, Oskar Vafek, and JIA Li, “Tuning electron correlation in magic-angle twisted bilayer graphene using coulomb screening,” Science 371, 1261–1265 (2021).

Myungchul Oh, Kevin P. Nuckolls, Dillon Wong, Ryan L Lee, Xiaoming Liu, Kenji Watanabe, Takashi Taniguchi, and Ali Yazdani, “Evidence for unconventional superconductivity in twisted bilayer graphene,” Nature , 1–10 (2021).

Yves H Kwan, Glenn Wagner, Nilotpal Chakraborty, Steven H Simon, and SA Parameswaran, “Domain wall competition in the chern insulating regime of twisted bilayer graphene,” Physical Review B 104, 115404 (2021).

Rafi Bistritzer and Allan H. MacDonald, “Moiré bands in twisted double-layer graphene,” Proceedings of the National Academy of Sciences 108, 12233–12237 (2011), https://www.pnas.org/content/108/30/12233.full.pdf.

J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, “Graphene bilayer with a twist: Electronic structure,” Phys. Rev. Lett. 99, 256802 (2007).

L. Brey, H. A. Fertig, R. Côté, and A. H. MacDonald, “Skyrmion crystal in a two-dimensional electron gas,” Phys. Rev. Lett. 75, 2562–2565 (1995).

A. G. Green, I. I. Kogan, and A. M. Tsvelik, “Skyrmions in the quantum Hall effect at finite zeeman coupling,” Phys. Rev. B 54, 16838–16849 (1996).

Thomas Bömerich, Lukas Heinen, and Achim Rosch, “Skyrmion and tetarton lattices in twisted bilayer graphene,” Phys. Rev. B 102, 100408 (2020).

David R. Nelson and J. M. Kosterlitz, “Universal jump in the superfluid density of two-dimensional superfluids,” Phys. Rev. Lett. 39, 1201–1205 (1977).

Patrick J. Ledwith, Eslam Khalaf, and Ashvin Vishwanath, “Strong coupling theory of magic-angle graphene: A pedagogical introduction,” (2021), arXiv:2105.08858 [cond-mat.str-el].

Patrick J. Ledwith, Eslam Khalaf, Ziyan Zhu, Stephen Carr, Efthimios Kaxiras, and Ashvin Vishwanath, “TB or not TB? Contrasting properties of twisted bilayer graphene and the alternating twist n-layer structures (n = 3, 4, 5,...),” arXiv e-prints , arXiv:2111.11060 (2021), arXiv:2111.11060 [cond-mat.str-el].

Nguyen N. T. Nam and Mikito Koshino, “Lattice relaxation and energy band modulation in twisted bilayer graphene,” Phys. Rev. B 96, 075311 (2017).

Stephen Carr, Shiang Fang, Ziyan Zhu, and Efthimios Kaxiras, “Exact continuum model for low-energy electronic states of twisted bilayer graphene,” Phys. Rev. Research 1, 013001 (2019).

Grigory Tarnopolsky, Alex Jura Kruchkov, and Ashvin Vishwanath, “Origin of magic angles in twisted bilayer graphene,” Phys. Rev. Lett. 122, 106405 (2019).

Zhen Bi, Noah F. Q. Yuan, and Liang Fu, “Designing flat bands by strain,” Phys. Rev. B 100, 035448 (2019).

Daniel E. Parker, Tomohiro Soejima, Johannes Hauschild, Michael P. Zaletel, and Nick Bultinck, “Strain-induced...
quantum phase transitions in magic-angle graphene,” Phys. Rev. Lett. 127, 027601 (2021).

60. Jeil Jung, Ashley M DaSilva, Allan H MacDonald, and Shaffique Adam, “Origin of band gaps in graphene on hexagonal boron nitride,” Nature communications 6, 1–11 (2015).

61. Liujun Zou, Hoi Chun Po, Ashvin Vishwanath, and T. Senthil, “Band structure of twisted bilayer graphene: Emergent symmetries, commensurate approximants, and wanner obstructions,” Phys. Rev. B 98, 085435 (2018).

62. Zhida Song, Zhijun Wang, Wujun Shi, Gang Li, Chen Fang, and B. Andrei Bernevig, “All magic angles in twisted bilayer graphene are topological,” Phys. Rev. Lett. 123, 036401 (2019).

63. Kasra Hejazi, Xiao Chen, and Leon Balents, “Hybrid wannier chern bands in magic angle twisted bilayer graphene and the quantized anomalous hall effect,” Phys. Rev. Research 3, 013242 (2021).

