

Correlated insulators, semimetals, and superconductivity in twisted trilayer graphene

Maine Christos,¹ Subir Sachdev,¹ and Mathias S. Scheurer²

¹*Department of Physics, Harvard University, Cambridge MA 02138, USA*

²*Institute for Theoretical Physics, University of Innsbruck, Innsbruck A-6020, Austria*

Motivated by recent experiments indicating strong superconductivity and intricate correlated insulating and flavor-polarized physics in mirror-symmetric twisted trilayer graphene, we study the effects of interactions in this system close to the magic angle, using a combination of analytical and numerical methods. We identify asymptotically exact correlated many-body ground states at all integer filling fractions ν of the flat bands. To determine their fate when moving away from these fine-tuned points, we apply self-consistent Hartree-Fock numerics and analytic perturbation theory, with good agreement between the two approaches. This allows us to construct a phase diagram for the system as a function of ν and the displacement field, the crucial experimental tuning parameter of the system, and study the spectra of the different phases. The phase diagram is dominated by a correlated semimetallic intervalley coherent state and an insulating sublattice-polarized phase around charge neutrality, $\nu = 0$, with additional spin-polarization being present at quarter ($\nu = -2$) or three quarter ($\nu = +2$) fillings of the quasi-flat bands. We further study the superconducting instabilities emerging from these correlated states, both in the absence and in the additional presence of electron-phonon coupling, also taking into account possible Wess-Zumino-Witten terms. In the experimentally relevant regime, we find triplet pairing to dominate, possibly explaining the observed violation of the Pauli limit. Our results have several consequences for experiments as well as future theoretical work and illustrate the rich physics resulting from the interplay of almost flat bands and dispersive Dirac cones in twisted trilayer graphene.

CONTENTS

I. Introduction	2	D. Finite TBG bandwidth	19
A. Connection to experiment and phase diagram	3	E. Comparison of energetics	19
B. Organization of the paper	4	F. Breaking $SU(2)_+ \times SU(2)_-$	19
II. Model and symmetries	4	VI. Numerics for $\nu = \pm 2$	20
A. Continuum model and symmetries	4	A. Procedure and results	20
B. Adding a Displacement Field	6	B. Connection to experiment	20
C. Projected low-energy model	6	VII. Superconductivity	21
D. Interactions and form factors	7	A. Pairing in the presence of polarization	21
III. Exact groundstates at $D_0 = 0$	9	B. Pairing without polarization	21
A. Hamiltonian and construction of eigenstates	9	C. Relevance of Dirac cones and topology	22
B. Discussion of ground states	10	VIII. Conclusion and discussion	23
C. Resultant candidate states	10	Acknowledgments	25
IV. Hartree-Fock numerics	12	References	25
A. Hartree-Fock Method	12	A. Hamiltonian for the system	28
B. Band Structures	12	1. Continuum model and low-energy degrees of freedom	28
1. Self-Consistent Band Structures at $D_0 = 0$	13	2. Single-particle symmetries	29
2. Self-Consistent Band Structures for $D_0 > 0$	14	3. Interaction and form factors	31
C. Energies and phase diagram	14	B. Exact statements about the interacting groundstates	34
V. Analytical perturbation theory	14	1. Turning on the coupling between the subsystems	34
A. Ordering in the graphene-like bands	16	2. Behavior of Goldstone modes	36
B. Mixing between the bands	17	C. Hartree-Fock functional	37
C. Deforming the form factors	18	1. General form	37

2. Numerical procedure	38
a. How we fix the phases in numerics	38
b. Iterative procedure	39
c. Ground State dependence on grid size	39
3. Varying Parameters	40
D. Numerics at $\nu = 2$	41
1. Stable Phases at $\nu = 2$	41
2. Connection to Experimental Phase Diagram	42
E. Band Structure and solution for $D_0 \sim w_1$	42
F. Different contributions to the energy	44
1. Deformation of the form factors	45
2. Finite bandwidth of the TBG-like bands	46
a. First order perturbation theory	46
b. Second order perturbation theory: superexchange	46
3. Ordering in the graphene-like bands	47
4. Mixing between the TBG-like and graphene-like bands	51
5. Hund's Coupling	58
G. Superconducting instabilities	59
1. Fluctuation-induced superconductivity	59
2. Order parameter in the polarized phase	61
H. WZW terms in MSTG	62

I. INTRODUCTION

Graphene-based moiré superlattice systems have attracted considerable interest in the last few years, motivated by the strongly correlated physics they display [1–6]. While originally driven by the experimental realization of near-magic-angle twisted bilayer graphene (TBG) [7, 8], related strongly-correlated moiré superlattices have emerged, such as twisted double-bilayer graphene [9–12] and ABC-trilayer graphene on hexagonal boron nitride [13–15]. All of these systems exhibit low-energy bands which are energetically separated from the rest of the spectrum and can be tuned to be flat [16–18], enhancing the impact of correlations. The resulting correlated phenomena include interaction-induced insulating states [7, 9–11, 13], superconductivity [8–10, 14], and nematic order [5, 19–22], which are also integral parts of the phase diagrams of paradigmatic strongly correlated materials such as the cuprates. Furthermore, TBG has demonstrated that graphene moiré systems can also exhibit additional cascades of transitions that “reset” the band structure in an entire range of electron filling fractions [23, 24]. This is likely related to the polarization of certain combinations of the internal “flavor” quantum num-

bers of the electrons in the quasi-flat bands [25, 26], with interesting consequences for superconductivity [25, 27].

In addition to the twist angle between the layers as a tuning knob in TBG, the spectrum of twisted double-bilayer graphene can be efficiently tuned by applying a perpendicular electric displacement field D_0 . Notwithstanding the interesting consequences for the correlated physics in the system [9–12, 21, 28, 29], including the possibility of electrical control of the nematic director [22], its superconducting properties [9, 10] have been found to be more fragile than in TBG. Fortunately, mirror symmetric twisted trilayer graphene (MSTG), which consists of three layers of graphene with alternating relative twist angles, see Fig. 2(a), has very recently been realized experimentally [30–32] and combines the best of both worlds: it can be tuned significantly by applying a perpendicular displacement field D_0 while exhibiting strong and reproducible superconductivity. These experiments indicate that MSTG also exhibits both interaction-induced resistive states around integer filling fractions ν as well as the aforementioned spontaneous band resetting in an extended region of ν with $2 \lesssim |\nu| \lesssim 3$, akin to the cascades in TBG [23, 24]. As a function of D_0 and ν , the largest superconducting region is found to emerge out of this reconstructed normal state. Remarkably, Ref. 32 finds that the superconductor can sustain an in-plane magnetic field much larger than simple estimates of the Pauli limit. One additional crucial difference between MSTG and all of the other graphene moiré systems mentioned above is that MSTG not only exhibits quasi-flat bands but, at the same time, dispersive Dirac cones [30, 31, 33–38]. The interplay between these different types of bands and interactions likely gives rise to rich physics but also provides theoretical challenges.

So far, only very few theoretical studies of interactions in MSTG exist [39–44] and a systematic understanding of the nature and origin of possible particle-hole and superconducting instabilities as a function of ν and D_0 is still missing. The goal of this work is to help fill this gap by providing a detailed theoretical study of electron-electron interactions in the normal state of MSTG, allowing for a large class of possible instabilities, which is then used to analyze the order parameter and origin of the different superconducting phases of the system. In order to tackle the challenges associated with the simultaneous presence of flat and dispersive bands, we use a combination of analytical and numerical approaches.

More specifically, we start from the limit without displacement field, $D_0 = 0$, where the *non-interacting* band structure of the system is just that of TBG and single-layer graphene [33–38]; these two “subsystems” are, however, coupled by the Coulomb interaction. In the limit where the TBG bands are perfectly flat, we construct exact eigenstates of the interacting Hamiltonian of MSTG at all integer ν , which are also shown to be groundstates

for a finite range of the strength of the interaction between the two subsystems. This analysis shows that, in certain limits, also the *interacting* groundstates of MSTG are given by those of TBG and single-layer graphene. This is consistent with experiment [30, 31], observing a graphene Dirac cone at small D_0 .

To be able to address $D_0 \neq 0$ and more realistic system parameters, we use Hartree-Fock (HF) numerics and analytic perturbation theory. Our HF approach is motivated by the success of this approach in TBG [45–50] and the fact that the exact groundstates at $D_0 = 0$ are Slater-determinant states (in the TBG sector). In the analytic perturbation theory, we start from the exact groundstates and study the D_0 -induced deformations of the interaction matrix elements and band structure, the possible ordering in the graphene bands, the mixing between the graphene and TBG sectors for $D_0 \neq 0$, as well as the finite bandwidth of the TBG bands as perturbations. This complements our HF numerics, as it allows us to pinpoint the novel energetic contributions in MSTG as compared to TBG and serves as an important validation of our numerics—in particular, concerning the fate of the graphene Dirac cones, which provides a challenge to any numerical study as they are energetically degenerate with the TBG bands only in a small fraction of the moiré Brillouin zone (MBZ). In our analytics, we take advantage of this fraction being small and use it as an expansion parameter. Overall, we find good agreement between the HF numerics and the perturbation theory, both confirming the form of the exact eigenstates at $D_0 = 0$ (but with realistic system parameters); discrepancies between the approaches are traced back to the impact on the energetics coming from additional remote bands only taken into account in the numerics.

Having established the correlated nature and possible particle-hole instabilities in the normal state of MSTG as a function of D_0 and ν , we analyze the consequences for superconductivity. We study superconductivity both in the additional presence and absence of flavor polarization. Motivated by experiments [51–53] on TBG, which indicate that electron-phonon coupling is important for pairing, we follow Ref. 54 and assume that electron-phonon coupling stabilizes superconductivity but leaves singlet and triplet almost degenerate; informed by results for particle-hole instabilities in MSTG, we can then investigate which of the two will be favored due to additional particle-hole fluctuations. We also discuss purely electronic pairing and comment on the relevance of Dirac cones and associated Wess-Zumino-Witten (WZW) terms, as studied previously in TBG [25, 55].

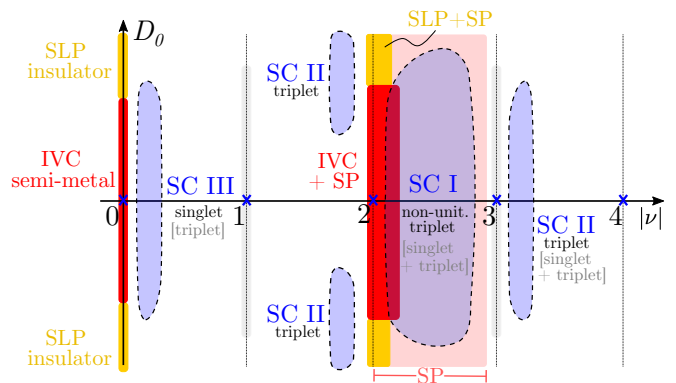


FIG. 1: Schematic phase diagram for MSTG based on our analysis, as a function of displacement field D_0 and filling fraction ν . Due to particle-hole symmetry of our model, the phase diagram only depends on $|\nu|$. We show the form of the dominant particle-hole instabilities, the intervalley coherent (IVC) semimetallic phase and a sublattice polarized (SLP) insulator at $\nu = 0$ (charge neutrality); the IVC and SLP state coexist with spin-polarization (SP) around $|\nu| = 2$. We discuss three different types of superconducting regimes, labelled SC I–III. We indicate in black their respective spin structure for the expected sign of the intervalley Hund’s coupling J_H resulting from Coulomb interactions alone [with opposite sign in gray].

A. Connection to experiment and phase diagram

To concisely summarize our main findings, we here briefly discuss their consequences for the phase diagram of MSTG, see Fig. 1, and their relation to experiment [30–32]. Based on our exact groundstates, perturbation theory, and HF numerics, the (for $\nu \neq 0$ slightly doped) semimetallic nature of the graphene sector is expected to be stable against interactions for $D_0 = 0$. At charge neutrality, $\nu = 0$, the leading instability is found to be an intervalley coherent state (IVC) in the TBG sector. While the two sectors start to mix when $D_0 \neq 0$, the semimetallic character of the bands is retained; the Landau fan sequence of this state, $\pm 2, \pm 6, \pm 10, \dots$, is consistent with that observed in Ref. 31. Once D_0 increases beyond a certain critical value D_0^c [which depends on the relaxation parameter w_0/w_1 in Eq. (2.3)], the IVC state transitions into a sublattice-polarized (SLP) phase, where all Dirac cones are gapped out, leading to an insulator. For realistic parameters, we expect D_0^c to be of the order of or potentially slightly larger than the range of the experimentally applied fields [30–32].

At $|\nu| = 2$, we find self-consistent HF solutions with spin polarization, which co-exists with additional particle-hole instabilities. At small D_0 , the leading additional symmetry-breaking state is again an IVC phase. Here, the TBG bands are completely spin polarized while the graphene cones’ polarization is only partial and approaches zero as $D_0 \rightarrow 0$. The spin polarization is

our proposed mechanism for the experimentally observed band reset and reduced flavor number [30, 31]. At larger D_0 , the IVC transitions into an SLP state. In Sec. VI B, we discuss how the bandstructures we find can give rise to the experimentally observed [30] additional sign change of the Hall density at sufficiently large D_0 . Our obtained behavior at $|\nu| = 0, 2$ also agrees with another experimental feature: the conductance is suppressed [30, 31] at $\nu = 0$ for both $D_0 = 0$ and $D_0 > 0$ while the suppression at $|\nu| = 2$ only sets in above a finite critical value of D_0 , which is consistent with the IVC (SLP) being semimetallic (insulating) at $\nu = 0$ and metallic (semimetallic) at $|\nu| = 2$.

As discussed in detail in Sec. V F, the relative spin-orientation of these orders in the two valleys depends on the sign of the intervalley Hund's coupling J_H (see Eq. (3.16)), which is not part of the Hamiltonian for our Hartree-Fock computations; the value or sign of J_H is not determined in our computations. For $J_H < 0$, which is expected for Coulomb interactions, both the SLP and SP have the same spin polarization in the two valleys, while the IVC order parameter does not carry any spin.

As indicated in blue in Fig. 1, we investigate three different regimes of superconducting phases. For superconductivity (SC I) in the presence of spin polarization with or without additional IVC order, the order parameter will be a non-unitary triplet for $J_H < 0$ and an admixed singlet-triplet phase [27] for $J_H > 0$. For superconducting phases (SC II) close to or in the range $2 \lesssim |\nu| \lesssim 3$ but not coexisting with additional particle-hole instabilities, we expect triplet (a singlet-triplet admixed phase) to dominate for $J_H < 0$ ($J_H > 0$). All of these superconducting states are consistent [27] with the observed [32] behavior in low to intermediate magnetic fields. Finally, our analysis indicates that a superconducting phase (if present) close charge neutrality (SC III) should be in a singlet state for $J_H < 0$.

B. Organization of the paper

The remainder of the paper is organized as follows. In Sec. II, we introduce the model for MSTG we study, its symmetries, and establish the basic notation used in this work. We begin our discussion of correlated physics in the limit of vanishing displacement field in Sec. III, where we construct exact groundstates. These build the starting point for our HF numerics, see Sec. IV for $\nu = 0$ and Sec. VI for $|\nu| = 2$, and our analytical perturbation theory, presented in Sec. V. Furthermore, superconducting instabilities will be analyzed in Sec. VII. Finally, Sec. VIII contains a short summary and discussion of results. Details of our analysis are provided in a set of appendices, labelled Appendix A through H.

II. MODEL AND SYMMETRIES

In this section, we introduce the interacting model we consider in this work, which consists of a continuum model to describe the moiré bands supplemented by Coulomb repulsion. We further discuss its symmetries and define the notation used in the remainder of the paper.

A. Continuum model and symmetries

MSTG is constructed from three parallel sheets of graphene where the top ($l = 1$) and bottom ($l = 3$) graphene layers are aligned with one another and the middle ($l = 2$) layer of graphene is twisted at a relative angle θ with respect to the top and bottom layers, as shown in Fig. 2(a). To compute the non-interacting band structure, we employ a continuum-model description, which is just a three-layer extension of the frequently applied continuum model for TBG [16–18]. Denoting the electronic creation operator for an electron at position $\mathbf{r} \in \mathbb{R}^2$, in sublattice $\rho = 1, 2$, layer $l = 1, 2, 3$, valley $\eta = \pm$, and of spin $s = \uparrow, \downarrow$ by $c_{\mathbf{r};\rho,l,\eta,s}^\dagger$ the Hamiltonian reads in the absence of a displacement field as [33–36]

$$H_{0,1}^{\text{Full}} = \int_{\mathbf{r}} c_{\mathbf{r};\rho,l,\eta,s}^\dagger \left(-iv_F (\boldsymbol{\rho}_{\theta_l}^\eta)_{\rho,\rho'} \cdot \nabla \right) c_{\mathbf{r};\rho',l,\eta,s} + \int_{\mathbf{r}} \left[c_{\mathbf{r};\rho,l,\eta,s}^\dagger (T_{\eta,\mathbf{r}})_{\rho,\rho'} c_{\mathbf{r};\rho',l+1,\eta,s} + \text{H.c.} \right], \quad (2.1)$$

where summation over repeated indices is implied, $\theta_l = (-1)^l \theta / 2$, $(\boldsymbol{\rho}_\theta^{\eta=+})_j = e^{i\theta\rho_z/2} \rho_j e^{-i\theta\rho_z/2}$, with ρ_j representing Pauli matrices in sublattice space (throughout we will use the same symbol for Pauli matrices and their indices, i.e., s_j, η_j are Pauli matrices in spin and valley space), and $\boldsymbol{\rho}_\theta^{\eta=-} = -(\boldsymbol{\rho}_\theta^{\eta=+})^*$. While the first term in Eq. (2.1) describes the Dirac cones of the two valleys, $\eta = \pm$, of each individual graphene layer in the continuum expansion around their respective K_g and K'_g points, the second line captures the tunneling between adjacent graphene layers parameterized by $T_{\eta,\mathbf{r}}$ (the direct hopping process between the outer two layers is neglected). As a consequence of the moiré superlattice, these tunneling matrix elements are modulated spatially. As is common, we focus on the lowest moiré-lattice harmonics for which symmetry allows for only two independent parameters, w_0 and w_1 , in $T_{\eta,\mathbf{r}}$. It can be written as [17]

$$T_{+,\mathbf{r}} = e^{-i\mathbf{q}_1 \mathbf{r}} \left[\mathcal{T}_1 + \mathcal{T}_2 e^{-i\mathbf{G}_1^M \mathbf{r}} + \mathcal{T}_3 e^{-i(\mathbf{G}_1^M + \mathbf{G}_2^M) \mathbf{r}} \right] \quad (2.2)$$

and $T_{-,\mathbf{r}} = (T_{+,\mathbf{r}})^*$. Here $\mathbf{q}_1 = k_\theta(0, -1)$, $k_\theta = 2|K_g| \sin(\theta/2)$, connects the K and K' points of the moiré lattice, $\mathbf{G}_1^M = -\sqrt{3}k_\theta(1, \sqrt{3})^T/2$ and $\mathbf{G}_2^M = \sqrt{3}k_\theta(1, 0)^T$

are the basis vectors of the reciprocal lattice (RL) of the moiré lattice, $\text{RL} := \{\sum_{j=1,2} n_j \mathbf{G}_j^M, n_j \in \mathbb{Z}\}$. The matrices \mathcal{T}_j in sublattice space in Eq. (2.2) are given by

$$\mathcal{T}_j = w_0 \rho_0 + w_1 \begin{pmatrix} 0 & \omega^{j-1} \\ \omega^{-(j-1)} & 0 \end{pmatrix}, \quad \omega = e^{-i\frac{2\pi}{3}}. \quad (2.3)$$

While rigidly rotated graphene layers correspond to $w_0 = w_1$, lattice relaxation [56, 57] leads to $w_0 \neq w_1$. We will here consider w_0 as a free parameter and study the physics as a function of it.

As a result of the moiré modulations in Eq. (2.2), the tunneling matrix elements between the layers couple momenta related by RL vectors, which reconstructs the graphene cones of the first line in Eq. (2.1), leading to a (technically infinite) set of bands; the band energies will be denoted by $\epsilon_{n,\eta}(\mathbf{k})$, $n \in \mathbb{Z}$, where \mathbf{k} is in the first MBZ and we already used that Eq. (2.1) is diagonal in the valley, η , and trivial in the spin index, s ; this makes the band energies independent of spin and allows to label them by their valley index.

MSTG is distinguished from other moiré systems, such as TBG or twisted double-bilayer graphene, by a reflection symmetry σ_h under exchange of the top and bottom layers of MSTG, with action $\sigma_h : c_{\mathbf{r};\rho,l,\eta,s} \rightarrow c_{\mathbf{r};\rho,\sigma_h(l),\eta,s}$ where $\sigma_h(2) = 2$, $\sigma_h(1) = 3$, and $\sigma_h(3) = 1$. Since it acts trivially in all internal indices (ρ, η, s) and \mathbf{r} , the Hamiltonian (2.1) can be decomposed into sectors with different mirror eigenvalues, $\sigma_h = \pm 1$, by performing a unitary transformation, $V \in \text{U}(3)$, in layer space only, $c_{\mathbf{r};\rho,l,\eta,s} = V_{l,\ell} \psi_{\mathbf{r};\rho,\ell,\eta,s}$. As was pointed out before [33] and detailed in Appendix A 1, the Hamiltonian in the mirror-even sector ($\sigma_h = +1$, $\ell = 1, 2$) turns out to be that of TBG with an interlayer hopping renormalized by a factor of $\sqrt{2}$. In the mirror-odd sector ($\sigma_h = -1$, $\ell = 3$) the Hamiltonian is that of single-layer graphene (without any moiré modulation). This can be clearly seen in the spectrum shown in Fig. 2(b), which exhibits two almost-flat TBG-bands (shown in red) per spin and valley and unreconstructed Dirac cones at K (K') associated with the graphene sector of valley $\eta = +$ ($\eta = -$). As such, there are three Dirac cones at K (and three at K'), one belonging to the graphene sector of MSTG in a single valley only and two Dirac crossings belonging to the TBG sector of both valleys. As can be seen, there are additional remote bands of both subspaces due to back-folding into the MBZ.

In addition to reflection symmetry, the Hamiltonian in Eq. (2.1) has more point symmetries, forming the group C_{6h} : the model is invariant under three-fold rotational symmetry C_{3z} , with action, $C_{3z} : c_{\mathbf{r}} \rightarrow e^{i\frac{2\pi}{3}\rho_z \eta_z} c_{C_{3z}\mathbf{r}}$, which is also an exact symmetry of the moiré lattice in Fig. 2(a). While not an exact lattice symmetry, two-fold rotation perpendicular to the plane, C_{2z} , is a good approximate symmetry for small twist angles; in fact, it is an exact symmetry of the continuum in Eq. (2.1)

which can be verified by applying its action, $C_{2z} : c_{\mathbf{r}} \rightarrow \eta_x \rho_x c_{-\mathbf{r}}$.

Besides these point symmetries, the model also exhibits the following exact internal symmetries: since there is no coupling between the valleys, it is invariant under a valley U(1) transformation, $\text{U}(1)_v : c_{\mathbf{r}} \rightarrow e^{i\eta_z \varphi} c_{\mathbf{r}}$. In combination with the absence of spin-orbit coupling, it is further invariant under the separate spin-rotation in each valley, forming the group $\text{SU}(2)_+ \times \text{SU}(2)_-$ with $\text{SU}(2)_{\pm}$ acting as $c_{\mathbf{r}} \rightarrow e^{i\boldsymbol{\varphi} \cdot \mathbf{s} (\eta_0 \pm \eta_z)/2} c_{\mathbf{r}}$. Furthermore, there is time-reversal symmetry, which is associated with the anti-unitary operator Θ with $\Theta c_{\mathbf{r}} \Theta^\dagger = \eta_x c_{\mathbf{r}}$; unless stated otherwise, we will always refer to this form of (spinless) time-reversal symmetry throughout the text, although combinations with spin-rotations (spinful time-reversal), Θ_s , and with $\text{U}(1)_v$ rotations, $\tilde{\Theta}$, will play a role further below. Since the mirror symmetry protects any mixing between the graphene and TBG subspaces and $\text{U}(1)_v$ any mixing between the two valleys, the aforementioned Dirac crossings at K and K' are protected by the combination $C_{2z}\tilde{\Theta}$, exactly as in graphene and TBG.

Focusing on specific limits, there are also additional internal symmetries, similar to TBG [46, 58, 59]: when $w_0 = 0$, $H_{0,1}^{\text{Full}}$ in Eq. (2.1) changes its sign, $H_{0,1}^{\text{Full}} \rightarrow -H_{0,1}^{\text{Full}}$, if we apply the chiral symmetry operator $C : c_{\mathbf{r}} \rightarrow \rho_z L_C c_{\mathbf{r}}$ where L_C is a matrix in layer space given by $L_C = V \text{diag}(1, 1, -1) V^\dagger$. Therefore, $w_0 = 0$ will be referred to as the *chiral limit* [58]. Among other consequences for the interaction terms to be discussed below, the band structure has to be symmetric about zero energy at every given momentum \mathbf{k} when $w_0 = 0$, i.e., $\text{spec}_{\mathbf{k},\eta} = -\text{spec}_{\mathbf{k},\eta}$, with $\text{spec}_{\mathbf{k},\eta} = \{\epsilon_{n,\eta}(\mathbf{k}), n \in \mathbb{Z}\}$.

Finally, when the additional rotation of the sublattice matrices in the first line of Eq. (2.1) is neglected, $\rho_{\theta_l}^\eta \rightarrow \rho_{\theta=0}^\eta$, the ‘‘unitary particle-hole symmetry’’, previously discussed in TBG [46, 59], can be extended to MSTG: defining the unitary operator P which acts as $P \psi_{\mathbf{r};l=1} P^\dagger = \eta_z \psi_{-\mathbf{r};l=2}$, $P \psi_{\mathbf{r};l=2} P^\dagger = -\eta_z \psi_{-\mathbf{r};l=1}$, and $P \psi_{\mathbf{r};l=3} P^\dagger = i \rho_y \eta_y \psi_{-\mathbf{r};l=3}$, one finds $P H_{0,1}^{\text{Full}} P^\dagger = -H_{0,1}^{\text{Full}}$, showing that the spectrum must obey $\text{spec}_{\mathbf{k}} = -\text{spec}_{-\mathbf{k}}$ where $\text{spec}_{\mathbf{k}} = \cup_\eta \text{spec}_{\mathbf{k},\eta}$. Since we will be focusing on small twist angles θ below, we will always assume that $\rho_{\theta_l}^\eta$ has been replaced by $\rho_{\theta=0}^\eta$ in Eq. (2.1) when analyzing symmetries. We will also later set $\theta = 0$ in our numerics.

For the detailed form of the model and the action of the symmetries in momentum space we refer the reader to Appendix A 1 and A 2, respectively. A brief list of the symmetries of H_0 and when they apply can be found in Table I.

The primary interaction term we shall consider, in Eq. (2.8), will preserve the symmetries of the continuum model. However, we will also consider the consequences of a Hund’s coupling in Eq. (3.16), which will break the $\text{SU}(2)_+ \times \text{SU}(2)_-$ symmetry down to a more

realistic $SU(2)_s$ (the group of simultaneous spin rotations in the two valleys).

B. Adding a Displacement Field

Motivated by recent experimental discoveries of electric-field-tunable correlated physics in MSTG [30–32], we now extend the model for MSTG in Eq. (2.1) and include a perpendicular displacement field D_0 . Finite D_0 results in a potential difference between the top and bottom layers of graphene. Suppressing all indices except layer l , a displacement field is represented in the continuum model as

$$H_{0,2}^{\text{Full}} = D_0 \sum_{\mathbf{r}} \left[c_{\mathbf{r};l=1}^\dagger c_{\mathbf{r};l=1} - c_{\mathbf{r};l=3}^\dagger c_{\mathbf{r};l=3} \right] \quad (2.4)$$

and the full Hamiltonian becomes $H_0^{\text{Full}} = H_{0,1}^{\text{Full}} + H_{0,2}^{\text{Full}}$. The displacement field D_0 breaks the mirror symmetry σ_h , resulting in hybridization between bands in the TBG and graphene sectors; this hybridization can only take place between graphene and TBG bands of the same valley and spin flavor, as $U_v(1)$ and spin-rotation symmetry [in fact the full $SU(2)_+ \times SU(2)_-$] are preserved by $H_{0,2}^{\text{Full}}$.

Furthermore, when D_0 is finite, the system retains C_{2z} , C_{3z} , leading to the point group C_6 . Time-reversal symmetry, Θ , and, in the limit $w_0 = 0$, chiral symmetry, C , also persist. However, the unitary particle hole symmetry P is no longer an exact symmetry for finite D_0 .

At the K (K') point of the MBZ, the Dirac crossings which belong to the valley $\eta = +$ ($\eta = -$) will hybridize as D_0 increases with the graphene bands of the same valley which also exhibit a Dirac cone at the Fermi level; this results in the Dirac points, which were pinned at the Fermi level at $D_0 = 0$, being pushed away from the Fermi level. As $C_{2z}\Theta$ symmetry is preserved for $D_0 \neq 0$, these Dirac crossings cannot be gapped out. As a consequence of the absence of a graphene cone to hybridize with, the Dirac point of TBG in the other valley, $\eta = -$ ($\eta = +$), remains pinned at the Fermi level as D_0 increases. These features can be clearly observed in the band structure of MSTG with finite displacement field shown in Fig. 2(b).

C. Projected low-energy model

In order to make the analytical and numerical study of interactions feasible in the system, we restrict our Hamiltonian to a finite set of bands in the vicinity of the Fermi level. Going forward, we will denote the two bands closest to the Fermi level per spin and valley as *TBG-like bands* and the next two closest bands per spin and valley as *graphene-like bands*. All other bands at energies further away from the Fermi level will be referred to as *remote bands*. When $D_0 = 0$, the TBG-like bands are

labeled by mirror symmetry eigenvalue $\sigma_h = +1$ and can be identified exactly with the two flat bands of TBG. The graphene-like bands at $D_0 = 0$ are identified exactly with the bands of the continuum model of graphene in the vicinity of the K and K' points. For finite D_0 , the bands with different eigenvalues hybridize but we will retain our naming convention for the four bands closest to the Fermi level, i.e., the bands indicated in red in Fig. 2 are TBG-like bands.

Note that away from the K points, the graphene-like bands are technically identified with the first remote bands of TBG with $\sigma_h = +1$ for $D_0 = 0$ within this convention. In our analytical calculations below, however, we will restrict the analysis to the TBG-like bands in the full MBZ and to the graphene-like bands only in the vicinity of the K and K' points where they retain their graphene-like identity. Denoting the associated creation operators in these two sets of bands by $b_{\mathbf{k};p,\eta,s}^\dagger$ and $g_{\mathbf{k};p,\eta,s}^\dagger$, where $\mathbf{k} \in \text{MBZ}$ is the momentum, $p = +$ ($p = -$) labels the upper (lower) band in each sector, η is the valley, and s the spin quantum number of the electrons, the non-interacting Hamiltonian projected to these bands becomes

$$H_0 = \sum_{\mathbf{k} \in \text{MBZ}} W_{\text{TBG}} \epsilon_{(b,p),\eta}(\mathbf{k}) b_{\mathbf{k};p,\eta,s}^\dagger b_{\mathbf{k};p,\eta,s} + \sum_{\mathbf{k} \in \text{MBZ}'} \epsilon_{(g,p),\eta}(\mathbf{k}) g_{\mathbf{k};p,\eta,s}^\dagger g_{\mathbf{k};p,\eta,s}. \quad (2.5)$$

Here the extra prime in MBZ' in the second line indicates that we restrict the graphene-like degrees of freedom to the vicinity of the K and K' points in our analytical calculations. Note that we further introduced the dimensionless parameter $W_{\text{TBG}} \in [0, 1]$ that will allow us to organize the perturbation theory of Sec. V D in the TBG-like bandwidth. While the physical system corresponds to $W_{\text{TBG}} = 1$, we define the *flat limit* as $W_{\text{TBG}} = 0$, which will play an important role below. Note that the Hamiltonian in Eq. (2.1) realizes the flat limit exactly, $\epsilon_{(b,p),\eta}(\mathbf{k}) = 0$, only when $w_0 = 0$, θ is at the magic angle, and when $D_0 = 0$.

In our numerics, we will keep all momenta of both sets of (in total 16) bands in Eq. (2.5) and continue to use $g_{\mathbf{k};p,\eta,s}^\dagger$ as creation operators for the second lowest set of bands throughout the MBZ. To check convergence, we will also study the impact of adding additional remote bands, as described in Appendix C 3.

In order to fix the phases of the wavefunctions of the TBG-like and graphene-like bands, we specify how they transform under the action of the various discrete symmetries discussed above or, equivalently how they act on the electronic operators $b_{\mathbf{k};p,\eta,s}^\dagger$ and $g_{\mathbf{k};p,\eta,s}^\dagger$. While a thorough discussion can be found in Appendix A 2, these representations are summarized in Table I.

Based on these symmetry representations, one finds

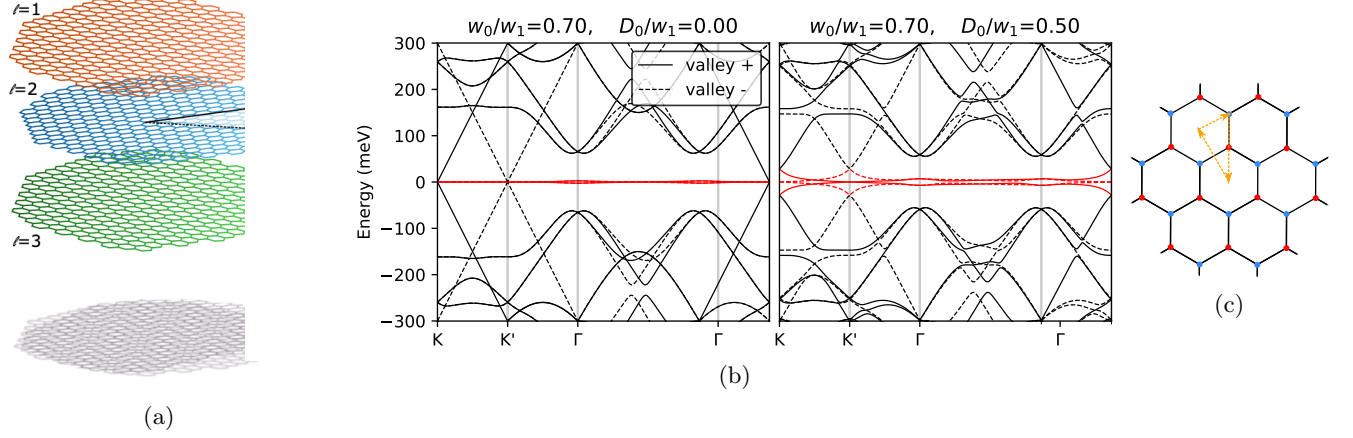


FIG. 2: (a) Schematic of the geometry of MSTG. The top and bottom layers shown in red and green respectively are aligned, while the middle layer (blue) is twisted at relative angle θ to the top and bottom layers. (b) Band structure of MSTG at $D_0 = 0$ (left) and $D_0 > 0$ (right) for $w_1 = 124$ meV and twist angle $\theta = 1.53^\circ$ along the one-dimensional cut through the MBZ shown in (c). We represent the spin-degenerate bands in the two different valleys $\eta = \pm$ with solid ($\eta = +$) and dotted ($\eta = -$) lines. At $D_0 = 0$, the band structure is given by that of TBG and single-layer graphene and we have colored the (two per spin and valley) quasi-flat TBG bands red. For $D_0 > 0$, the bands of these two subsystems mix but we continue to label the two bands per spin and valley closest to the Fermi as “TBG-like” bands, indicated in red.

TABLE I: Summary of discrete single-particle symmetries (denoted S) of our model and when they apply (last column). For convenience of the reader and future reference, we show a redundant set of symmetries. We also indicate whether they are unitary or anti-unitary symmetries (second column) and whether they commute + or anti-commute – (third column) with the non-interacting Hamiltonian H_0^{Full} . Finally, we specify their action on the low-energy field operators in order to fix the phase of the Bloch states, using σ_j to denote Pauli matrices acting on the upper/lower bands subspace.