64. Ming Xie and A. H. MacDonald, “Nature of the correlated insulator states in twisted bilayer graphene,” Phys. Rev. Lett. 124, 097001 (2020).

65. Shang Liu, Eslam Khalaf, Jong Yeon Lee, and Ashvin Vishwanath, “Nematic topological semimetal and insulator in magic-angle bilayer graphene at charge neutrality,” Phys. Rev. Research 3, 013033 (2021).

66. Patrick J. Ledwith, Grigory Tarnopolsky, Eslam Khalaf, and Ashvin Vishwanath, “Fractional Chern insulator states in twisted bilayer graphene: An analytical approach,” Phys. Rev. Research 2, 023237 (2020).

67. Alexander Kerelsky, Leo J McGilly, Dante M Kennes, Lede Xian, Matthew Yankowitz, Shaowen Chen, K Watanabe, T Taniguchi, James Hone, Cory Dean, et al., “Maximized electron interactions at the magic angle in twisted bilayer graphene,” Nature 572, 95 (2019).

68. Youngjoon Choi, Jeannette Kemmer, Yang Peng, Alex Thomson, Harpreet Arora, Robert Polski, Yiran Zhang, Hechen Ren, Jason Alicea, Gil Refael, et al., “Electronic correlations in twisted bilayer graphene near the magic angle,” Nature Physics 15, 1174–1180 (2019).

69. Yuhang Jiang, Xinyuan Lai, Kenji Watanabe, Takashi Taniguchi, Kristjan Haule, Jinhai Mao, and Eva Y Andrei, “Charge order and broken rotational symmetry in magic-angle twisted bilayer graphene,” Nature 573, 91–95 (2019).

70. Yonglong Xie, Biao Lian, Berthold Jäck, Xiaomeng Li, Cheng-Li Chiu, Kenji Watanabe, Takashi Taniguchi, B Andrei Bernevig, and Ali Yazdani, “Spectroscopic signatures of many-body correlations in magic angle twisted bilayer graphene,” Nature 572, 101 (2019).

71. Dillon Wong, Kevin P. Nuckolls, Myungchul Oh, Biao Lian, Yonglong Xie, Sangjun Jeon, Kenji Watanabe, Takashi Taniguchi, B. Andrei Bernevig, and Ali Yazdani, “Cascade of electronic transitions in magic-angle twisted bilayer graphene,” Nature 582, 198–202 (2020).

72. K. Lejnell, A. Karlhede, and S. L. Sondhi, “Effective-action studies of quantum hall spin textures,” Phys. Rev. B 59, 10183–10193 (1999).

73. Konstantin N. Kudin, Gustavo E. Scuseria, and Eric Cancès, “A black-box self-consistent field convergence algorithm: One step closer,” The Journal of Chemical Physics 116, 8255–8261 (2002), https://aip.scitation.org/doi/pdf/10.1063/1.1470195.

74. Ming Xie and Allan H. MacDonald, “Weak-field Hall Resistivity and Spin/Valley Flavor Symmetry Breaking in MAtBG,” arXiv e-prints, arXiv:2010.07928 (2020), arXiv:2010.07928 [cond-mat.mes-hall].

75. Tommaso Cea and Francisco Guinea, “Band structure and insulating states driven by coulomb interaction in twisted bilayer graphene,” Phys. Rev. B 102, 045107 (2020).

76. Yi Zhang, Kun Jiang, Ziqiang Wang, and Fuchun Zhang, “Correlated insulating phases of twisted bilayer graphene at commensurate filling fractions: A hartree-fock study,” Phys. Rev. B 102, 035136 (2020).

77. Jian Kang and Oskar Vafek, “Non-abelian dirac node braiding and near-degeneracy of correlated phases at odd integer filling in magic-angle twisted bilayer graphene,” Phys. Rev. B 102, 035161 (2020).

78. Fang Xie, Aditya Cowsik, Zhi-Da Song, Biao Lian, B. Andrei Bernevig, and Nicolas Regnault, “Twisted bilayer graphene: vi. an exact diagonalization study at nonzero integer filling,” Phys. Rev. B 103, 205416 (2021).

79. Tomohiro Soejima, Daniel E. Parker, Nick Bultnick, Johannes Hauschild, and Michael P. Zaletel, “Efficient simulation of moire materials using the density matrix renormalization group,” Phys. Rev. B 102, 205111 (2020).

80. Pawel Potasz, Ming Xie, and Allan H. MacDonald, “Exact Diagonalization for Magic-Angle Twisted Bilayer Graphene,” arXiv e-prints, arXiv:2102.02256 (2021), arXiv:2102.02256 [cond-mat.str-el].

81. Jianpeng Liu and Xi Dai, “Theories for the correlated insulating states and quantum anomalous hall effect phenomena in twisted bilayer graphene,” Phys. Rev. B 103, 035427 (2021).

82. Gal Shavit, Erez Berg, Ady Stern, and Yuval Oreg, “Theory of correlated insulators and superconductivity in twisted bilayer graphene,” (2021), arXiv:2107.08486 [cond-mat.str-el].

83. Shihao Zhang, Xin Lu, and Jianpeng Liu, “Correlated insulators, density wave states, and their nonlinear optical response in magic-angle twisted bilayer graphene,” arXiv preprint arXiv:2109.11441 (2021).

84. Kasra Hejazi, Chunxiao Liu, Hassan Shapourian, Xiao Chen, and Leon Balents, “Multiple topological transitions in twisted bilayer graphene near the first magic angle,” Phys. Rev. B 99, 035111 (2019).

85. Jian Kang, B. Andrei Bernevig, and Oskar Vafek, “Cascades between light and heavy fermions in the normal state of magic angle twisted bilayer graphene,” (2021), arXiv:2104.01145 [cond-mat.str-el].

86. Andrew T. Pierce, Yonglong Xie, Jeong Min Park, Eslam Khalaf, Seung Hwan Lee, Yuan Cao, Daniel E. Parker, Patrick R. Forrester, Shaowen Chen, Kenji Watanabe, Takashi Taniguchi, Ashvin Vishwanath, Pablo Jarillo-Herrero, and Amir Yacoby, “Unconventional sequence of correlated chern insulators in magic-angle twisted bilayer graphene,” (2021), arXiv:2101.04123 [cond-mat.mes-hall].

87. Francisco Guinea and Niels R. Walet, “Electrostatic effects, band distortions, and superconductivity in twisted graphene bilayers,” Proceedings of the National Academy of Sciences 115, 13174–13179 (2018).

88. Louk Rademaker, Dmitry A. Abanin, and Paula Mellado, “Charge smoothening and band flattening due to hartree corrections in twisted bilayer graphene,” Phys. Rev. B 100, 205114 (2019).

89. Tommaso Cea, Niels R. Walet, and Francisco Guinea, “Electronic band structure and pinning of fermi energy to van hove singularities in twisted bilayer graphene: A self-
Appendix A: Interacting BM model and Hartree-Fock

1. BM Model

Here we review the definition of the Bistritzer-MacDonald (BM) model for the case of unstrained TBG. Modifications relevant to the strained case are described in the Appendix of Ref. 20.

The basis direct lattice vectors (LVs) and reciprocal lattice vectors (RLVs) for the unstrained TBG moiré pattern (where layers 1 and 2 are rotated counterclockwise by \( \theta/2 \) and \(-\theta/2\) respectively) are chosen to be

\[
\begin{align*}
\mathbf{a}_1^M &= \frac{4\pi}{3k_\theta} \left( \frac{\sqrt{3}}{2}, 1 \right), \quad \mathbf{a}_2^M = \frac{4\pi}{3k_\theta} (0, 1), \quad \mathbf{G}_1^M = \sqrt{3}k_\theta (1, 0), \quad \mathbf{G}_2^M = \sqrt{3}k_\theta \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)
\end{align*}
\]

where \( k_\theta = 2k_D \sin \frac{\theta}{2} \), and \( a \) is the graphene C-C bond length. The intracell coordinates for the A and B atoms of graphene are \( \tau_A = a \left( \frac{\sqrt{3}}{2}, \frac{1}{2} \right) \) and \( \tau_B = a \left( \frac{\sqrt{3}}{2}, -\frac{1}{2} \right) \). \( \mathbf{K}_D = \frac{4\pi}{3\sqrt{3}a} (1, 0) \) is the valley-\( K \) (i.e. \( \tau = + \)) Dirac point for untwisted graphene. The Dirac point in valley \( K' \) (\( \tau = - \)) is found by time reversal.

The intralayer kinetic term in valley \( \tau \) is given by the Dirac cones

\[
\langle k, l | \hat{H}^{BM} | k', l \rangle = \hbar v_F \mathbf{k} \cdot \left( \begin{array}{cc} \tau\sigma_x & \delta_{kk'} \\ -\sigma_y & \delta_{kk'} \end{array} \right)
\]

where the momenta \( \mathbf{k} \) above are measured from the Dirac points of the respective layer. Note that we have neglected the small Pauli rotation terms which is required to strictly satisfy particle-hole symmetry \( \mathcal{P} \).