S	unitary	$[\cdot, \cdot]_{\pm}$	$Sb_{\mathbf{k}}S^\dagger$	$Sg_{\mathbf{k}}S^\dagger$	condition
C_{2z}	✓	+	$\eta_1 \sigma_0 b_{-\mathbf{k}}$	$\eta_1 \sigma_0 g_{-\mathbf{k}}$	—
C_{3z}	✓	+	$\eta_0 \sigma_0 b_{C_{3z}\mathbf{k}}$	$\eta_0 \sigma_0 g_{C_{3z}\mathbf{k}}$	—
σ_h	✓	+	$\eta_0 \sigma_0 b_{\mathbf{k}}$	$-\eta_0 \sigma_0 g_{\mathbf{k}}$	$D_0 = 0$
Θ	✗	+	$\eta_1 \sigma_0 b_{-\mathbf{k}}$	$\eta_1 \sigma_0 g_{-\mathbf{k}}$	—
Θ_s	✗	+	$s_2 \eta_1 \sigma_0 b_{-\mathbf{k}}$	$s_2 \eta_1 \sigma_0 g_{-\mathbf{k}}$	—
$\tilde{\Theta}$	✗	+	$\eta_2 \sigma_0 b_{-\mathbf{k}}$	$\eta_2 \sigma_0 g_{-\mathbf{k}}$	—
C	✓	–	$\eta_3 \sigma_2 b_{\mathbf{k}}$	$\eta_3 \sigma_2 g_{\mathbf{k}}$	$w_0 = 0$
P	✓	–	$-i\eta_3 \sigma_2 b_{-\mathbf{k}}$	$i\eta_1 \sigma_2 g_{-\mathbf{k}}$	$D_0 = 0, \rho_{\theta \rightarrow 0}$
$C_{2z}\Theta$	✗	+	$\eta_0 \sigma_0 b_{\mathbf{k}}$	$\eta_0 \sigma_0 g_{\mathbf{k}}$	—
$C_{2z}P$	✓	–	$\eta_2 \sigma_2 b_{\mathbf{k}}$	$i\sigma_2 g_{\mathbf{k}}$	$D_0 = 0, \rho_{\theta \rightarrow 0}$

that the dispersions are constrained to have the form

$$\begin{aligned} \epsilon_{(b,p),\eta}(\mathbf{k}) &= p \bar{E}_0^b(\mathbf{k}) + \eta \bar{E}_1^b(\mathbf{k}) w_0 \\ &+ p \eta \bar{E}_2^b(\mathbf{k}) D_0 + \bar{E}_3^b(\mathbf{k}) w_0 D_0 \end{aligned} \quad (2.6)$$

and

$$\epsilon_{(g,p),\eta}(\mathbf{k}) = p \bar{E}_0^g(\mathbf{k}; \eta) + \bar{E}_1^g(\mathbf{k}; \eta) w_0 D_0. \quad (2.7)$$

To illustrate the changes of the form of the band structures when symmetries are broken by $w_0 \neq 0$ and $D_0 \neq 0$, we introduced $\bar{E}_j^b(\mathbf{k})$ and $\bar{E}_j^g(\mathbf{k}; \eta)$ which are functions of w_0 and D_0 that stay finite when $w_0, D_0 \rightarrow 0$.

As can be seen, the two terms in the first row of Eq. (2.6) just correspond to the TBG band structure (with, $w_0 = 0$, or without, $w_0 \neq 0$, chiral symmetry). The second line shows that the hybridization with the graphene bands for $D_0 \neq 0$ distorts the TBG band structure in a way not present in TBG, which is related to the D_0 -induced breaking of $C_{2z}P$ symmetry. One of these additional terms leads to superexchange processes in the TBG-like bands that are not present in TBG and will be discussed in Sec. VD. In Eq. (2.7), the first term, $\bar{E}_0^g(\mathbf{k}; \eta)$, simply becomes the graphene Dirac dispersion for $D_0 \rightarrow 0$. We further see that the property $\epsilon_{(g,+),\eta}(\mathbf{k}) = -\epsilon_{(g,-),\eta}(\mathbf{k})$ is only violated if w_0 and D_0 are simultaneously non-zero. This is expected, as the graphene-like bands can only “feel” the broken chiral symmetry if they hybridize with the TBG-like bands. For the parameters in Fig. 2(b), this is a rather weak effect around the Dirac cones.

D. Interactions and form factors

As our goal is to study the interacting behavior of MSTG, we next add interaction terms to the Hamiltonian. We assume that these interactions in the full con-

tinuum model are of the density-density form. Our full interaction Hamiltonian thus reads as

$$H_1^{\text{Full}} = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) : \rho_{\mathbf{q}}^{\text{Full}} \rho_{-\mathbf{q}}^{\text{Full}} : \quad (2.8)$$

where $: \dots :$ denotes normal ordering, N is the number of moiré sites, and $\rho_{\mathbf{q}}^{\text{Full}}$ is the density operator of momentum \mathbf{q} of the continuum electron operators $c_{\mathbf{r}}$ in Eq. (2.1). Note $\mathbf{q} \in \mathbb{R}$ in the sum in Eq. (2.8) which is *not* restricted to the MBZ. For our analytical discussions below, we will not have to specify the explicit form of $V(\mathbf{q})$ but only use $V(\mathbf{q}) > 0$ and $V(\mathbf{q}) = V(-\mathbf{q})$; consequently, our analytical results will be valid regardless of the details of the screening processes at high energies and of nearby gates and/or substrates, that affect the detailed form $V(\mathbf{q})$. For our numerics, we will use the single-gate-screened Coulomb potential:

$$V(\mathbf{q}) = \frac{1 - e^{-2|\mathbf{q}|d_s}}{2A_{\text{moiré}}\epsilon_0\epsilon|\mathbf{q}|}, \quad (2.9)$$

where we have normalized by the real space area of the moiré unit cell $A_{\text{moiré}}$. We vary the screening length d_s in our numerical calculations from $d_s = 10$ to 80 nm and find relatively little dependence of the relative energies between phases and no qualitative change in the ground states. We will therefore take $d_s = 40$ nm as our default value. We also vary the dielectric constant ϵ from $\epsilon = 4$ to $\epsilon = 15$. We find the phase boundaries depend on ϵ , though no new phases emerge as ϵ is varied [the effect of varying interaction parameters is further discussed in Appendix C3, especially Fig. 9]. Unless otherwise specified, we will take $\epsilon = 7$.

We note that the above interaction term has all the symmetries of the continuum model introduced in Sec. II, including the $SU(2)_+ \times SU(2)_-$ symmetry. In the realistic system, there is a Hund's coupling, $J_H \neq 0$, which breaks $SU(2)_+ \times SU(2)_-$ down to $SU(2)_s$. We discuss the form of the Hund's coupling in Appendix F5 and Eq. (3.16), and will note its consequences in Sec. III C.

Neglecting normal ordering and replacing $\rho_{\mathbf{q}}^{\text{Full}} \rightarrow \tilde{\rho}_{\mathbf{q}}^{\text{Full}} = \rho_{\mathbf{q}}^{\text{Full}} - \text{const.} \times \delta_{\mathbf{q},0}$ in Eq. (2.8) leads to a new form of the interaction, $\tilde{H}_1^{\text{Full}}$, which, however, can be rewritten as H_1^{Full} by a redefinition of the chemical potential (and energy reference point). This is not the case anymore after projecting $\rho_{\mathbf{q}}^{\text{Full}}$ and $\tilde{\rho}_{\mathbf{q}}^{\text{Full}}$ in H_1^{Full} and $\tilde{H}_1^{\text{Full}}$, respectively, to a finite set of bands of H_0 . As described in more detail in Appendix A3, we follow [46, 59] and rewrite the interaction before projection such that it exhibits particle-hole symmetry with respect to the charge neutrality point before and after projection. Denoting the electronic creation operators for band n , of valley η , spin s , and with momentum $\mathbf{k} \in \text{MBZ}$ by $f_{\mathbf{k};n,\eta,s}^\dagger$, the projected interaction becomes

$$\tilde{H}_1 = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta\rho_{\mathbf{q}} \delta\rho_{-\mathbf{q}}, \quad (2.10a)$$

with the symmetrized density operators

$$\delta\rho_{\mathbf{q}} = \sum_{\mathbf{k} \in \text{MBZ}} \left[f_{\text{MBZ}(\mathbf{k}+\mathbf{q})}^\dagger F_{\mathbf{k},\mathbf{q}} f_{\mathbf{k}} - \frac{1}{2} \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} \text{tr}(F_{\mathbf{k},\mathbf{G}}) \right]. \quad (2.10b)$$

Here $\text{MBZ}(\mathbf{k}) := \mathbf{k} - \mathbf{G}_{\mathbf{k}} \in \text{MBZ}$ for the unique reciprocal lattice vector $\mathbf{G}_{\mathbf{k}} \in \text{RL}$. We suppressed all indices of $f_{\mathbf{k}}$ and $f_{\mathbf{k}}^\dagger$, which should be viewed as column and row vectors in band, valley, and spin space and introduced the matrix-valued *form factors* $F_{\mathbf{k},\mathbf{q}}$, which contain all the microscopic details of the wavefunctions of the bands of H_0 [see Eq. (A25) for a formal definition]. Note that Eq. (2.10) holds for any subset of bands that we want to keep. As already discussed in Sec. II C above, we will only keep the TBG-like bands and graphene-like bands (around the K/K' points) in the analytics, while we will allow for more bands in the HF numerics.

We refer the interested reader to Appendix A3, where a detailed discussion of constraints on $F_{\mathbf{k},\mathbf{q}}$ resulting from symmetries, Hermiticity, and the structure of H_0 can be found, and here only state a few properties of the form factors that we will explicitly refer back to in the main text. First, as a consequence of $U(1)_v$ and spin rotation symmetry, the form factors can only have non-diagonal matrix structure in band space,

$$(F_{\mathbf{k},\mathbf{q}})_{(n,\eta,s),(n',\eta',s')} = \delta_{s,s'} \delta_{\eta,\eta'} \left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{n,n'}. \quad (2.11)$$

In accordance with our notation in Eq. (2.5), we will use the multi-index notation $n = (t,p)$ where the ‘‘type’’ t distinguishes between the graphene-like, $t = g$, and TBG-like, $t = b$, bands and $p = +$ ($p = -$) labels the upper (lower) band; for instance, for the electronic operators, it holds $f_{\mathbf{k};(b,p),\eta,s} = b_{\mathbf{k};p,\eta,s}$ and $f_{\mathbf{k};(g,p),\eta,s} = g_{\mathbf{k};p,\eta,s}$. As a consequence of the σ_h symmetry for $D_0 = 0$, the form factors become block-diagonal in the sectors with different σ_h eigenvalue. As such, $F_{\mathbf{k},\mathbf{q}}^{tt'}$ defined via

$$\left(F_{\mathbf{k},\mathbf{q}}^{tt'} \right)_{(p,\eta,s),(p',\eta',s')} = \delta_{s,s'} \delta_{\eta,\eta'} \left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{(t,p)(t',p')} \quad (2.12)$$

obeys

$$F_{\mathbf{k},\mathbf{q}}^{tt'} \propto \delta_{t,t'}, \quad \text{for } D_0 = 0. \quad (2.13)$$

Consequently, the entire interacting Hamiltonian $H_0 + \tilde{H}_1$ preserves the charge in the graphene and TBG system separately. This leads to simplifications, which will be exploited in Sec. III below. For $D_0 \neq 0$, this is not the case anymore, as the form factors in Eq. (2.12) and, hence, the density operators in the interaction will *not* be diagonal in the t index and scatter electrons between the two different types of bands.

In the limit $D_0 = 0$, one can also compute the form factors in the graphene sector analytically [see Eq. (A39) for

the full expression]. Most importantly for our purposes here, one finds that

$$F_{\mathbf{k},\mathbf{G}}^{gg} = \delta_{\mathbf{G},0} \sigma_0 \eta_0 s_0, \quad \mathbf{G} \in \text{RL}, \quad \text{for } D_0 = 0, \quad (2.14)$$

where σ_0 are Pauli matrices in the ‘‘band space’’ (with indices $p = \pm$). Here only the $\mathbf{G} = 0$ component is finite, which is related to the fact that the graphene bands do not ‘‘feel’’ the moiré superlattice for $D_0 = 0$.

III. EXACT GROUNDSTATES AT $D_0 = 0$

We begin our discussion of the interacting physics in the *decoupled limit*, defined as $D_0 = 0$, where all bands can be labeled by their mirror eigenvalue $\sigma_h = \pm 1$, and the low-energy bands are those of TBG and single-layer graphene. While one might intuitively expect that the presence of the additional graphene Dirac cones, which have a much lower density of states than the (almost) flat bands close to the magic angle of TBG, is not strong enough to change the symmetry of the correlated insulating phase of TBG, it is *a priori* not clear whether the density-density coupling between the two subsystems can also induce the same symmetry-breaking order (and potentially gap out) the graphene Dirac cones. It is further not clear whether *exact* interacting ground states of MSTG can be identified in certain limits, similar to TBG [46, 60, 61]. These aspects will be addressed in this section.

A. Hamiltonian and construction of eigenstates

To this end, let us focus on the flat limit, $W_{\text{TBG}} = 0$, and postpone the perturbative treatment of the finite TBG bandwidth to Sec. VD. From our discussion in Sec. II, we can read off that the low-energy Hamiltonian of MSTG in the flat-decoupled limit is given by

$$H_{\text{FD}} = H^g + H^b + \lambda H^{\text{gb}}, \quad (3.1)$$

consisting of three terms: H^g is the Hamiltonian of (both valleys of) single-layer graphene with Coulomb repulsion [see Eq. (B4)], H^b the interacting Hamiltonian of TBG in the flat limit, given by

$$H^b = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta \rho_{\mathbf{q}}^b \delta \rho_{-\mathbf{q}}^b \quad (3.2)$$

where we defined the projected subsystem density operators

$$\delta \rho_{\mathbf{q}}^t = \sum_{\mathbf{k} \in \text{MBZ}} \left[t_{\text{MBZ}(\mathbf{k}+\mathbf{q})}^\dagger (F_{\mathbf{k},\mathbf{q}}^{tt}) t_{\mathbf{k}} - \frac{1}{2} \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} \text{tr}(F_{\mathbf{k},\mathbf{G}}^{tt}) \right] \quad (3.3)$$

where $t = g, b$. Finally, the last term in Eq. (3.1) describes the coupling between the two subsystems via a density-density interaction,

$$\lambda H^{\text{gb}} = \frac{\lambda}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta \rho_{-\mathbf{q}}^g \delta \rho_{\mathbf{q}}^b, \quad (3.4)$$

where the additional prefactor $\lambda \in [0, 1]$ has been introduced to adiabatically turn on the H^{gb} interaction (the physical system corresponds to $\lambda = 1$).

For $\lambda = 0$, the Hamiltonian is just the sum $H^g + H^b$ of the two commuting subsystem Hamiltonians. So its eigenstates are just given by (all combinations) of the individual eigenstates of graphene and TBG with the correct particle number: let us fix a certain integer filling $\nu_b = 0, \pm 1, \pm 2, \pm 3, \pm 4$ of the TBG system. Then the graphene system will be at a corresponding filling, which we write formally as $\nu_g = \nu_g(\nu_b)$ and its ground state will be a semimetal with ($\nu_g = 0$) or without ($\nu_g \neq 0$) doping, that exhibits correlations but does not spontaneously break any symmetries; these properties are well established (theoretically and experimentally) for graphene [62]. Let us denote the ground-state of the graphene system at filling $\nu_g = \nu_g(\nu_b)$ by $|\Psi_0^g(\nu_b)\rangle$ and its (gapless) excited states with the same particle number by $|\Psi_j^g(\nu_b)\rangle$, $j > 0$. In the hypothetical absence of any correlations in the graphene subspace, $|\Psi_0^g(\nu_b)\rangle$ ($|\Psi_j^g(\nu_b)\rangle$) would just be Slater-determinant state(s) with the Dirac cones filled up to to the chemical potential (and some additional particle-hole excitations).

Exact ground states of the flat-band TBG Hamiltonian in Eq. (3.2) have been discussed previously [60], which we will very briefly review here using our notation, in order to set the stage for the extension to MSTG. Upon defining new operators according to [referred to as ‘‘chiral basis’’ in [59]; in this basis the chiral form factors, see Eq. (B7), are diagonal]

$$\tilde{b}_{\mathbf{k};c,\eta,s} = U_{c,p} b_{\mathbf{k};p,\eta,s}, \quad U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}, \quad (3.5)$$

consider the set of states

$$|\Psi_0^b(\nu_b)\rangle = \prod_{\mathbf{k} \in \text{MBZ}} \prod_{c=\pm} \prod_{j_c=1}^{\nu_c} \tilde{b}_{\mathbf{k};c,\eta_{j_c}^c,s_{j_c}^c}^\dagger |0^b\rangle \quad (3.6)$$

with arbitrary combinations of occupied flavors $\{\eta_j^\pm, s_j^\pm\}$ such that $\nu_+ + \nu_- = 4 + \nu_b$. It was shown in [60] that

$$\delta \rho_{\mathbf{q}}^b |\Psi_0^b(\nu_b)\rangle = \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} R_{\mathbf{G}} |\Psi_0^b(\nu_b)\rangle, \quad (3.7)$$

with $R_{\mathbf{G}} = \nu_b \sum_{\mathbf{k}} \text{tr}[F^{bb}(\mathbf{k},\mathbf{G})]/8$ when $w_0 = 0$, i.e., in the chiral limit. As such, all of these states are exact eigenstates of H^b in Eq. (3.2). Ref. 60 further showed that these states will always be ground states of H^b for

$\nu_b = 0$; the same holds for all other integer ν_b as long as the flat-metric condition,

$$F_{\mathbf{k},\mathbf{G}}^{bb} = \mathbb{1}f(\mathbf{G}), \quad \forall \mathbf{k}, \mathbf{G}, \quad (3.8)$$

is not violated by a significantly large amount. Furthermore, when turning on $w_0 \neq 0$, the subset of states in Eq. (3.6) with $\eta_j^+ = \eta_j^-$ and $s_j^+ = s_j^-$, which are necessarily at even integer ν_b , still obey Eq. (3.7) and remain ground states of H_b [unless $\nu_b = \pm 2$ and Eq. (3.8) is sufficiently violated].

Having established the spectrum of H_{FD} in Eq. (3.1) for $\lambda = 0$, let us next discuss what happens once λ is turned on. Using the fact that the graphene form factors obey Eq. (2.14), we show in Appendix B 1 that H^{gb} can be rewritten in the low-energy spectrum of MSTG as

$$H^{\text{gb}} = \frac{1}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta \rho_{-\mathbf{q}}^{\text{g}} \left(\delta \rho_{\mathbf{q}}^{\text{b}} - \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} R_{\mathbf{G}} \right) + E_0(\nu_b), \quad (3.9)$$

where $E_0(\nu_b)$ is just a constant energy. So we immediately see that the property (3.7) of all of the exact TBG states $|\Psi_0^{\text{b}}(\nu_b)\rangle$ defined above, implies

$$H^{\text{gb}} |\Psi_j^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle = E_0(\nu_b) |\Psi_j^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle. \quad (3.10)$$

Consequently, all of the states $|\Psi_j^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$ remain *exact* eigenstates of the full MSTG Hamiltonian H_{FD} in the flat-decoupled limit, at arbitrary λ .

Whether the states $|\Psi_0^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$ will also remain the exact ground states is a more subtle question: since the states $|\Psi_0^{\text{b}}(\nu_b)\rangle$ break symmetries, $H_{\text{FD}}|_{\lambda=0}$ will have a gapless Goldstone spectrum. In principle, an arbitrarily small λ could lower the energy of some of those states below that of $|\Psi_0^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$. However, we show in Appendix B 2 that is not the case if $\nu_b = 0$ or $w_0 = 0$ or Eq. (3.8) holds. Therefore, a finite $\lambda > 0$ is required before $|\Psi_0^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$ cease to be the exact ground states.

B. Discussion of ground states

Taken together, we have shown that the states $|\Psi_j^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$ where $|\Psi_j^{\text{g}}(\nu_b)\rangle$ is just the spectrum of single-layer graphene at filling $\nu_g = \nu_g(\nu_b)$ and $|\Psi_0^{\text{b}}(\nu_b)\rangle$ is any of the states in Eq. (3.6) are exact eigenstates of the MSTG Hamiltonian in the chiral-flat-decoupled limit ($w_0 = W_{\text{TBG}} = D_0 = 0$), $H_{\text{FD}}|_{w_0=0}$, for any integer ν_b . Furthermore, there is a finite region of λ for which $|\Psi_0^{\text{g}}(\nu_b)\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$ will remain a groundstate of MSTG if $|\Psi_0^{\text{b}}(\nu_b)\rangle$ is a groundstate of TBG (recall that $|\Psi_0^{\text{b}}(\nu_b)\rangle$ is guaranteed to be a groundstate of TBG for $\nu_b = 0$ without further assumptions while it requires that the flat-metric condition is not too strongly violated for $\nu_b \neq 0$). Finally, away from the chiral limit $w_0 \neq 0$, the subset of states in Eq. (3.6) with $\eta_j^+ = \eta_j^-$ and $s_j^+ = s_j^-$ are known

to be ground states of TBG in the flat limit for integer ν_b and if the flat metric condition holds [60]. Our analysis shows that these states remain exact eigenstates for $\lambda \neq 0$ and also ground states for $|\lambda| < \lambda_c > 0$ in the non-chiral-flat-decoupled limit ($W_{\text{TBG}} = D_0 = 0$, $w_0 \neq 0$).

In all of these limits, we see that the graphene subsystem retains its (correlated but symmetry-unbroken and, depending on ν_b , doped) semimetallic properties for all integer filling fractions ν_b . This is consistent with experiment, where quantum oscillations indicate a dispersive Dirac cone at $D_0 = 0$ [30, 31]. Furthermore, the exact eigenstates established above will be used as our starting point for further analytical considerations in Sec. V and their product-state nature motivates our HF numerical study of the problem in Sec. IV. Both numerics and analytics will complement the discussion presented above by (i) validating the stability of the Dirac cones at $D_0 = 0$ in schemes that do not rely on λ being small and (ii) by tuning away from the exactly solvable limits ($W_{\text{TBG}} = D_0 \neq 0$) and (iii), for the numerics, including additional remote bands.

C. Resultant candidate states

To build the foundation for these additional analytical and numerical computations, we will use the exact (and highly degenerate) ground states established above in the chiral-flat-decoupled limit to construct a finite set of candidate phases and their respective order parameters.

We first define the *correlation matrix* $P_{\mathbf{k}}$ with elements

$$(P_{\mathbf{k}})_{(t,p,\eta,s),(t',p',\eta',s')} := \langle \Psi_0 | f_{\mathbf{k};(t,p),\eta,s}^\dagger f_{\mathbf{k};(t',p'),\eta',s'} | \Psi_0 \rangle \quad (3.11)$$

to characterize a given ground state $|\Psi_0\rangle$ of MSTG. Hermiticity implies $P_{\mathbf{k}}^\dagger = P_{\mathbf{k}}$. As is common, we further write $P_{\mathbf{k}} = \frac{1}{2}(\mathbb{1} + Q_{\mathbf{k}})$ and will use $Q_{\mathbf{k}}$ as our ‘‘order parameter’’ to characterize the (potentially symmetry-broken) structure of $|\Psi_0\rangle$. It must obey

$$Q_{\mathbf{k}}^\dagger = Q_{\mathbf{k}}, \quad \frac{1}{N} \sum_{\mathbf{k}} \text{tr}[Q_{\mathbf{k}}] = \nu_b + \nu_g \equiv \nu. \quad (3.12)$$

As we have seen above, the ground states of MSTG in the flat-decoupled limit ($W_{\text{TBG}} = D_0 = 0$) obey

$$(Q_{\mathbf{k}})_{(t,p,\eta,s),(t',p',\eta',s')} = \delta_{t,t'} (Q_{\mathbf{k}}^t)_{(p,\eta,s),(p',\eta',s')}, \quad (3.13)$$

i.e., do not exhibit any ‘‘coherence’’ between the graphene and TBG sectors. This is expected as the presence of σ_h requires any order parameter to be either even (diagonal in t space) or odd (off-diagonal) under σ_h ; due to the large density of states in the TBG sector, we expect the former to dominate. Once $D_0 \neq 0$, mixing is allowed, as we will see in our numerics below and discuss in detail analytically in Sec. VB.

TABLE II: We list the different candidate phases in the TBG-like subspace, constructed as the discrete set of states that are part of the large manifold of exact ground states in the chiral-flat-decoupled limit, see Sec. III, but transform under the irreducible representations of the symmetries of the real system. Here $\mathbf{1}$ ($\mathbf{3}$) is the singlet (triplet) representation of $SU(2)_s$ and 0 (1) the one(two)-dimensional representation of $U(1)_v$. For future reference in Sec. VII, we list the behavior (\pm denoting even/odd, and \times indicating absence) under both spinful (Θ_s), and valley ($\tilde{\Theta}$) time-reversal symmetry, see Table I. The last two columns indicate which states are Hund's partners [27], i.e., transform into each other when reversing the sign of the Hund's coupling J_H while being exactly degenerate in the $SU(2)_+ \times SU(2)_-$ -symmetric limit, and which sign of J_H favors the respective state.

Type	Short form	Q^b	\tilde{Q}^b	$SU(2)_s$	$U(1)_v$	C_{2z}	$\Theta_s/\tilde{\Theta}$	Hund's part.	J_H
spin polarized	SP	$\sigma_0\eta_0\mathbf{s}$	$\tilde{\sigma}_0\eta_0\mathbf{s}$	$\mathbf{3}$	0	\checkmark	$-/\times$	SVP	< 0
valley polarized	VP	$\sigma_0\eta_z s_0$	$\tilde{\sigma}_0\eta_z s_0$	$\mathbf{1}$	0	\times	$-/-$	—	0
Θ -even IVC	IVC ₊	$\sigma_0\eta_{x,y}s_0$	$\tilde{\sigma}_0\eta_{x,y}s_0$	$\mathbf{1}$	1	\checkmark	$+/-$	SIVC ₊	> 0
Θ -odd IVC	IVC ₋	$\sigma_y\eta_{x,y}s_0$	$\tilde{\sigma}_z\eta_{x,y}s_0$	$\mathbf{1}$	1	\checkmark	$-/+$	SIVC ₋	> 0
Θ -odd, sublattice pol./Hall	SLP ₋	$\sigma_y\eta_0 s_0$	$\tilde{\sigma}_z\eta_0 s_0$	$\mathbf{1}$	0	\checkmark	$-/-$	—	0
Θ -even, sublattice pol./valley Hall	SLP ₊	$\sigma_y\eta_z s_0$	$\tilde{\sigma}_z\eta_z s_0$	$\mathbf{1}$	0	\times	$+/+$	—	0
Θ -odd, spin-sublattice-pol./spin Hall	SSLP ₋	$\sigma_y\eta_0\mathbf{s}$	$\tilde{\sigma}_z\eta_0\mathbf{s}$	$\mathbf{3}$	0	\checkmark	$+/\times$	SSLP ₊	< 0
spin-valley polarized	SVP	$\sigma_0\eta_z\mathbf{s}$	$\tilde{\sigma}_0\eta_z\mathbf{s}$	$\mathbf{3}$	0	\times	$+/\times$	SP	> 0
Θ -even, spin-pol. IVC	SIVC ₊	$\sigma_0\eta_{x,y}\mathbf{s}$	$\tilde{\sigma}_0\eta_{x,y}\mathbf{s}$	$\mathbf{3}$	1	\checkmark	$-/\times$	IVC ₊	< 0
Θ -odd, spin-pol. IVC	SIVC ₋	$\sigma_y\eta_{x,y}\mathbf{s}$	$\tilde{\sigma}_z\eta_{x,y}\mathbf{s}$	$\mathbf{3}$	1	\checkmark	$+/\times$	IVC ₋	< 0
Θ -even, spin-subl. pol./spin-valley Hall	SSLP ₊	$\sigma_y\eta_z\mathbf{s}$	$\tilde{\sigma}_z\eta_z\mathbf{s}$	$\mathbf{3}$	0	\times	$-/\times$	SSLP ₋	> 0

Furthermore, the analysis above reveals that the ground state in the TBG sector will be of the form of Eq. (3.6). For instance, for $\nu_b = 0$ with $\eta_j^+ = \eta_j^- = (-1)^j$ and $s_j^+ = s_j^- = \uparrow$, $j = 1, 2$, it holds $\tilde{Q}_{\mathbf{k}}^b = \tilde{\sigma}_0\eta_0 s_z$ where $\tilde{Q}_{\mathbf{k}}^b = U^* Q_{\mathbf{k}}^b U^T$ is the order parameter $Q_{\mathbf{k}}^b$ in Eq. (3.13) in the TBG subspace transformed to the chiral basis of Eq. (3.5); in the basis of Eq. (3.11), it holds $Q_{\mathbf{k}}^b = \sigma_0\eta_0 s_z$, which we will refer to as spin polarized (SP) state. Here and in the following we will use σ_j ($\tilde{\sigma}_j$) to denote Pauli matrices in the band-space with index p (in the chiral basis with index c). Besides the SP state, Eq. (3.6) describes many other possible ground states, that are exactly degenerate in the chiral-flat-decoupled limit. A systematic way of seeing this proceeds by noting that the $U(4) \times U(4)$ identified [46, 59, 61] for TBG also persists as a symmetry of MSTG in chiral-flat-decoupled limit [36]; this immediately follows from the structure of H_{FD} in Eq. (3.1). We will here refer to this symmetry group as $(U(4) \times U(4))_{\text{b,cf}}$ and its action is particularly simple in the chiral basis [59],

$$\tilde{b}_{\mathbf{k}} \rightarrow \mathcal{U} \tilde{b}_{\mathbf{k}}, \quad \mathcal{U} = e^{i \sum_{j=0,3} \sum_{\mu,\mu'=0}^3 \varphi_{j,\mu,\mu'} \tilde{\sigma}_j \eta_{\mu} s_{\mu'}}. \quad (3.14)$$

The form of these transformations is readily inferred from Eq. (B8) which indicates that all \mathcal{U} with $[\mathcal{U}, \tilde{\sigma}_3] = 0$ will leave H_{FD} invariant. Under Eq. (3.14), the order parameter defined above transforms as $\tilde{Q}_{\mathbf{k}}^b \rightarrow \mathcal{U}^* \tilde{Q}_{\mathbf{k}}^b \mathcal{U}^T$, which allows us to generate the entire (continuous) set of exactly degenerate ground states from one “seed” state, such as the SP state, $\tilde{Q}_{\mathbf{k}}^b = \tilde{\sigma}_0\eta_0 s_z$. Since this seed state and \mathcal{U} commute with $\tilde{\sigma}_3$, we know that $[\tilde{Q}_{\mathbf{k}}^b, \tilde{\sigma}_3] = 0$ for

all ground states. Further noting that $(\tilde{Q}_{\mathbf{k}}^b)^2 = \mathbb{1}$ (physically related to the Slater-determinant nature), we can, thus, summarize the properties in the original basis as

$$Q_{\mathbf{k}}^b = Q^b, \quad [Q^b, \sigma_2] = 0, \quad (Q^b)^2 = \mathbb{1}, \quad \text{tr}[Q^b] = 0, \quad (3.15)$$

at charge neutrality.

The actual Hamiltonian of MSTG is not in the chiral-flat-decoupled limit and does not exhibit an exact $(U(4) \times U(4))_{\text{b,cf}}$ symmetry. Intuitively, this can be thought of as generating an easy axis in this multi-dimensional space of degenerate states, favoring a specific (subspace of) state(s) in Eq. (3.15). While energetics is required to decide which phase is ultimately preferred by the system—the aim of the subsequent sections—we can use symmetries to derive the discrete and finite set of possible “candidate states”: to this end, we impose only $U(1)_v$ and global spin rotations, $SU(2)_s$, as exact continuous symmetries. We then know that the candidate order parameters must transform under the irreducible representations of these symmetry groups (and be even or odd under the exact discrete symmetries C_{2z} and Θ), leading to the 11 options listed in Table II. In order to connect smoothly to the limit $D_0 \rightarrow 0$, we take here $Q_{\mathbf{k}}^g = -\sigma_z$ in the graphene subspace but emphasize that these order parameters are only used to define the different states and characterize their symmetries; for our numerical and analytical discussion below, they are only taken to be the starting point and we will allow for (and also find) mixing between the TBG-like and graphene-like sectors when $D_0 \neq 0$ as well as momentum dependence in $Q_{\mathbf{k}}$.

Finally, we point out that the model introduced in Sec. II, and which we study energetics in below, has an exact $SU(2)_+ \times SU(2)_-$ symmetry. Therefore, certain pairs of states, which we call *Hund's partners* following [27], have to be exactly degenerate, see Table II. As we noted earlier, in the realistic system, there is a non-zero *intervalley Hund's coupling*, $J_H \neq 0$, which will break $SU(2)_+ \times SU(2)_-$ down to $SU(2)_s$, albeit weakly, and favor one member of each of the pairs over the other. For most of the following study we will focus on the $SU(2)_+ \times SU(2)_-$ limit and, hence, can, without loss of generality, restrict the discussion to the first 7 states above the line in Table II. However, one has to keep in mind that the real system will realize only one state of each Hund's pair, which will depend on the (unknown) sign of J_H and precise form of the Hund's coupling [see Appendix F 5 for more details]. Taking, for concreteness, the intervalley Hund's coupling to be of the form

$$H_2 = \frac{J_H}{N} \sum_{\mathbf{q}} \mathbf{S}_{\mathbf{q}}^+ \cdot \mathbf{S}_{\mathbf{q}}^-, \quad \mathbf{S}_{\mathbf{q}}^{\pm} = \frac{1}{2} c_{\mathbf{k}+\mathbf{q}}^{\dagger} (\mathbb{1} \pm \eta_z) c_{\mathbf{k}}, \quad (3.16)$$

we can, in the chiral-flat-decoupled limit, uniquely associate a single state of each Hund's pair with a given sign of J_H ; this is indicated in the last column in Table II.

IV. HARTREE-FOCK NUMERICS

As it facilitates the presentation of the results, we will begin the discussion of correlated phases away from the chiral-flat-decoupled limit with the HF numerics and postpone the complementary analytics to Sec. V. Furthermore, we will first focus on the charge-neutrality point, $\nu = 0$.

A. Hartree-Fock Method

In the HF approximation, one focuses on Slater-determinant states $|\Psi[P_{\mathbf{k}}]\rangle$ characterized by the correlation matrix $P_{\mathbf{k}}$ as defined in Eq. (3.11). Consequently, it holds $P_{\mathbf{k}}^2 = P_{\mathbf{k}}$ or, equivalently, $Q_{\mathbf{k}}^2 = \mathbb{1}$; this is also true for the exact candidate ground states constructed in Sec. III C for the chiral-flat-decoupled limit and, hence, the HF approximation is expected to provide reliable results.

The goal of our HF numerics will be to determine the optimal $Q_{\mathbf{k}}$ that yields the lowest energy expectation value with respect to the interacting Hamiltonian for MSTG introduced in Sec. II. To be more specific, we start from the full Hamiltonian, $H = H_{0,1}^{\text{Full}} + H_{0,2}^{\text{Full}} + H_1$, consisting of the continuum model in Eq. (2.1), the displacement-field term in Eq. (2.4), supplemented by the density-density interaction in Eq. (2.8), and perform a mean-field decoupling. Using the same notation as

in Eq. (2.10b), the resulting HF mean-field Hamiltonian reads as

$$\begin{aligned} H^{\text{MF}} &= \sum_{\mathbf{k} \in \text{MBZ}} \epsilon_{n,\eta}(\mathbf{k}) f_{\mathbf{k};n,\eta,s}^{\dagger} f_{\mathbf{k};n,\eta,s} \\ &+ \sum_{\mathbf{k} \in \text{MBZ}} f_{\mathbf{k}}^{\dagger} [h_{\text{H}}[P](\mathbf{k}) + h_{\text{F}}[P](\mathbf{k})] f_{\mathbf{k}} \\ &- \frac{1}{2} \sum_{\mathbf{k} \in \text{MBZ}} \text{Tr} [h_{\text{H}}[P](\mathbf{k}) P_{\mathbf{k}}^T + h_{\text{F}}[P](\mathbf{k}) P_{\mathbf{k}}^T], \end{aligned} \quad (4.1)$$

where the Hartree and Fock contributions to the mean-field Hamiltonian can be written in terms of the projector $P_{\mathbf{k}}$ as

$$h_{\text{H}}[P](\mathbf{k}) = \frac{1}{N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) F_{\mathbf{k},\mathbf{G}} \sum_{\mathbf{k}' \in \text{MBZ}} \text{Tr} [F_{\mathbf{k}',\mathbf{G}}^* P_{\mathbf{k}'}] \quad (4.2)$$

and

$$h_{\text{F}}[P](\mathbf{k}) = -\frac{1}{N} \sum_{\mathbf{q}} V(\mathbf{q}) F_{\mathbf{k},\mathbf{q}}^{\dagger} P_{\mathbf{k}+\mathbf{q}}^T F_{\mathbf{k},\mathbf{q}}, \quad (4.3)$$

respectively. This form of the HF mean-field Hamiltonian is valid for an arbitrary number of bands kept. In the numerics presented here, we will focus on the four bands for each spin and valley flavor that are closest to the Fermi level, which contains the graphene-like and TBG-like bands we focus on in the analytics. We verify for representative values of D_0 and w_0 that the solutions we obtain are stable against doubling the number of remote bands in our self-consistent calculation in Appendix C 3.