The interlayer hopping term for valley \( + \) is

\[
\langle k, 1 | \hat{H}^{BM} | k', 2 \rangle = T_1 \delta_{k-k',0} + T_2 \delta_{k-k',\mathbf{G}_1^M+\mathbf{G}_2^M} + T_3 \delta_{k-k',\mathbf{G}_2^M}
\]

where the matrices act in sublattice space. The equations in valley \( K' \) can be found by time-reversal. Note that now, \( \mathbf{k} \) in both layers is measured with respect to a common basepoint. Therefore the hopping term is moiré-periodic.

By solving the BM model with a momentum cutoff and projecting to the central bands, one obtains a set of Bloch functions (8 bands in total including spin and valley). The Chern basis is constructed by diagonalizing the sublattice \( \alpha \) operator \( \sigma_z \) projected onto the central subspace. Combining the ‘sublattice’, valley, and spin indices into a common index \( \alpha \), the Bloch functions and associated creation operators can be written as \( |\psi_\alpha(\mathbf{k})\rangle = e^{ik\cdot\mathbf{r}} |u_\alpha(\mathbf{k})\rangle \) and \( \hat{d}_\alpha^\dagger (\mathbf{k}) \). The Chern number of the bands are \( C = \tau_x \sigma_z \), and the form factors are \( [\Lambda_\alpha(\mathbf{k})]_{\alpha\beta} = \langle u_\alpha(\mathbf{k}) | u_\beta(\mathbf{k}+\mathbf{q}) \rangle \) which satisfy \( \Lambda_\alpha(\mathbf{k}) = \Lambda_\alpha^\dagger (\mathbf{k}+\mathbf{q}) \).

In the Chern basis (with appropriate gauge-fixing, see Ref. 5), the form factors are heavily constrained by \( \hat{C}_x \hat{T} = \sigma_x \hat{\mathcal{C}} \) and (anticommuting) \( \hat{P} \hat{T} = 2g \sigma_y \) symmetries. The sublattice-symmetric and antisymmetric parts of the form factors are

\[
\Lambda_\alpha^S(\mathbf{k}) = F_\alpha^S(\mathbf{k}) e^{i\phi_\alpha^S(\mathbf{k})\tau_z}, \quad \Lambda_\alpha^A(\mathbf{k}) = \sigma_x \tau_z F_\alpha^A(\mathbf{k}) e^{i\phi_\alpha^A(\mathbf{k})\tau_z}.
\]

(A8)
2. Strong-coupling form

Consider first the normal-ordered interaction Hamiltonian in the projected subspace (this will later be augmented by interaction-induced corrections)

\[ \hat{H} = \hat{H}^{BM} + \hat{H}^{\text{int,n-o}} \]

\[ = \sum_k d^\dagger(k)h^{BM}(k)d(k) + \frac{1}{2A} \sum_q V_q : \hat{\rho}^\dagger_q \hat{\rho}_q : \]

\[ = \sum_k d^\dagger(k)h^{BM}(k)d(k) + \frac{1}{2A} \sum_{qkk'}(\alpha) V_q [\Lambda_q(k)\Lambda^\dagger_{-q}(k')]_{\gamma\delta} \delta^\dagger_{\alpha}(k)\hat{d}^\dagger_{\gamma}(k')\hat{d}_{\delta}(k'-q)\hat{d}_{\beta}(k+q) \]

\[ \hat{\rho}_q = \sum_k d^\dagger(k)\Lambda_q(k)d(k+q) \]

where \( h^{BM} \) are the BM kinetic matrix elements in the Chern basis. Consider decoupling the normal-ordered interaction term with the projector \( P_{\alpha\beta}(k) = \langle \hat{d}^\dagger_{\alpha}(k)\hat{d}_{\beta}(k) \rangle \)

\[ h^{HF,\text{int}}[P](k) = \frac{1}{A} \sum_G V_G\Lambda_G(k) \sum_{k'} \text{Tr}[P(k')\Lambda^\dagger_{G'}(k')] - \frac{1}{A} \sum_q V_q\Lambda_q(k)P(k+q)\Lambda^\dagger_q(k). \]  

To complete the Hamiltonian, we need to account for the interaction contribution of the remote bands, as well as the fact that some interactions have been double-counted (for example the normal-ordered interaction is not particle-hole symmetric about neutrality due to the self direct and exchange interaction of the filled central bands). There is no unique prescription for doing this, but a convenient parameterization is in terms of some reference projector \( P^0 \). Assuming that the net contributions arising from Bloch states in the remote bands cancel, the total Hamiltonian can be written

\[ \hat{H} = \hat{H}' - \hat{H}^{HF,\text{int}}[P^0] \]

\[ = \sum_k d^\dagger(k)h(k)d(k) + \frac{1}{2A} \sum_q V_q : \hat{\rho}^\dagger_q \hat{\rho}_q : \]

\[ h(k) = h^{BM}(k) - h^{HF,\text{int}}[P^0](k). \]

In this way at the reference density \( P = P^0 \), the HF spectrum is the same as that of the BM model.