As pointed out in several HF works on TBG [45–48], it is important to note that the continuum model $H_{0,1}^{\text{Full}}$ already references electron-electron interactions in the experimentally determined values for microscopic model parameters and therefore we must define a reference subtraction projector P_0 such that interactions will not be double counted in our numerics. We here choose P_0 such that the projected low-energy Hamiltonian exhibits the manifestly particle-hole symmetric interaction in Eq. (2.10). As shown in [59] for TBG, this ansatz has the natural interpretation of effectively taking into account the HF contributions from all remote bands that have been projected out.

To determine the optimal $P_{\mathbf{k}}$, we start with an initial guess for it with the symmetries of a given candidate order in Table II. We then use the HF Hamiltonian in Eq. (4.1) to compute a new projector $P_{\mathbf{k}}$ and iterate until $P_{\mathbf{k}}$ converges. More details on our iterative HF procedure and subtraction point are given in Appendix C 2.

B. Band Structures

In this section, we will discuss the band structures of the self-consistent solutions we find at $D_0 = 0$ for each

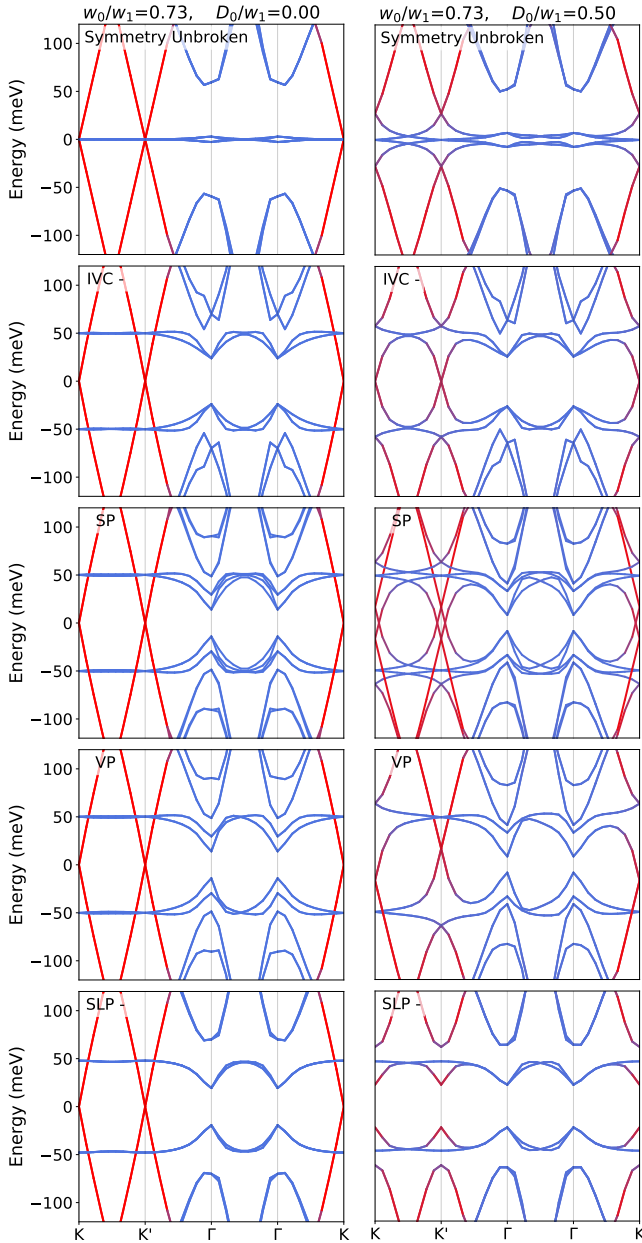


FIG. 3: HF band structures in the limit $D_0 = 0$ (left) and nonzero D_0 (right) as obtained from Eq. (4.1) for the optimal $P_{\mathbf{k}}$ in the respective symmetry-breaking channel defined in Table II. We show only the states which are either ground states or subleading energy states for some region of our phase diagram, with a full set of band structures available in Appendix C3. Each band of the HF Hamiltonian is colored according to the expectation value of the mirror symmetry operator within each band as a function of \mathbf{k} , with the expectation value ranging from -1 (red) to +1 (blue).

of the states in Table II and how these band structures evolve as D_0 increases.

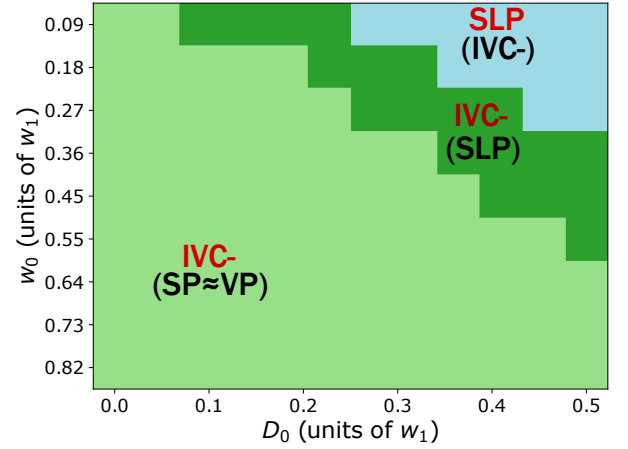


FIG. 4: Phases obtained by self-consistent HF method as a function of w_0 and D_0 . Both the leading instability (red) and the subleading ones (black, in parenthesis) are shown. The phases are identified by the symmetries they break, see Table II, and representative band structures can be found in Fig. 3. We show only one state as a representative of each pair of Hund's partners which are degenerate in our Hartree-Fock procedure. We label the nearly degenerate SLP_+ , SLP_- , $SSLP_+$, and $SSLP_-$ phases with "SLP", though note the preference in our numerics is for the SLP_- state, from a Hartree contribution.

1. Self-Consistent Band Structures at $D_0 = 0$

At $D_0 = 0$, we can separately describe the behavior of the graphene-like and TBG-like bands for each type of state, since, for all states we consider, the graphene-like bands near the Fermi-level at $D_0 = 0$ do not mix with any other bands near the K and K' points. They have a bandwidth larger than the scale of Coulomb interactions. For this reason, the Dirac cones of the graphene-like bands prefer a $Q_{\mathbf{k}}^g$ which equally fills the lower bands of the continuum model and preserves all point group symmetries for every class of solution we study. The Dirac cones thus remain semimetallic and are, in this sense, "spectators" at $D_0 = 0$, in agreement with Sec. III. On the other hand, the TBG-like bands have a bandwidth (5-10 meV) smaller than the scale of the Coulomb energy at $D_0 = 0$ and therefore become insulating as they are polarized for a given symmetry-breaking $P_{\mathbf{k}}$.

We find converged solutions for each ansatz in Table II and show representative band structures for those states which have the lowest energy in the leftmost panels of Fig. 3. Additional band structures for solutions not shown in Fig. 3 can be found in Appendix C3. We note the similarity of the IVC_- band structure in the TBG-like bands to the band structure of the ground state in Ref. 46.

2. Self-Consistent Band Structures for $D_0 > 0$

All the solutions described for $D_0 = 0$ are insulating in the TBG-like bands and semimetallic in the graphene-like bands. However, as D_0 increases and the TBG-like and graphene-like bands begin to hybridize, the graphene-like bands begin to play a more important role.

For the spin and valley polarized states (SP and VP) which preserve $C_{2z}\Theta$ and $U(1)_v$, the hybridized Dirac crossings are protected (and pinned to the K/K' points due to C_{3z}), meaning if the TBG-like bands acquire a spin or valley polarization, the Dirac crossings of the graphene-like bands must connect to these polarized bands as they are pushed away from the Fermi level. Away from the K points of the MBZ, the graphene-like bands will likely still prefer to fill the lowest bands of the non-interacting model. Therefore, the VP and SP are generically expected to be metallic for $D_0 > 0$. This is indeed what we find, as shown in the right panels for the SP and VP states in Fig. 3.

The band structures of the IVC_{\pm} states also retain Dirac crossings at the K/K' points for nonzero D_0 , as they exhibit C_{3z} and a \mathbf{k} -local anti-unitary symmetry that commutes with C_{3z} (for the IVC_+ and IVC_- these are $C_{2z}\Theta$ and $C_{2z}\tilde{\Theta}$, respectively). However, unlike the SP and VP states, the IVC_{\pm} states also preserve $SU(2)_s$ and C_{2z} , which pin the Dirac crossings at the Fermi level at $\nu = 0$. We therefore expect the intervalley coherent states will remain semimetallic as D_0 increases. We observe this to be true for the self consistent solutions, as can be seen in the right IVC_- panel in Fig. 3.

The last class of states are the sublattice-polarized, $C_{2z}\Theta$ -symmetry-breaking states (SLP_{\pm} , $SSLP_-$) which preserve $U(1)_v$. We expect these states will generally be insulating for nonzero D_0 as there are no protected Dirac crossings and both the TBG-like and graphene-like bands can be gapped out. This is indeed seen in our numerics, with insulating band structures for the SLP_+ , SLP_- , and $SSLP_-$ states. The band structure of the SLP_- state is shown in Fig. 3.

C. Energies and phase diagram

Having established the band structures of the different possible phases, we next turn to their relative energetics and discuss which states are expected to be favored energetically.

The evolution of the energies of each of our self-consistent solutions as a function of w_0 and D_0 is shown in Figs. 5 and 6. As mentioned before, at $D_0 = 0$, the Hamiltonian of the system is given by the sum of the Hamiltonian of TBG and that of graphene, both with Coulomb interactions, which are further coupled to each other by a density-density interaction. While we only

constructed exact eigenstates in Sec. III for the chiral-flat-decoupled limit, we expect a similar picture when $w_0, W_{\text{TBG}} \neq 0$: given the bandwidth of the graphene-like bands is large compared to the scale of the Coulomb interactions, we expect the graphene bands will prefer to fill the lower bands of the continuum model. As the graphene density of states is small compared to that of the flat bands of TBG, it should not crucially alter the ground state in the TBG sector—at least close to the magic angle. Based on previous work [46, 48], we thus expect that the IVC_- state has the lowest energy for $D_0 = 0$ (though with a smaller energy difference than in previous works between our IVC_- and spin polarized phase due to our choice of subtraction point). Both expectations for the graphene-like and TBG-like bands are confirmed by our numerics which finds the IVC_- state has the lowest energy of all our candidates for all values of w_0 studied in the decoupled limit, $D_0 = 0$. We will also recover these observations analytically in Sec. V.

The lowest energy state for $D_0 \neq 0$ cannot be directly inferred from knowledge of the physics of TBG as a finite D_0 induces hybridization between the TBG-like and graphene-like bands near the K/K' points of the MBZ; it further breaks symmetries in the TBG sector and, hence, changes the basic form of its dispersion and interaction matrix elements (form factors). Consequently, it is not clear whether the ground state in the TBG-like and/or graphene-like sector changes with increasing D_0 . As can be seen in Figs. 5 and 6 as well as in the corresponding phase diagram in Fig. 4, we find within HF that the IVC_- remains the ground state for an extended range of D_0 , which increases with w_0 . For reference, the range of D_0 in Figs. 4, 5, and 6 when combined with additional studies at larger D_0 in Appendix E corresponds roughly to the range of displacement fields studied experimentally in Ref. 30.

Beyond the critical value of D_0 for the IVC_- , the sublattice-polarized (SLP) group of states (SLP_{\pm} , $SSLP_-$) dominates. While these latter three states are almost degenerate for all parameters D_0 , w_0 studied, there is a slight preference towards the time-reversal-odd SLP_- (quantum Hall), predominantly associated with the Hartree energy. While a SP or VP phase does not appear as a ground state in Fig. 4, we find the energetically close SP or VP states are either the second or third lowest energy state to the IVC_- and SLP group across the phase diagram. In the next subsection, we will recover many of these features analytically by investigating the aforementioned energetic contributions perturbatively.

V. ANALYTICAL PERTURBATION THEORY

Finally, we complement the HF numerics with an analytical study of the behavior of the energies and the order

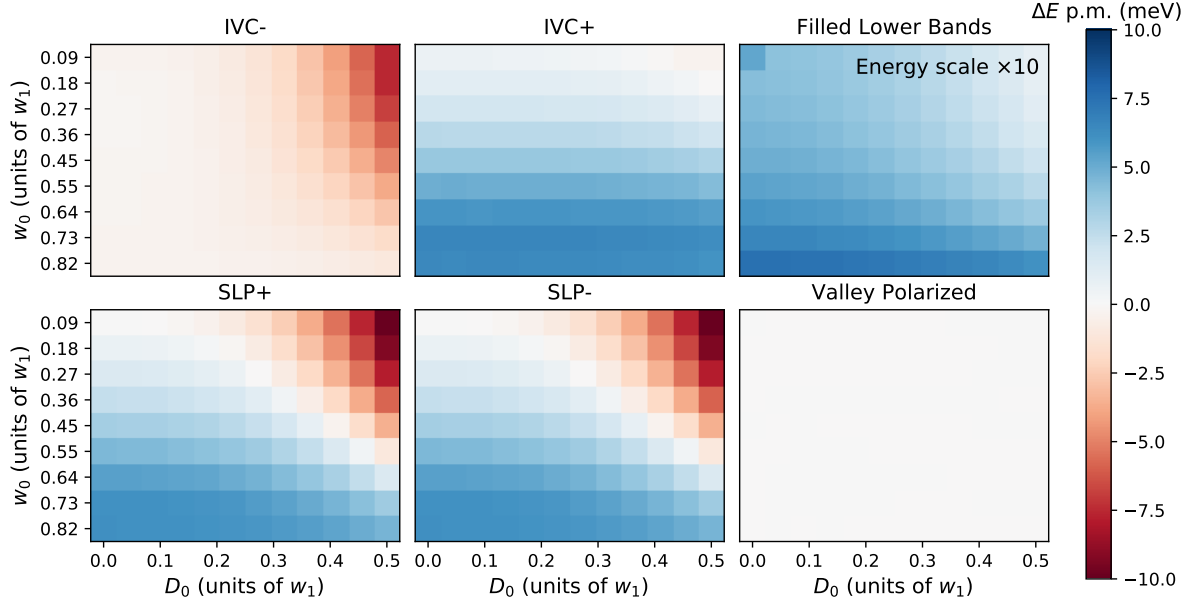


FIG. 5: Energies obtained from self-consistent HF calculations for selected candidate states of Table II. All energies shown are relative to the energy of the self-consistent SP solution and measured in meV per moiré unit cell. We take $\epsilon = 7$, $d_s = 40$ nm, and $w_1 = 124$ meV and 4 bands per spin and valley. We find the IVC_- is lower in energy than the SP state for all values of w_0 and D_0 . The IVC_+ state is always higher in energy than the SP in agreement with Table III. The VP state is nearly exactly degenerate with the SP state, even as D_0 increases. Both the SLP_+ and SLP_- states have a region of energy lower than both the SP and IVC_- states at small w_0 and large D_0 . While they appear nearly degenerate, a small Hartree contribution favors the time-reversal odd SLP_- over the SLP_+ state. However, this difference between the two sublattice polarized states (< 1 meV) is not visible in the energies we show.

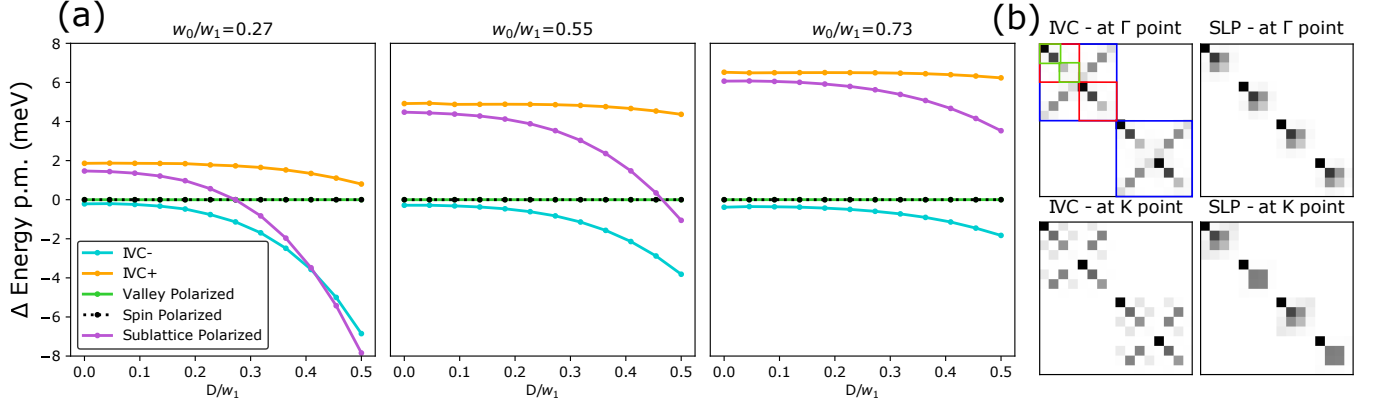


FIG. 6: (a) Linecuts for fixed w_0 showing the energy of our self consistent solutions relative to a SP state as a function of D_0 . We use the same parameters as in Fig. 5. Note that the SLP_{\pm} and $SSLP_-$ are not exactly degenerate, as discussed in the main text, but are plotted as a single line here since the splitting is too small to be visible on the scale of the plot. (b) Matrix form of our converged IVC_- and SLP_- orders $P_{\mathbf{k}}$ for \mathbf{k} near the Γ point and near the K point. The matrix structure of $P_{\mathbf{k}}$ is organized such that the largest block outlined in blue denotes spin flavor, the next largest block outlined in red denotes valley flavor, the green block denotes upper and lower bands in the continuum model and the final two boxes denote the graphene-like and TBG-like band in the upper-band box and the TBG-like then graphene-like bands in the lower-band box.

parameters of the degenerate ground states of the chiral-flat-decoupled limit ($w_0 = W_{\text{TBG}} = D_0 = 0$) of Sec. III and Table II when turning on W_{TBG} , D_0 , and w_0 . We first investigate the ordering tendencies of the graphene-like bands (Sec. VA) and their mixing with the TBG-like

bands (Sec. VB), before addressing the energetic contributions coming from the D_0, w_0 -induced distortion of the TBG-like form factors (Sec. VC) and band structure (Sec. VD).

A. Ordering in the graphene-like bands

Since the graphene and graphene-like bands in Fig. 2(b) are highly dispersive, starting from a flat-band limit, as is natural for the TBG-like sector, is not possible for the graphene(-like) bands. Since their bands are coupled, without any band gap, to the TBG-like bands with a high density of states, treating the interactions between the two subsystems as a perturbation is also not necessarily controlled for realistic parameters. Instead, we here use a different control parameter: based on the band structure, we expect the effect of the TBG(-like) bands on the graphene(-like) bands to be the strongest around the K/K' points and very weak away from it. To formalize this, let us assume that the graphene-like bands remain in their filled-lower-bands state away from the K/K' points, but allow them to be “deformed” in the region $\mathcal{A} = \mathcal{A}_+ \cup \mathcal{A}_-$ of the MBZ, where \mathcal{A}_+ (\mathcal{A}_-) are simply connected and centered around the K (K') point. As explained in detail in Appendix F 3, we study the energy of symmetry-allowed ordering tendencies of the graphene-like bands for the different candidate phases in Table II in the limit where the area of \mathcal{A} is small (compared to that of the MBZ).

To illustrate this procedure, let us consider the SP state. Since the TBG-like bands break spin-rotation symmetry, it is natural to assume that the same happens to the graphene-like bands in \mathcal{A} ; postponing the discussion of mixing between the bands to Sec. V B, this means that $(Q_{\mathbf{k}}^g)_{\eta,\eta'} = \delta_{\eta,\eta'}\sigma_0\mathbf{s}$, $\mathbf{k} \in \mathcal{A}_\eta$, in the notation introduced in Sec. III C, while $(Q_{\mathbf{k}}^g)_{\eta,\eta'} = -\delta_{\eta,\eta'}\sigma_3$ for all other \mathbf{k} . Here we already anticipated (as is also readily checked within this formalism) that it is energetically more favorable if only the graphene valley $\eta = +$ ($\eta = -$) that is at low energies in the region \mathcal{A}_+ (\mathcal{A}_-) exhibits spin polarization. Denoting the linear size of \mathcal{A}_\pm by $\Delta k > 0$, the structure of the energetic change associated with the deformation of the graphene order is asymptotically given by

$$\Delta E \sim d_1|D_0|(\Delta k)^2 + d_2(\Delta k)^3 + I_{gg}(\Delta k)^3 - I_{gb}D_0^2(\Delta k)^2 \quad (5.1)$$

for small Δk and D_0 ; here, $d_{1,2} > 0$ are positive constants (independent of Δk and D_0 , but dependent on w_0) associated with the graphene-like dispersion, while I_{gg} depends on the graphene-graphene form factors F^{gg} and $I_{gb} > 0$ on F^{gb} and F^{gg} . The explicit form and derivation can be found in Appendix F 3.

From Eq. (5.1), we can read off the behavior of the graphene-like bands of the SP state. For $D_0 = 0$, we see that $\Delta E > 0$ (deformation is energetically disfavored) as long as $d_2 > I_{gg}$. While this is what our numerics shows, we point out that $d_2 < I_{gg}$ would imply that single-layer graphene spontaneously magnetizes, which is known to be not the case. This agrees with our result in Sec. III

based on adiabatically turning on the coupling between the TBG and graphene system and our numerics which displays unpolarized graphene Dirac cones in Fig. 3 at $D_0 = 0$. When $D_0 \neq 0$, we see that the energetic cost coming from the dispersion in the first line of Eq. (5.1) will always overcompensate the energetic gain coming from the scattering between the graphene- and TBG-like bands (at least in a finite range of $D_0 \neq 0$). This is also consistent with our HF numerics: as can be seen in the band structure in Fig. 3, the graphene-like bands are not spin-polarized for non-zero D_0 [see also Fig. 17 where $Q_{\mathbf{k}}$ as obtained in HF are shown].

The ordering tendencies in the graphene-like bands for all other states in Table II can be analyzed in a similar way, see Appendix F 3. For instance, for the VP we find that the graphene-like bands are found to not develop valley polarization either—both for $D_0 = 0$ and $D_0 \neq 0$ and again consistent with the HF numerics. The same holds for the IVC $_{\pm}$ states, where breaking of $U(1)_v$ in the graphene-like bands is strongly suppressed due to the fact that, at the K (K') point, the $\eta = -$ ($\eta = +$) graphene-like band energies are large (already at $D_0 = 0$), see Fig. 2.

The situation is different for the SLP group of states (SLP $_{\pm}$ and SSLP $_{\pm}$). To begin with the SLP $_{\pm}$ states, we have $(Q_{\mathbf{k}}^g)_{\eta,\eta'} = \delta_{\eta,\eta'}(-\sigma_3 \cos \theta_{\mathbf{k}} + \sigma_2 \sin \theta_{\mathbf{k}})$, $\mathbf{k} \in \mathcal{A}_\eta$, and we need to determine $\theta_{\mathbf{k}}$, with $\theta_{\mathbf{k}} = \mp \theta_{-\mathbf{k}}$ for SLP $_{\pm}$ due to Θ and C_{2z} , by minimizing the energy. The change of the energy as a consequence of this “deformation” is of the form

$$\Delta E_{gg}[\theta_{\mathbf{k}}] = \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_\eta} [A_{\mathbf{k},\eta}(1 - \cos \theta_{\mathbf{k}}) + B_{\mathbf{k},\eta} \sin \theta_{\mathbf{k}}] \quad (5.2)$$

for both SLP $_+$ and SLP $_-$. The explicit expressions for $A_{\mathbf{k},\eta}$ and $B_{\mathbf{k},\eta}$, which are functions of w_0 and D_0 , are given in Appendix F 3. From these expressions, it follows that $B_{\mathbf{k},\eta} = 0$ and $A_{\mathbf{k},\eta} > 0$ for $D_0 = 0$ such that $\Delta E[\theta_{\mathbf{k}}]$ is minimized when $\sin \theta_{\mathbf{k}} = 0$ (with $\Delta E = 0$) and there is no order in the graphene bands—again in agreement with Sec. III and the HF numerics. Once $D_0 \neq 0$, we get $B_{\mathbf{k},\eta} \neq 0$ and $\Delta E < 0$ by choosing a profile with $\sin \theta_{\mathbf{k}} \neq 0$. In other words, the graphene-like bands will develop SLP $_{\pm}$ order for any non-zero D_0 , as is visible in the $Q_{\mathbf{k}}$ for the SLP $_-$ state close to the K point shown in Fig. 6(b). This gaps out the graphene cone, as can also be seen in our HF band structure in Fig. 3. The energetic gain scales as $g_{\pm}(\Delta k)^2|D_0|^3$, $g_{\pm} > 0$, for small D_0 and Δk for the SLP $_{\pm}$ state; the prefactors differ, $g_+ \neq g_-$, due to the symmetry-imposed constraint $\theta_{\mathbf{k}} = \mp \theta_{-\mathbf{k}}$. For small w_0 , we can show that $B_{\mathbf{k},\eta} > 0$ such that $g_- > g_+$, i.e., the SLP $_-$ state can gain more energy than the SLP $_+$.

For the SSLP $_-$ state and its Hund’s partner SSLP $_+$, which we discuss here explicitly for reasons that will become clear shortly, we have $(Q_{\mathbf{k}}^g)_{\eta,\eta'} = \delta_{\eta,\eta'}(-\sigma_3 \cos \theta_{\mathbf{k}} + \sigma_2 \mathbf{s} \sin \theta_{\mathbf{k}})$, $\mathbf{k} \in \mathcal{A}_\eta$. Symmetry imposes $\theta_{\mathbf{k}} = \mp \theta_{-\mathbf{k}}$ for

SSLP $_{\pm}$. The deformation-related energy change ΔE is found to be again of the form of Eq. (5.2); however, we here obtain $B_{\mathbf{k},\eta} = 0$ for any D_0 or w_0 . This is consistent with $SU(2)_+ \times SU(2)_-$ which requires that SSLP $_{\pm}$ have the same energy. We, hence, have shown that the graphene-like bands do not develop any direct SSLP $_{\pm}$ order and gain energy in the process. They will, however, hybridize with the TBG-like bands, as we discuss in the next subsection, which also gaps out the Dirac cones.

In Table III, we summarize our findings that only the SLP $_{\pm}$ states can benefit from ordering the graphene-like bands and how the respective energy gain scales.

B. Mixing between the bands

As anticipated above, we next look into the mixing between the TBG-like and graphene-like bands. While it is clear by symmetry that the mixing has to vanish for $D_0 = 0$ and be generically present for $D_0 \neq 0$, we here investigate the associated energetic gain and precise form of the band mixing for our candidate states in Table II. As before, we here outline the basic strategy and discuss the major results of this calculation and defer the details to Appendix F 4.

Our starting point are product states characterized by a correlator with $Q_{\mathbf{k}}$ as given in Eq. (3.13), i.e., without any coherence between the TBG-like and graphene-like bands. In accordance with our analysis of Sec. III, we will take $Q_{\mathbf{k}}^g = -\sigma_z$ and let $Q_{\mathbf{k}}^b$ be any of the candidate orders. To introduce momentum-dependent coherence between these sets of bands, we “deform” $Q_{\mathbf{k}}$ by a unitary transformation, $U_{\mathbf{k}}$, and take

$$Q'_{\mathbf{k}} = U_{\mathbf{k}} Q_{\mathbf{k}} U_{\mathbf{k}}^{\dagger}, \quad U_{\mathbf{k}} = e^{i\Lambda_{\mathbf{k}}}, \quad \Lambda'_{\mathbf{k}} = \Lambda_{\mathbf{k}} \in \mathbb{C}^{16 \times 16}, \quad (5.3)$$

as ansatz for the correlator. Our goal will be to find the optimal momentum-dependent $\Lambda_{\mathbf{k}}$ to minimize the energy. Since we are interested in band mixing, we will restrict $\Lambda_{\mathbf{k}}$ to act as a superposition of ζ_1 and ζ_2 , $\Lambda_{\mathbf{k}} = \sum_{j=1,2} \mathcal{M}_{\mathbf{k},j} \zeta_j$, $\mathcal{M}_{\mathbf{k},j} \in \mathbb{C}^{8 \times 8}$, with ζ_j denoting Pauli matrices acting in the space of TBG-like and graphene-like bands (with index $t = b, g$). Furthermore, $\Lambda_{\mathbf{k}}$ will be constrained by the symmetries of the state under consideration.

To illustrate the procedure, let us focus on the IVC $_-$ since this state was found to be dominant in the HF numerics. Choosing $Q_{\mathbf{k}}^b = \sigma_y \eta_y$ for concreteness, this state preserves the $C_{2z} \Theta$ symmetry of Table I, which forces $\Lambda_{\mathbf{k}}$ to obey $(\Lambda_{\mathbf{k}})^* = -\Lambda_{\mathbf{k}}$. Furthermore noting that the IVC $_-$ state does not break the $SU(2)_s$ symmetry (while postponing the consequences of C_{2z} which will relate $\Lambda_{\mathbf{k}}$ and $\Lambda_{-\mathbf{k}}$), it follows that $\Lambda_{\mathbf{k}}$ has to be a momentum-dependent superposition of the 16 generators

$$\zeta_x \sigma_y \eta_{0,x,z}, \quad \zeta_x \sigma_{0,x,z} \eta_y, \quad \zeta_y \sigma_{0,x,z} \eta_{0,x,z}, \quad \zeta_y \sigma_y \eta_y. \quad (5.4)$$

To simplify further, we can focus on those 8 linear combinations [see Eq. (F46) for their explicit form] of the terms in Eq. (5.4) that anti-commute rather than commute with $Q_{\mathbf{k}}$ of the IVC $_-$ state.

The first energetic constraint we will take into account is related to the fact that the graphene-like bands of valley $\eta = -$ ($\eta = +$) are far away from the Fermi level at the K (K') point, see Fig. 2. This means that all mixing processes in $\Lambda_{\mathbf{k}}$ that induce a finite occupation of the upper or unoccupied states in the lower graphene-like band of valley $\eta = -$ ($\eta = +$) at the K (K') point are suppressed. In Appendix F 4, we show that this is equivalent to demanding that $(\mathcal{M}_{\mathbf{k}})_{(p',\eta',s'),(p,\mp,s)} = 0$ for $\mathbf{k} \in \mathcal{A}_{\pm}$. This reduces the number of generators further from 8 to only 4 [given in Eq. (F52)]. For each of these 4 generators, Λ_j , $j = 1, 2, 3, 4$, we compute the change of energy associated with the deformation in Eq. (5.3) where $U_{\mathbf{k}} = e^{i\varphi_{\mathbf{k}} \Lambda_j}$. It is found to be of the form

$$\Delta E_{gb}[\varphi_{\mathbf{k}}] = \sum_{\mathbf{k} \in \mathcal{A}_+} \alpha_{\mathbf{k}} \sin^2 \varphi_{\mathbf{k}} + \beta_{\mathbf{k}} \sin \varphi_{\mathbf{k}} \cos \varphi_{\mathbf{k}}, \quad (5.5)$$

where $\alpha_{\mathbf{k}}$ and $\beta_{\mathbf{k}}$ are expressions involving the form factors, the interaction $V(\mathbf{q})$, and the band structure [see Appendix F 4 for explicit form], and thus depend on D_0 and w_0 . In Eq. (5.5), we have already taken into account the C_{2z} constraint that allowed us to write it as a sum over $\mathbf{k} \in \mathcal{A}_+$ only.

First, we find $\beta_{\mathbf{k}} = 0$ for $D_0 = 0$ and, hence, no mixing between the bands as expected by symmetry. However, even when $D_0 \neq 0$, we still obtain $\beta_{\mathbf{k}} = 0$ for two of the four generators, while $\beta_{\mathbf{k}} \neq 0$ for the other two (say $\Lambda_{1,2}$) in the chiral limit ($w_0 = 0$). Consequently, at least when w_0 is not too large, the generators $\Lambda_{1,2}$ have to dominate. Which of those remaining two is dominant, cannot be determined purely analytically as it will depend on non-universal values of the form factors, so let us compare with numerics. The *mixing matrix*, defined as

$$\left(M_{\mathbf{k}}^Q \right)_{(p,\eta,s),(p',\eta',s')} := (Q_{\mathbf{k}})_{(b,p,\eta,s),(g,p',\eta',s')} \quad (5.6)$$

and can be computed straightforwardly for the two candidate generators $\Lambda_{1,2}$. We find

$$M_{\mathbf{k}}^Q = 2\sqrt{2}i (\sigma_x P_{\pm} \pm \sigma_0 \eta^{\mp}) \varphi_{\mathbf{k}}, \quad \mathbf{k} \in \mathcal{A}_{\pm}, \quad (5.7a)$$

and

$$M_{\mathbf{k}}^Q = 2\sqrt{2}i (\sigma_z P_{\pm} \pm i\sigma_y \eta^{\mp}) \varphi_{\mathbf{k}}, \quad \mathbf{k} \in \mathcal{A}_{\pm}, \quad (5.7b)$$

for Λ_1 and Λ_2 , respectively. Here, we defined $P_{\pm} = (\eta_0 \pm \eta_z)/2$ and $\eta^{\pm} = (\eta_x \pm i\eta_y)/2$. Intuitively, the first mixing matrix means that the lower (upper) graphene-like bands of the valley that is at low energies at K or K' mixes with the upper (lower) and lower (upper) TBG-like bands of the same and opposite valley, respectively. The second option in Eq. (5.7b) describes the “twisted” situation where the lower (upper) graphene-like band mixes

TABLE III: Summary of the different energetic contributions, as discussed in Sec. V, for the candidate orders in Table II when tuning away from the chiral-flat-decoupled limit ($w_0 = W_{\text{TBG}} = D_0 = 0$). By construction, the energies are identical for the respective Hund's partners and are, hence, omitted. Here $\Delta E_{bb}(W_{\text{TBG}} = 0)$ is the change of energy, relative to SP phase, coming from the modifications of the form factors in the TBG-like bands when turning on w_0 and D_0 in the flat limit ($W_{\text{TBG}} = 0$). As indicated in the column labelled ΔE_{gg} , only the SLP_\pm states gain energy by ordering in the graphene-like bands immediately when $D_0 \neq 0$. In all cases, $D_0 \neq 0$ leads to mixing between the bands, with energetic gain as listed in the chiral limit, $w_0 = 0$, in the column ΔE_{gb} ; here Δk is the linear size of the fraction \mathcal{A}_\pm of the MBS where ordering in the graphene-like bands and mixing take place. The four columns with \bar{E}_j^b indicate which of the four contributions to the TBG dispersion in Eq. (2.6) can lower the energies to second order in W_{TBG} . Finally, in the last columns, we list the energy change associated with a finite value of the Hund's coupling in Eq. (3.16) for the respective state (ΔE_J) and its Hund's partner (ΔE_J^H), if it exists, in the chiral-flat-decoupled limit. All coefficients obey $c_j, g_\pm, g_j, \beta_j > 0$ and explicit expressions can be found in Appendix F.

Type	Q^b	\tilde{Q}^b	$\Delta E_{bb}(W_{\text{TBG}} = 0)$	ΔE_{gg}	ΔE_{gb}	\bar{E}_0^b	$\bar{E}_1^b w_0$	$\bar{E}_2^b D_0$	$\bar{E}_3^b w_0 D_0$	ΔE_J	ΔE_J^H
SP	$\sigma_0 \eta_0 \mathbf{s}$	$\tilde{\sigma}_0 \eta_0 \mathbf{s}$	0 (by definition)	0	$-g_1(\Delta k)^2 D_0^2$	\times	\times	\times	\times	$2J_H \beta_1$	$-2J_H \beta_1$
VP	$\sigma_0 \eta_z s_0$	$\tilde{\sigma}_0 \eta_z s_0$	0	0	$-g_1(\Delta k)^2 D_0^2$	\times	\times	\times	\times	0	—
IVC ₊	$\sigma_0 \eta_{x,y}$	$\tilde{\sigma}_0 \eta_{x,y}$	$c_1 D_0^4 + c_3 w_0^2$	0	$-g_2(\Delta k)^2 D_0^2$	\times	\checkmark	\checkmark	\times	$-3J_H \beta_2$	$J_H \beta_2$
IVC ₋	$\sigma_y \eta_{x,y}$	$\tilde{\sigma}_z \eta_{x,y}$	$c_1 D_0^4 + c_4 w_0^2 D_0^4$	0	$-g_2(\Delta k)^2 D_0^2$	\checkmark	\checkmark	\times	\times	$-3J_H \beta_2$	$J_H \beta_2$
SLP ₋	σ_y	$\tilde{\sigma}_z$	$c_3 w_0^2 + c_4 w_0^2 D_0^4$	$-g_-(\Delta k)^2 D_0 ^3$	$-g_1(\Delta k)^2 D_0^2$	\checkmark	\times	\checkmark	\times	0	—
SLP ₊	$\sigma_y \eta_z$	$\tilde{\sigma}_z \eta_z$	$c_2 D_0^4 + c_3 w_0^2 + c_4 w_0^2 D_0^4$	$-g_+(\Delta k)^2 D_0 ^3$	$-g_3(\Delta k)^2 D_0^2$	\checkmark	\times	\checkmark	\times	0	—
SSLP ₋	$\sigma_y s_z$	$\tilde{\sigma}_z s_z$	$c_3 w_0^2 + c_4 w_0^2 D_0^4$	0	$-g_1(\Delta k)^2 D_0^2$	\checkmark	\times	\checkmark	\times	$2J_H \beta_3$	$-2J_H \beta_3$

with the lower (upper) TBG-like bands of the same valley and the upper (lower) band in the opposite valley. The HF result for $Q_{\mathbf{k}}$ close to the K point, shown in Fig. 6(b), is consistent with Eq. (5.7a).