We now recast the Hamiltonian in ‘strong-coupling’ form, which makes the hierarchy of scales more transparent. Letting \( P^0(k) = \frac{1}{2}[1 + Q^0(k)] \), the expression for \( h(k) \) becomes

\[ h(k) = h^{BM}(k) - \frac{1}{2A} \sum_G V_G\Lambda_G(k) \sum_{k'} \text{Tr}[\Lambda^\dagger_{G'}(k')] + \frac{1}{2A} \sum_q V_q\Lambda_q(k)\Lambda^\dagger_q(k) + \frac{1}{2A} \sum_q V_q\Lambda_q(k)Q^0(k+q)^T\Lambda^\dagger_q(k) \]

where a term vanished because \( P^0 \) is assumed to be a neutral projector arising from a \( \tilde{C}_{2z}\tilde{T} \) and \( \tilde{P}\tilde{T} \) symmetric Hamiltonian. The normal-ordered interaction can be rewritten

\[ \hat{H}^{\text{int,n-o}} = \frac{1}{2A} \sum_{qkk'} V_q[\hat{d}^\dagger_k\Lambda_q(k)d(k+q)][\hat{d}^\dagger_{k'}\Lambda_{-q}(k')d_{k'-q}] - \frac{1}{2A} \sum_{qk} V_q[\hat{d}^\dagger_k\Lambda_q(k)\Lambda_{-q}(k+q)d_k]. \]

The second term of Eq. A18 cancels with the third term of Eq. A17. Collecting the rest of the terms, we obtain the strong-coupling form of the total Hamiltonian

\[ \hat{H} = \sum_k d^\dagger(k)h^{SP}(k)d(k) + \frac{1}{2A} \sum_q V_q \delta^\dagger \hat{\rho}_q \delta \hat{\rho}_q \]

\[ h^{SP}(k) = h^{BM}(k) + \frac{1}{2A} \sum_q V_q\Lambda_q(k)Q^0(k+q)^T\Lambda^\dagger_q(k) \]

\[ \delta \hat{\rho}_q = \hat{\rho}_q - \bar{\rho}_q, \quad \bar{\rho}_q = \frac{1}{2} \sum_{Gk} \delta_{G,q} \text{Tr}\Lambda_G(k). \]
Dividing the effective single-particle term into inter-sublattice and (sub-dominant) intra-sublattice pieces
\[ h_{SP}^\rho(k) = h_0(k)\tau_z + h_x(k)\sigma_x + h_y(k)\sigma_y \tau_z, \]  
we can define the energy scales of the strong-coupling hierarchy
\[ U_S = \frac{1}{2AN} \sum_{kq} V_q |F^S_q(k)|^2 \]  
\[ U_A = \frac{1}{2AN} \sum_{kq} V_q (|F^S_q(k)|^2 + |F^A_q(k)|^2) \]  
\[ t_S = \frac{1}{N} \sum_k |h_x(k) + ih_y(k)| \]  
\[ t_A = \frac{1}{N} \sum_k |h_0(k)|. \]  
Similarly, the NLSM parameters are
\[ E_J = JA_{UC} = \frac{1}{N} \sum_{kk'} [h_x(k) + ih_y(k)][\mathcal{H}_{ch}^{-1}]_{kk'} [h_x(k') + ih_y(k')] \]  
where \[ [\mathcal{H}_{ch}]_{kk'} = \frac{1}{A} \sum_q V_q |F^S_q(k)|^2 \left[ \delta_{kk'} - \delta_{k',k+q} \epsilon^{2ioq(k)} \right] \]  
\[ E_A = \lambda A_{UC} = \frac{1}{2AN} \sum_{kq} V_q |F^A_q(k)|^2 \]  
\[ \rho_{ps} = -\frac{1}{8A^2} \sum_{k,q} V_q M_k - q - q' M_k(k + q', -q') \]  
\[ M_k(q, q') = \Lambda_q(k)\Lambda_q'(k + q) - \Lambda_q'(k)\Lambda_q(k + q'). \]  
The pseudospin stiffness \( \rho_{ps} \) is a property of the fully symmetric limit.