Minimizing Eq. (5.5) and noting that $\beta_{\mathbf{k}}$ scales linearly with D_0 for small D_0 while $\alpha_{\mathbf{k}} > 0$ at $D_0 = 0$, we find the scaling of the energetic gain to be $\Delta E_{gb} \sim -g_3(\Delta k)^2 D_0^2$, where $g_3 > 0$; this is indicated in Table III.

In the same way, all other candidate states can be studied, see Appendix F 4. In accordance with our expectation based on symmetry, we find no mixing and vanishing energetic gain, $\Delta E_{gb} = 0$, when $D_0 = 0$, while mixing immediately sets in and $\Delta E_{gb} < 0$ once $D_0 \neq 0$. In all cases, ΔE_{gb} scales as $\Delta E_{gb} \sim -g_j(\Delta k)^2 D_0^2$, $g_j > 0$, as $D_0 \rightarrow 0$. While $g_j = g_j(w_0)$ depend on w_0 , we list ΔE_{gb} in the chiral limit, $w_0 = 0$, in Table III as it reveals some structure: we see that the energetic gain coming from hybridization is identical for the two IVCs and it is the same for the four states SP, VP, SLP₋, and SSLP₋ in the chiral limit.

The mixing matrices M^Q in Eq. (5.6) for these states provide some additional consistency checks between analytics and numerics. For instance, the IVC₊'s analysis closely parallels the one outlined above for the IVC₋: out of 16 symmetry-allowed generators, only two candidate combinations remain. Interestingly, their associated $\alpha_{\mathbf{k}}$ and $\beta_{\mathbf{k}}$ in Eq. (5.5) are the same (to leading order in D_0) as those of the two IVC₋ candidates. Having identified $M_{\mathbf{k}}^Q$ in Eq. (5.7a) as being dominant by comparison with numerics, we can read off which of the two analytical candidates for the IVC₊ must be realized; indeed, we

find the same one in numerics [compare Eq. (F59) and Fig. 17]. To provide a second example, we find a mixing matrix for the SLP₋ given by

$$M_{\mathbf{k}}^Q = 2i(\sigma_y + \sigma_z)P_{\pm}\varphi_{\mathbf{k}}, \quad \mathbf{k} \in \mathcal{A}_{\pm}. \quad (5.8)$$

This means that there is only mixing in the valley for which the graphene-like bands are at low energies and that both upper and lower graphene-like bands mix with both TBG-like bands. This is consistent with the $Q_{\mathbf{k}}$ in Fig. 6(b) we found for the SLP₋ in HF close to the K point.

C. Deforming the form factors

Apart from the mixing with and the ‘‘proximity-induced’’ order in the graphene-like bands, there are also important energetic contributions directly within the TBG-like bands, when tuning away from the chiral-flat-decoupled limit. While some of these contributions are analogous to TBG [46, 60], some others are not: non-zero D_0 strongly breaks P (see Table I) and, hence, induces terms in the TBG-like band structure, see second line in Eq. (2.6), without any analogue in TBG; it also leads to terms in the form factors [see Eq. (A33a) for details] that cannot be present in TBG and have not been studied in the literature.

We here begin with the impact of the form factors and restrict ourselves, for now, to the flat limit, $W_{\text{TBG}} = 0$. In the column labeled ΔE_{bb} in Table III we show the energy of the candidate states for non-zero w_0 and D_0

relative to the SP [see Appendix F 1 for explicit expressions for the prefactors $c_j > 0$]. First, for $D_0 = 0$, we recover the previous result [46] that the IVC₋ and VP are the only states besides the SP that are not penalized when turning on w_0 . Second, we see that this changes once $D_0 \neq 0$: if $w_0 = 0$, it is the SLP group of states that is not suppressed by breaking P with D_0 [63], while both IVCs have increasing energy. When both w_0 and D_0 are simultaneously non-zero, also the SLP group is suppressed (by the exact same amount as the IVC₋) compared to the SP and VP. Algebraically, this is related to the fact that the SP and VP order parameters are the only ones that commute with all form factors once $w_0, D_0 \neq 0$ [cf. Eq. (A33a)].

D. Finite TBG bandwidth

Finally, we take into account the finite bandwidth of the TBG-like bands by doing perturbation theory in W_{TBG} in Eq. (2.5), starting from the product states associated with the candidate orders of Table II. We will only outline the results here and refer the interested reader to Appendix F 2.

Since the correction to the relative energies of the different candidate orders vanishes to first order in W_{TBG} , we focus on second order perturbation theory, which, if non-zero, will always lower the energy of the states and can be thought of as “superexchange”. We consider the superexchange processes associated with all four terms, E_j^b , $j = 0, 1, 2, 3$, in Eq. (2.6) of the TBG-like dispersion; the corresponding energetic gain will scale as W_{TBG}^2/U for $j = 0$, $W_{\text{TBG}}^2 w_0^2/U$ for $j = 1$, $W_{\text{TBG}}^2 D_0^2/U$ for $j = 2$, and $W_{\text{TBG}}^2 w_0^2 D_0^2/U$ for $j = 3$ to leading order in D_0 and w_0 , where U is the energy scale associated with occupying a \mathbf{k} -state of an unoccupied flavor in a given ground state. Which of these four superexchange processes are “active” for the candidate states is listed in Table III. We find that, by virtue of being proportional to the identity in Eq. (2.6), E_3^b does not affect the energy of the states. Furthermore, we see that the displacement-induced superexchange process favors the IVC₊ and the SLP group of states.

E. Comparison of energetics

Taken together, the energetics obtained analytically as summarized in Table III agrees well with the numerics in several aspects: we can see that the energies of the SLP₊, SLP₋, and SSLP₋ are expected to be very close, exactly as seen in HF, see Fig. 6. There is a very small splitting between the three associated with the fact that the SSLP₋ cannot benefit from ordering in the graphene-like bands, that that effect is weaker for the SLP₊ than for

the SLP₋ (recall $g_- > g_+$, for small w_0), and that the SLP₊ is slightly suppressed by the Hartree term. We have seen that there is a small energetic preference towards the SLP₋ in the numerics as well. In addition, we can also read off from Table III that the SLP group of states should be preferred for small w_0 and large D_0 , which is consistently seen in our HF phase diagram in Fig. 4.

Notwithstanding the good agreement between the HF and the analytics, concerning the graphene-ordering, mixing between the bands, and the energetics, there are also differences. These can be traced back to the additional presence of remote bands, in particular, the bands with σ_h eigenvalues $+1$ just above (below) the almost flat TBG-like band $p = +$ ($p = -$) away from the K/K' points. While these are included in our HF numerics, they are not taken into account in the analytics. Most notably, we have seen that mixing between these bands and the almost flat TBG-like bands further lowers the energy of IVC₋ state relative to the SP and VP states.

F. Breaking $\text{SU}(2)_+ \times \text{SU}(2)_-$

Finally, we come back to the fact that certain Hund's partners of states are degenerate in the model we have focused on so far and defined in Sec. II; this can be traced back to the presence of the $\text{SU}(2)_+ \times \text{SU}(2)_-$ spin symmetry. Turning on a finite Hund's coupling, $J_H \neq 0$, will lift this degeneracy and also slightly affect the relative energetics of the candidate states. For the form of the Hund's coupling defined in Eq. (3.16), we have computed the respective change of energy. With more general and explicit expressions available in Appendix F 5, we present their impact on the energies of the states in the last two columns of Table III in the chiral-flat-decoupled limit. Here β_j are positive constants obeying $\beta_2 > \beta_1, \beta_3$. While estimates based on the Coulomb interaction yield an energy change (per unit cell) of $2J_H\beta_1 \simeq -0.2$ meV for the SP state in TBG [64], its actual effective value might be smaller due to screening processes and additional electron-phonon coupling, which can also change the sign of J_H .

However, irrespective of the sign of J_H and its precise magnitude, we see that it will favor the SSLP_± states in the otherwise almost degenerate SLP group of states, since the SLP_± cannot lower their energy with $J_H \neq 0$. This is due to the fact that SLP_± are their own Hund's partners. Similarly, the SP and SVP states will be favored over the VP phase. Finally, we note that the energetic gain coming from J_H is of the same order of magnitude (even larger) for the SIVC₋ (IVC₋) compared to the SP (SVP) state. Consequently, we do not expect that J_H will change the fact that the IVC₋ dominates over the SP in the phase diagram in Fig. 4.

VI. NUMERICS FOR $\nu = \pm 2$

In this section, we will discuss the numerical results for correlated states at half-filling of the lower or upper TBG-like bands, $\nu = \pm 2$.

A. Procedure and results

We find self-consistent HF solutions with energies lower than the symmetry-unbroken normal state where, at $D_0 = 0$ ($D_0 \neq 0$), the TBG (TBG-like and graphene-like) bands are spin polarized and, on top of this, exhibit any of the candidate orders defined for $\nu = 0$ in Table II. The obtained spin polarization is our explanation for the observed [30, 31] reduced flavor degeneracy setting in around $|\nu| = 2$.

To be more explicit, our self-consistent solutions are found by starting, say for $\nu = -2$, from a correlator $P_{\nu=-2}^b = \frac{1}{2}(\mathbb{1} + s_z)\frac{1}{2}(\mathbb{1} + Q^b)$ in the TBG-like subspace, where Q^b is any of the candidate orders in Table II; the initial correlator for the graphene-like sector is taken to be filled lower bands of the continuum model at charge neutrality, $P^g = \frac{1}{2}(\mathbb{1} - \sigma_z)$. The corresponding correlator at $\nu = +2$ in the TBG-like sector is simply given by $P_{\nu=+2}^b = \mathbb{1}_{8 \times 8} - P_{\nu=-2}^b$. We apply the same iterative procedure as for $\nu = 0$ to obtain self-consistent solutions.

We show the band structures resulting from our self-consistent calculation for the spin-polarized IVC₋ and spin-polarized SLP states in Fig. 7. The energies we obtain up to $D_0/w_1 = 0.5$ are shown in Appendix D, but are qualitatively similar to the numerical energies we computed at $\nu = 0$; we note though the spin-valley-polarized and spin-polarized IVC₋ are closer in energy at $\nu = 2$ than the SVP and IVC₋ state at $\nu = 0$. As can be seen in Fig. 7, the graphene bands at $D_0 = 0$ remain unpolarized, in agreement with the analytics in Sec. III. Once a finite displacement field is applied, also the graphene-like bands develop some small polarization. However, exactly as for $\nu = 0$, the Dirac cones of the IVC₋ state are not gapped out. We note the IVC₋, which was semimetallic at $\nu = 0$, is metallic at $\nu = \pm 2$ (albeit with small Fermi surfaces). This follows from the additional spin polarization at $|\nu| = 2$, breaking $SU(2)_s$ which guaranteed its semimetallic character at $\nu = 0$. In Fig. 7(a,b), we see that the SLP state, which was insulating at $\nu = 0$, becomes semimetallic at $|\nu| = 2$, since the spin-polarization of the TBG-like bands only allows for SLP order in one of the spin species of the graphene-like bands.

B. Connection to experiment

An important feature in the experimental data on MSTG [30, 31], which may be related to our numerics, is

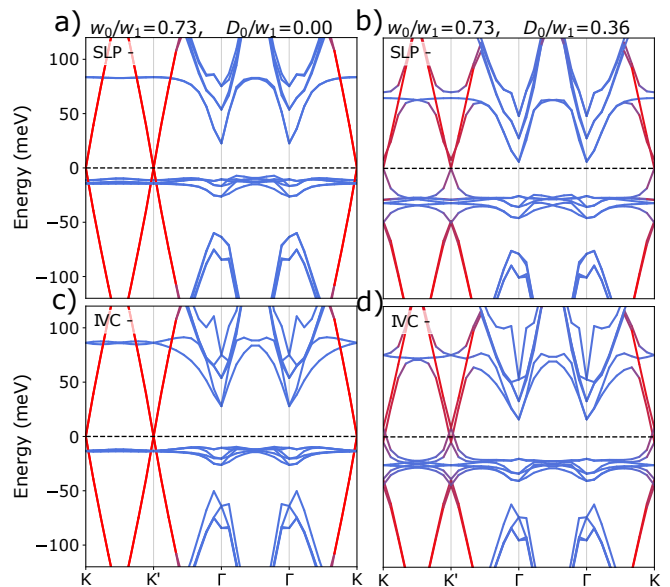


FIG. 7: HF band structures at $\nu = 2$ for the spin polarized SLP₋ state (a) at $D_0 = 0$ and (b) $D_0 > 0$ and spin polarized IVC₋ state (c) at $D_0 = 0$ and (d) $D_0 > 0$.

the observation of enhanced resistivity at $\nu = 0$ for both $D_0 = 0$ and $D_0 > 0$ and a state with high resistivity observed only at a finite value of D_0 at $\nu = 2$. Taking our results for band structures at $\nu = 0$ and $|\nu| = 2$ together, we note that at $\nu = 0$, the leading phase in our numerics is either a semimetallic IVC₋ state or an insulating SLP state, both of which may be compatible with the observed high resistivity state at $\nu = 0$. At $\nu = 2$, we find our leading solution is a metallic IVC₋ state for small values of D_0 and a semimetallic SLP state for large enough D_0 —a possible explanation for why high resistivity at $|\nu| = 2$ only sets in above a finite value of D_0 in experiment.

Another notable experimental observation is that the reset of the band structure at $|\nu| = 2$ splits into a Dirac-like feature at $|\nu| = 2$ and a van-Hove singularity, associated with a rapid increase and sign change in the Hall density, at $|\nu| = 2 - \delta(D_0)$, $\delta(D_0) \ll 1$ [30]. If we assume that the spin polarization we found persists for a finite range of $|\nu|$ below 2, the bandstructures in Fig. 7 provide a natural explanation: lowering the chemical potential, e.g., in Fig. 7(d), until it hits the lower, almost flat set of bands will lead to a Lifshitz transition where the hole pockets around the K and K' points merge. This could explain the observed behavior of the Hall density. We note that the tendency, visible in Fig. 7, that increasing D_0 pushes these almost flat bands away from the Fermi level at $|\nu| = 2$, is consistent with this feature only being visible at non-zero D_0 and the associated $\delta(D_0)$ increasing with $|D_0|$ in experiment [30]. Numerical results for larger values of D_0 are discussed in Appendix E.

VII. SUPERCONDUCTIVITY

Having established the nature of the correlated phases in MSTG at various filling fractions, we next study the consequences for the superconducting states.

A. Pairing in the presence of polarization

Let us begin with the range of electron filling, $2 < |\nu| < 3$, where superconductivity is most prominently observed in experiment [30–32]. We have seen in Sec. VI that spin polarization supplemented with IVC₋ order is favored for realistic parameters at $|\nu| = 2$. Because the associated reduction of the number of flavors in the normal state is seen in experiment over the entire or most of the superconducting range of ν , we assume that spin polarization and superconductivity co-exist at least in part of the phase diagram. As follows from the analysis in Ref. 27, where a classification of pairing in almost SU(2)₊ × SU(2)₋-symmetric graphene moiré systems in the presence and absence of flavor polarization can be found, the superconducting state has to be in a non-unitary triplet phase—irrespective of the precise pairing mechanism. We emphasize that this also holds if the additional IVC₋ ordering found in the HF at $|\nu| = 2$ co-exists with superconductivity in a finite range of ν : the bands above the Fermi level in the IVC₋ band structure in Fig. 7 still exhibit Kramers partners at momenta \mathbf{k} and $-\mathbf{k}$ with the same spin, due to the preserved spinless $\tilde{\Theta}$ symmetry; these degenerate states can form Cooper pairs with non-unitary triplet vector. Note that this would not be the case, e.g., for the VP state [shown in Fig. 13], which does not exhibit exactly degenerate energy levels at \mathbf{k} and $-\mathbf{k}$ above the Fermi level. Given the strong tendency of MSTG towards superconductivity, this VP phase is, thus, a less natural candidate order. Indeed, we find it to be subleading in our HF numerics, see Sec. VI.

It is important to note, however, that the spin polarization (and additional IVC₋ state) is only realized for $J_H < 0$. While this is expected to be the case if J_H stems entirely from the Coulomb interaction, its sign is unknown. For $J_H > 0$, we would, instead, obtain the SVP phase with additional SIVC₋ order (cf. Table II). In that case, the associated superconductor would also be the corresponding Hund’s partner, which is an admixture between a singlet and unitary triplet [27]. In the presence of a magnetic field there is a crucial difference between the two scenarios: while, for $J_H < 0$, the SP will just align with the Zeemann field and the superconductor will remain a non-unitary triplet, the two antiparallel spin polarizations in the two valleys of the SVP state will be canted gradually; a coexisting singlet-unitary-triplet superconductor will continuously transition into a non-unitary triplet [27]. For completeness, we demonstrate

this explicitly in Appendix G 2. Based on the estimate $J_H \lesssim 0.2$ meV [64], we obtain 3 T as the characteristic magnetic field strength of the transition.

B. Pairing without polarization

In experiment [30–32], there are also regions of ν and D_0 with superconductivity but without any signs of flavor polarization in the corresponding normal state (SC II in Fig. 1). We here discuss the nature and origin of these superconducting phases.