**Appendix B: Effective skyrmion hopping/hybridization models**

In this appendix, we describe the construction of effective hopping/hybridization models for localized skyrmions about the KIVC state at even integer filling.

### 1. One- and two-body matrix elements

We first review the general expressions for the matrix elements of one- and two-body operators between different Slater determinants. Consider two general Slater determinants with the same particle number
\[ |A\rangle = \prod_{m \leq N} a_m^\dagger |0\rangle, \quad |B\rangle = \prod_{n \leq N} b_n^\dagger |0\rangle \]  
\[ a_m^\dagger = \sum_\alpha A_{\alpha m} d_\alpha^\dagger \]  
\[ b_n^\dagger = \sum_\alpha B_{\beta n} d_\beta^\dagger \]  
where the \( d_\alpha^\dagger \) operators represent some natural underlying basis of the Hamiltonian. The \( m, n \) indices above only run up to \( N \), so the transformation matrices \( A, B \) are of dimension \( N_b \times N \), where \( N_b \) is the Hilbert space dimension. An important quantity is the \( N \times N \) matrix \( S_{mn} \) which relates the natural bases of the two states \( |A\rangle \) and \( |B\rangle \)
\[ \{ a_m^\dagger, b_n \} = (A^T B^*)_{mn} \equiv S^*_{mn} \]  
\[ \{ a_m, b_n^\dagger \} = (A^\dagger B)_{mn} \equiv S_{mn} \]  
\[ \langle A | B \rangle = \det S^{A,B}, \]
Hence we obtain the final result
\[ P_{\alpha,\beta}^{A,B} = \delta_{\alpha,\beta} \det S^{A,B} - \langle A | d^{\dagger}_{\alpha} d_{\beta} | B \rangle \] (B7)

where \( \cdot \) indicates vector/matrix structure in that index.

Now the two-body matrix elements
\[ P_{\alpha,\beta,\gamma,\delta}^{A,B} = \delta_{\alpha,\beta} \delta_{\gamma,\delta} - \langle A | d^{\dagger}_{\alpha} d^{\dagger}_{\beta} d_{\gamma} d_{\delta} | B \rangle \] (B12)

But using \( S^{(\delta,\gamma), (\alpha,\beta)} = \delta_{\alpha,\beta} \delta_{\gamma,\delta} - \delta_{\alpha,\gamma} \delta_{\beta,\delta} \) and an identity for determinants of block matrices
\[ \det \left( \begin{array}{cc} S^{(\delta,\gamma), (\alpha,\beta)} & S^{(\delta,\gamma),B} \\ S^{A,(\alpha,\beta)} & S^{A,B} \end{array} \right) = \langle A | d^{\dagger}_{\alpha} d^{\dagger}_{\beta} d_{\delta} d_{\gamma} | B \rangle \] (B14)

Hence we obtain the final result
\[ P_{\alpha,\beta,\gamma,\delta}^{A,B} = P_{\alpha,\beta}^{A,B} \delta_{\gamma,\delta} + P_{\beta,\gamma}^{A,B} \delta_{\alpha,\delta} - P_{\alpha,\gamma}^{A,B} \delta_{\beta,\delta} - P_{\beta,\delta}^{A,B} \delta_{\alpha,\gamma} - \delta_{\alpha,\gamma} \delta_{\beta,\delta} \] (B17)

\[ + \det (A^{\dagger} B) \left( \begin{array}{cc} \delta_{\alpha,\delta} & \delta_{\beta,\delta} \\ \delta_{\alpha,\gamma} & \delta_{\beta,\gamma} \end{array} \right) - [B(A^{\dagger} B)^{-1} A^{\dagger}]_{(\delta,\gamma), (\alpha,\beta)} \] (B18)

2. Application to TBG — single-component hopping

We first discuss the case of a single type of localized skyrmion (e.g. we neglect partners related by point group symmetries in the case of paired skyrmions with substantial ellipticity). Recall how a localized skyrmion \( |\phi_{0}\rangle \) is constructed from the eigenvectors \( v^{n}_{k,\tau,a}(0) \) of the HF Hamiltonian (the \( 0 \) denotes that the skyrmion is centered at some arbitrarily chosen origin moiré site)
\[ \sum_{k'\tau\delta} \mathcal{H}_{k'\tau\delta}^{HF}[P(0)] v^{n}_{k',\tau',a}(0) = \epsilon^{n} v^{n}_{k,\tau,a}(0) \] (B19)

\[ \text{i.e. fill orbitals } \sum_{k\tau a} v^{n}_{k,\tau,a}(0) d^{\dagger}_{k\tau a} \text{ up to some } n. \] (B20)

We construct skyrmions \( |\phi_{R}\rangle \) at all other moiré sites \( R \) by boosting the eigenvectors \( v^{n}_{k,\tau,a}(R) = e^{ikR} v^{n}_{k,\tau,a}(0) \). Treating the set of eigenvectors as a matrix \( [v(R)]_{\alpha,n} = v^{n}_{\alpha}(R) \), where \( \alpha = (k, \tau, a) \), we obtain expressions for the
overlap, one-body, and two-body correlators

\[ S(R) = \langle \phi_0 | \phi_R \rangle = \det \left[ v^\dagger(0)v(R) \right] \]  

(B21)

\[ P^1_{\alpha,\beta}(R) = \langle \phi_0 | d^\dagger_\alpha d_\beta | \phi_R \rangle \]

(B22)

\[ = S(R) \left[ \left( v(R) \left( v(0)^\dagger v(R) \right)^{-1} v(0) \right)^\dagger \right]_{\beta,\alpha} \]  

(B23)

\[ P^2_{(\alpha,\beta):(\gamma,\delta)}(R) = \langle \phi_0 | d^\dagger_\alpha d^\dagger_\beta d_\gamma d_\delta | \phi_R \rangle \]

(B24)

\[ = S(R)(\delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma}) + P^1_{\alpha,\beta}(R)\delta_{\beta\gamma} + P^1_{\beta,\gamma}(R)\delta_{\alpha\delta} - P^1_{\beta,\delta}(R)\delta_{\alpha\gamma} - P^1_{\alpha,\gamma}(R)\delta_{\beta\delta} \]

(B25)

\[ + \det(S(R)) \det \left( \begin{pmatrix} \delta_{\alpha\delta} & \delta_{\beta\gamma} \\ \delta_{\alpha\gamma} & \delta_{\beta\delta} \end{pmatrix} - \left[ v(R)(v(0)^\dagger v(R))^{-1} v(0)^\dagger \right]_{(\delta,\gamma),(\alpha,\beta)} \right) . \]  

(B26)

The other correlators (say between a skyrmion at \( R_1 \) and \( R_2 \) can be obtained by translation invariance of the Hamiltonian.

The hopping model is constructed using the overlaps and matrix elements of the localized skyrmions

\[ S_{R,R'} = \langle \phi_R | \phi_{R'} \rangle = S(R - R') \]  

(B27)

\[ H_{R,R'} = \langle \phi_R | \hat{H} | \phi_{R'} \rangle = H(R - R') \]  

(B28)

where the Hamiltonian matrix elements are

\[ H(R) = \sum_{\alpha\beta} H^{SP}_{\alpha,\beta} P^1_{\alpha,\beta}(R) \]

(B29)

\[ + \frac{1}{2\lambda} \sum_{\tau' abcd} \sum_{k^\alpha k^\beta qG} V(q, G) \lambda_{\tau,ab}(k^\alpha;q,G)\lambda_{\tau',dc}(k^\beta;q,G) P^2_{(\alpha,\beta):(\gamma,\delta)}(R) \]

(B30)

\[ \alpha = (k^\alpha, \tau, a) \]

(B31)

\[ \beta = (k^\beta + q, \tau', c) \]

(B32)

\[ \gamma = (k^\alpha + q, \tau, d) \]

(B33)

\[ \delta = (k^\beta, \tau', b) \]

(B34)

and \( H^{SP} \) is the effective one-body term which includes interaction-induced dispersion corrections.

The resulting model is solved by simple diagonalization

\[ | \psi \rangle = \sum_R c_R | \phi_R \rangle \]

(B35)

\[ \sum_{R'} [S^{-1}H]_{R,R'} c_{R'} = Ec_R. \]

(B36)

Due to moiré translation invariance, this leads to ‘Bloch skyrmions’ with plane wave coefficients and a skyrmion bandstructure in the mBZ.

3. Application to TBG — multi-component hybridization

Next we discuss the case of multiple components of localized skyrmion, for example the \( \hat{C}_3 \)-related partners of a given paired skyrmion. Here we ignore the hopping problem, and instead consider how the components (all localized at the same moiré site) hybridize. We consider \( N_c \) components indexed by \( l \), with corresponding skyrmion Slater determinants \( | \phi_l \rangle \) and HF eigenvectors \( v^n_{k,\tau,a}(l) \). Treating the set of eigenvectors as a matrix \( |v(l)\rangle_{\alpha,n} = v^n_\alpha(l) \), we
obtain expressions for the overlap, one-body, and two-body correlators

\[ S(l, l') = \langle \phi_l | \phi_{l'} \rangle = \det \left[ v^\dagger(l) v(l') \right] \]  
(B37)

\[ P^{1}_{\alpha,\beta}(l, l') = \langle \phi_l | d^\dagger_\alpha d_\beta | \phi_{l'} \rangle = S(l, l') \left( v(l') (v(l)^\dagger v(l'))^{-1} v(l) \right)_{\beta,\alpha} \]  
(B38)

\[ P^{2}_{(\alpha,\beta);(\gamma,\delta)}(l, l') = \langle \phi_l | d^\dagger_\alpha d^\dagger_\gamma d_\delta d_\beta | \phi_{l'} \rangle = S(l, l') (\delta_{\alpha\gamma} \delta_{\beta\delta} - \delta_{\alpha\delta} \delta_{\beta\gamma}) + P^{1}_{\alpha,\delta}(l, l') \delta_{\beta\gamma} + P^{1}_{\beta,\gamma}(l, l') \delta_{\alpha\delta} - P^{1}_{\beta,\delta}(l, l') \delta_{\alpha\gamma} - P^{1}_{\alpha,\gamma}(l, l') \delta_{\beta\delta} \]  
(B39)

\[ + \det(S(l, l')) \det \left( \begin{pmatrix} \delta_{\alpha\delta} & \delta_{\beta\delta} \\ \delta_{\alpha\gamma} & \delta_{\beta\gamma} \end{pmatrix} - \left[ v(l') (v(l)^\dagger v(l'))^{-1} v(l) \right]_{(\delta,\gamma),(\alpha,\beta)} \right). \]  
(B40)

The Hamiltonian is constructed in an analogous way to App. B 2.

### 4. Application to TBG — multi-component hopping

The most general case if we consider multiple components and include hopping. We consider skyrmions with position \( R \) with \( N_c \) components indexed by \( l \). We have corresponding skyrmion Slater determinants \( |\phi_{R,l}\rangle \) and HF eigenvectors \( \psi_{k,l,a}(R;l) \). We obtain expressions for the overlap, one-body, and two-body correlators

\[ S(R; l, l') = \langle \phi_{0,l} | \phi_{R;l'} \rangle = \det \left[ w^\dagger(0; l) w(R;l') \right] \]  
(B41)

\[ P^{1}_{\alpha,\beta}(R; l, l') = \langle \phi_{0,l} | d^\dagger_\alpha d_\beta | \phi_{R;l'} \rangle = S(R; l, l') \left( w(R;l') \left( w(0; l) w(R;l') \right)^{-1} w(0; l) \right)_{\beta,\alpha} \]  
(B42)

\[ P^{2}_{(\alpha,\beta);(\gamma,\delta)}(R; l, l') = \langle \phi_{0,l} | d^\dagger_\alpha d^\dagger_\gamma d_\delta d_\beta | \phi_{R;l'} \rangle = S(R; l, l') (\delta_{\alpha\gamma} \delta_{\beta\delta} - \delta_{\alpha\delta} \delta_{\beta\gamma}) + P^{1}_{\alpha,\delta}(R; l, l') \delta_{\beta\gamma} + P^{1}_{\beta,\gamma}(R; l, l') \delta_{\alpha\delta} - P^{1}_{\beta,\delta}(R; l, l') \delta_{\alpha\gamma} - P^{1}_{\alpha,\gamma}(R; l, l') \delta_{\beta\delta} \]  
(B43)

\[ + \det(S(R; l, l')) \det \left( \begin{pmatrix} \delta_{\alpha\delta} & \delta_{\beta\delta} \\ \delta_{\alpha\gamma} & \delta_{\beta\gamma} \end{pmatrix} - \left[ w(R;l') \left( w(0; l) w(R;l') \right)^{-1} w(0; l) \right]_{(\delta,\gamma),(\alpha,\beta)} \right). \]  
(B44)

The effective model can be easily solved with a Bloch ansatz

\[ |\psi_k\rangle = \frac{1}{\sqrt{N_{VC}}} \sum_{l} f_l e^{ikR} |\phi_{R;l}\rangle \]  
(B45)

\[ \sum_{l'} \left[ S^{-1}(k) \tilde{H}(k) \right]_{l,l'} f_{l'} = E f_l \]  
(B46)

\[ \tilde{S}_{l,l'}(k) = \sum_{R} S_{l,l'}(R) e^{-ikR} \]  
(B47)

\[ \tilde{H}_{l,l'}(k) = \sum_{R} H_{l,l'}(R) e^{-ikR} \]  
(B48)