In light of recent experiments in TBG [51–53], which indicate that electron-phonon coupling plays an important role for pairing in graphene moiré systems, the picture proposed in Ref. 54 provides a very promising microscopic scenario: while electron-phonon coupling is likely important to stabilize superconductivity and crucially determines the critical temperature, it might predominantly mediate an SU(2)₊ × SU(2)₋ symmetric pairing interaction. This leads to the near degeneracy of singlet and triplet pairing. In the additional presence of flavor polarization, its structure determines whether singlet or triplet is realized, as discussed above. In the absence of polarization, electron-phonon coupling by itself can never favor triplet over singlet and at most make the two degenerate [65, 66]; however, additional Cooper-channel interactions coming from the fluctuation of particle-hole orders can determine whether the superconductor will be in a singlet or triplet state [54]. To analyze this, we will consider the action

$$S = S_0 + S_\phi + S_{\phi f} + S_{\text{phonon}}, \quad (7.1)$$

which consists of the bare non-interacting action S_0 associated with H_0 in Eq. (2.5) and S_ϕ , given by

$$S_\phi = \frac{1}{2} \int_q \phi_q^j [\chi^{-1}(i\Omega_n, \mathbf{q})]_{j,j'} \phi_{-q}^{j'}, \quad (7.2)$$

where ϕ_q^j is the set of real bosonic fields (labelled by j) describing the fluctuations of a given candidate particle-hole order in Table II. We use $q = (i\Omega_n, \mathbf{q})$ labeling bosonic Matsubara frequencies Ω_n and momentum \mathbf{q} , introduce the short-cut $\int_q \dots \equiv T \sum_{\Omega_n} \sum_{\mathbf{q}} \dots$, and denote the (fully renormalized, low-energy) susceptibility in the particle-hole channel under consideration by $\chi(i\Omega_n, \mathbf{q})$. The bosons ϕ_q^j are coupled to the low-energy electron fields $f_{k;n,\eta,s}$, with $k = (i\omega_n, \mathbf{k})$, via

$$S_{\phi f} = \int_q \int_k f_{k+q}^\dagger \lambda^j(\mathbf{k} + \mathbf{q}, \mathbf{k}) f_k \phi_q^j. \quad (7.3)$$

Note that the coupling vertex $\lambda^j(\mathbf{k} + \mathbf{q}, \mathbf{k})$ is in general a matrix in valley, spin, and band space. For instance, a minimal description of fluctuations of the IVC₋ state is given by the two-component boson ϕ_q^j , $j = x, y$, with

$[\chi(i\Omega_n, \mathbf{q})]_{j,j'} \propto \delta_{j,j'} / (\Omega_n^2 + c^2 \mathbf{q}^2 + \xi^{-2})$, where the coherence length ξ parametrizes the proximity to the critical point. Furthermore, $\lambda^j(\mathbf{k}, \mathbf{k}') = s_0 \eta_j f_{\mathbf{k}, \mathbf{k}'}$, where $f_{\mathbf{k}, \mathbf{k}'}$ is a matrix in band space ($p = \pm, t = b, g$) only, which obeys $f_{\mathbf{k}, \mathbf{k}'} = f_{\mathbf{k}', \mathbf{k}}^\dagger = -f_{\mathbf{k}', \mathbf{k}}^T$ due to Hermiticity and ΘC_{2z} symmetry. Finally, the last part $\mathcal{S}_{\text{phonon}}$ in Eq. (7.1) stands for the action of the phonons and their coupling to the electrons.

We integrate out the bosonic modes ϕ_q^j in the action of Eq. (7.1), leading to an interaction between the fermions f . In the saddle-point equations of the Cooper channel, this interaction can be viewed as a $SU(2)_+ \times SU(2)_-$ -symmetry breaking correction to the $SU(2)_+ \times SU(2)_-$ -symmetric interaction coming from the phonons in $\mathcal{S}_{\text{phonon}}$.

As shown in Ref. 54, whether this tips the balance towards singlet or triplet is determined by the behavior of the fluctuating modes ϕ_q^j under *spinful* time-reversal Θ_s : if the bosonic mode is even (odd) under Θ_s , as indicated by + (−) in the column Θ_s in Table II, it will generically favor singlet (triplet) over triplet (singlet) pairing—in the presence of fine-tuning or additional symmetries, the two might remain degenerate. We emphasize that this conclusion does not depend on microscopic details such the precise form of χ in Eq. (7.2) or of λ^j in Eq. (7.3). Due to $U(1)_v$ symmetry, the system also exhibits the spinless time-reversal symmetry $\tilde{\Theta}$, see Table I. We have generalized the analysis of Ref. 54 in Appendix G to also include this form of time-reversal symmetry and proved that any ϕ_q^j that is even (odd) under $\tilde{\Theta}$ will generically favor triplet (singlet). This also implies that any ϕ_q^j which has the same behavior under Θ_s and $\tilde{\Theta}$ will keep singlet and triplet degenerate. As expected, this is precisely the case for all order parameters in Table II that are their own Hund’s partner.

First, we focus on the superconducting domes (indicated by SC II in Fig. 1) that emerge in the vicinity of but outside the region with flavor polarization and reconstructed band structure [30]. As mentioned above, with Coulomb interactions only, the IVC_- order with additional spin polarization is favored over the Hund’s partner, the $SIVC_-$ with additional SVP order. We can read off from Table II that triplet will then be favored over singlet pairing. As shown in Ref. 27, the resulting triplet will be unitary within mean-field theory, but can become non-unitary when spin fluctuations corrections become significant. These two scenarios can be distinguished experimentally since the unitary triplet will exhibit a BKT transition (of a charge-4e, spin-rotation-invariant order parameter combination), while the non-unitary state will not. For the other sign of J_H , singlet pairing (if mean-field theory applies) or a mixed singlet-triplet phase (if the SVP fluctuations dominate) will be realized [27].

Second, we also comment on possible pairing phases close to charge neutrality, although these have not been

seen experimentally. For the same sign, $J_H < 0$, that favors SP at $\nu = 2$, we obtain the $SIVC_-$ at $\nu = 0$ as the dominant, semimetallic instability at small D_0 and large w_0 , see Fig. 4. As can be seen from Table II, it is even under Θ_s and will, hence, favor singlet pairing. Among the SLP group of states, realized for larger D_0 , only fluctuations of the $SSLP_\pm$ states will break the $SU(2)_+ \times SU(2)_-$ -symmetry; the $SSLP_-$, realized for $J_H < 0$, will also favor singlet.

Finally, if electron-phonon coupling does not play any role for pairing in MSTG, i.e., the term $\mathcal{S}_{\text{phonon}}$ in Eq. (7.1) can be neglected, we still obtain the same results (see Appendix G 1 for a derivation): if ϕ_q^j are even/odd (odd/even) under Θ_s ($\tilde{\Theta}$), singlet/triplet pairing will be favored. Consequently, the above statements about the singlet-triplet competition still apply in this scenario as well.

C. Relevance of Dirac cones and topology

Motivated by recent theoretical works [25, 55, 67] discussing the potential importance of WZW terms for superconductivity and insulating behavior in TBG, we here analyze under which conditions this can also be relevant for MSTG. We note that the WZW term is a purely kinematic term associating a Berry phase with spatio-temporal textures of the orders defined on the TBG-like bands, and is independent of the interactions between the electrons in these bands. The presence of such a topological term has crucial consequences for direct transitions between superconductivity and correlated insulators, without closing a single-particle gap at integer fillings. We explain in Appendix H how the following results can be formally derived from the exhaustive classification of WZW physics in TBG in Ref. 25 and focus here on the main picture and implications.

Although the main superconducting phase is found between $|\nu| = 2$ and $|\nu| = 3$ in experiment [30–32], we begin our analysis at the charge neutrality point, $\nu = 0$, with associated superconducting phase labelled SC III in Fig. 1. Since we expect superconductivity to survive finite displacement fields, we focus on Dirac cones that remain close to the Fermi level for $D_0 \neq 0$. Inspection of the band structure in Fig. 2(b) reveals that the effect of D_0 is to push the two Dirac cones of the graphene and TBG bands of valley $\eta = +$ ($\eta = -$) at K (K') away from the Fermi level, leaving only a single TBG-like Dirac cone of valley $\eta = -$ ($\eta = +$) at low energies. Denoting the fermionic fields by ψ_q of these two Dirac cones (per spin), their Hamiltonian reads as

$$H_0^D = \sum_{\mathbf{q}} \psi_{\mathbf{q}}^\dagger [\rho_x \mu_z s_0 q_x + \rho_y s_0 q_y] \psi_{\mathbf{q}}, \quad (7.4)$$

where \mathbf{q} is momentum measured relative to the K (K') point for the Dirac cone at “mini-valley” $\mu_z = +$ ($\mu_z =$

TABLE IV: Possible particle-hole, m_j , and superconducting order parameters, Δ , that can exhibit mutual WZW terms at the indicated filling fractions $\nu = 0$ and $|\nu| = 2$. We here use the Dirac notation of Eqs. (7.4) and (7.5). As in Ref. 25, moiré density wave (MDW) indicates that the state breaks moiré translational symmetry; A and B refer to the irreducible representations of the point group C_6 of the superconductor (SC). The MDW with singlet pairing row below corresponds to the first row of Table IV in Ref. 25, while the MDW with triplet SC corresponds to the first row of Table VII (or more explicitly, in lines 5 and 9 of Table XV).

$ \nu $	m_j	type	Δ	SC type
0, 2	$\mu_z \rho_z \mathbf{s}$	SSLP ₋	$\mathbb{1}$	A singlet
0, 2	$\rho_x(\mu_x, \mu_y); \rho_z$	MDW; SLP ₊	$\mathbb{1}$	A singlet
2	$\rho_z(s_x, s_y); \rho_z \mu_z s_z$	SSLP ₊ /SSLP ₋	$\mu_z s_z$	B unit. triplet
2	$\rho_x s_z(\mu_x, \mu_y); \rho_z$	MDW/SLP ₊	$\mu_z s_z$	B unit. triplet

–). Furthermore, ρ_j are Pauli matrices acting in the Dirac space, which are related to the band-space matrices σ_j used above. In this notation, C_{2z} and (anti-unitary) spinfull time-reversal act as $C_{2z} : \psi_{\mathbf{q}} \rightarrow \rho_x \mu_x \psi_{-\mathbf{q}}$ and $\Theta_s : \psi_{\mathbf{q}} \rightarrow T \psi_{-\mathbf{q}}$, $T = s_y \mu_x$, respectively.

Similar to Ref. 25, we ask what different types of particle-hole orders, m_j , and superconducting order parameters, Δ , coupling to $\psi_{\mathbf{q}}$ as

$$H_1^D = \sum_{\mathbf{q}, j} \psi_{\mathbf{q}}^\dagger m_j \psi_{\mathbf{q}} + \sum_{\mathbf{q}} \left(\psi_{\mathbf{q}}^\dagger \Delta T \psi_{-\mathbf{q}}^\dagger + \text{H.c.} \right), \quad (7.5)$$

can form WZW terms. The Hamiltonian $H_0^D + H_1^D$ is equivalent to the low-energy Dirac Hamiltonian for TBG projected into a valley-minivalley locked subspace, meaning that all pairings and insulating orders with WZW derived in [25] which survive projection to the same subspace will also be viable in MSTG. This is also the reason why the number of possible WZW terms we find here is significantly reduced as compared to Ref. 25.

Based on our analysis of particle-hole instabilities, the Dirac cones in Eq. (7.4) can either be those of the non-interacting bands or those of our leading instability at small D_0 —the semimetallic IVC₋ state with band structure shown in Fig. 3—which could persist for an extended range of $|\nu| > 0$; the resulting WZW terms we discuss next are identical in both scenarios.

Due to spin-rotation invariance, triplet pairing is not consistent with a WZW term [25]. We find that singlet pairing, $\Delta = \rho_0 \mu_0 s_0$, which transforms under the irreducible representation A of the point group, is the only possible superconducting state with WZW term at $\nu = 0$. This is consistent with the singlet we established in Sec. VII B near $\nu = 0$ due to electron-phonon coupling and particle-hole fluctuations. Considering all possible

compatible insulating orders m_j , we find there are two types: a spin Hall order of the form $\mu_z \rho_z s_{x,y,z}$, which aligns with our SSSLP₋ order in Table II, and a moiré density wave (MDW) state, $\rho_x \mu_{x,y}$, which breaks translations on the moiré lattice scale together with an SLP₊ state (ρ_z), see Table IV. We note that since this order is defined in the minivalley-valley locked space and, in the full space of MSTG, thus also breaks $U(1)_v$ symmetry in the same way as our IVC states. Interestingly, the SSSLP₋ is precisely our leading instability for larger D_0 and, hence, constitutes indeed a natural candidate m_j . The MDW state goes beyond our analysis in this work, as we have not considered states which break moiré translational symmetry. We will leave this for future work.

At $|\nu| = 2$, the situation is more complicated, because the entire band structure is reconstructed as a consequence of interaction-induced flavor polarization. Nonetheless, we expect the low-energy Dirac theory to be still of the form of Eq. (7.4). While there are other possible microscopic realizations, this is true for our leading instability of the SP (or SVP) with additional IVC₋ (or SIVC₋) order with band structure in Fig. 7(c,d). Neglecting the small D_0 -induced SP (or SVP) in the graphene-like bands, Eq. (7.4) still applies.

As summarized in Table IV, all options for insulators with singlet pairing then carry over to $|\nu| = 2$. A difference between the two cases is the broken $SU(2)_s$ spin symmetry, which also allows for triplet pairing. In this case, there are two additional options compatible with the unitary triplet $\Delta = \mu_z s_z$: one set with insulating orders $\rho_z s_{x,y}$ and $\rho_z \mu_z s_z$ and another set with $\rho_x s_z \mu_{x,y}$ and ρ_z . The first two orders correspond to our SSSLP₊ and SSSLP₋, respectively. The second two states are a spin polarized MDW state and the SLP₊ state. We find no non-unitary state to be consistent with WZW terms. Since only either singlet or unitary triplet are possible, the underlying flavor polarization of the TBG-like bands in Fig. 3(c,d) must be SVP rather than SP. Consequently, WZW terms at $|\nu| = 2$ are more likely relevant if $J_H > 0$.

VIII. CONCLUSION AND DISCUSSION

To summarize, we have studied particle-hole instabilities and superconductivity in MSTG, for different filling fractions ν and displacement field values D_0 , using a combination of analytical arguments and HF numerics. We started, in Sec. III, in the limit $D_0 = 0$, where the non-interacting band structure is just given by the spectrum of TBG and single-layer graphene. In the interacting Hamiltonian (3.1), these two subsystems are coupled by the density-density interaction in Eq. (3.4). We showed that states of the form $|\Psi_j^e(\nu_b)\rangle |\Psi_0^b(\nu_b)\rangle$, where $|\Psi_j^e(\nu_b)\rangle$ are the correlated semimetallic eigenstates of single-layer graphene and $|\Psi_0^b(\nu_b)\rangle$ are the eigenstates [60] in Eq. (3.6)

of the TBG Hamiltonian in the flat limit, are also exact eigenstates of the MSTG in Eq. (3.1). Furthermore, if $|\Psi_0^b(\nu_b)\rangle$ are groundstates of the TBG Hamiltonian, $|\Psi_j^g(\nu_b)\rangle$, $|\Psi_0^b(\nu_b)\rangle$ are shown to be exact groundstates of MSTG for a finite range of the coupling strength, quantified by λ in Eq. (3.4), between the two sectors. In this sense, not only the bare band structure but also the *interacting* physics of MSTG can be separated into that of single-layer graphene and TBG in the flat-decoupled limit.

We used these results to construct the set of candidate particle-hole orders summarized in Table II, which are exactly degenerate in the chiral-flat-decoupled limit [defined by further setting $w_0 = 0$ in Eq. (2.3)]. We took these states as our starting point of the HF numerics and analytical perturbation theory, which are not based on λ being small and allow to tune away from the chiral-flat-decoupled limit. The resulting band structures for $\nu = 0$ of the most important candidate orders are shown in Fig. 3 for zero and non-zero D_0 , with w_0 close to what is believed to describe the real system [56, 57] and finite TBG bandwidth. We see three distinct types of behavior, with different experimental signatures, when turning on D_0 : the IVC_{\pm} and their Hund’s partners $SIVC_{\pm}$ retain their semimetallic behavior for $D_0 \neq 0$, while the graphene Dirac cones of the SLP_{\pm} and $SSLP_{-}$ (and its Hund’s partner $SSLP_{+}$) are gapped out when turning on D_0 . Interestingly, for the VP and SP (and its Hund’s partner SVP), a finite D_0 induces small Fermi surfaces. In combination with further transport and, in particular, scanning tunneling microscopy experiments, which are sensitive to the local spectrum of the system, the computed spectra could help shed light on the correlated physics of MSTG.

Also the relative energetics between these candidate states as a function of D_0 , w_0 , and the TBG bandwidth is very rich: as summarized in Table III, there are many contributions without any analogue in TBG—the mixing between the TBG and graphene bands, additional ordering in the graphene bands, and the D_0 -induced breaking of symmetries in the TBG sector, which changes interaction matrix elements and induces new superexchange processes. The resultant phase diagram in the D_0 - w_0 plane with leading and subleading phases is presented in Fig. 4: at small D_0 , a semimetallic intervalley coherent phase is favored, which transitions into a sublattice polarized phase at larger D_0 ; among the latter set of states, we expect the $SSLP_{\pm}$ to dominate as they are the only states in this otherwise almost degenerate manifold that can benefit from the intervalley Hund’s coupling J_H in Eq. (3.16).

At $|\nu| = 2$, we found self-consistent HF solutions for all of the candidate states in Table II, which coexists with additional spin polarization (either of the SP or SVP type, depending on the sign of J_H). This can explain

the experimentally observed [30, 31] band resetting for $2 \lesssim |\nu| \lesssim 3$. For instance, the spectrum for the intervalley coherent state is shown in Fig. 7(c) and (d) for $\nu = 2$: increasing the filling fraction ν slightly will lead to Fermi surfaces of completely spin-polarized TBG-like bands, i.e., with half the number of flavors. In magnetic fields, the SP band resetting ($J_H < 0$) will not change its form, while the SVP-related resetting ($J_H > 0$) will continuously develop a finite canting and transform into a SP configuration; we estimate the associated magnetic-field scale to be of order of $3T$. Since $J_H < 0$ follows for pure Coulomb interactions [64], we expect the SP to be a more natural candidate; however, also $J_H > 0$ is possible, both theoretically and experimentally, and so we have studied both signs of J_H in our analysis.

Building on our results for the correlated normal states of MSTG, we analyzed the superconducting order parameters in Sec. VII in the different regimes indicated by SC I–III in Fig. 1: when superconductivity co-exists with flavor polarization (SC I), the nature of the pairing state depends crucially on the form of the flavor reduction; for the SP ($J_H < 0$) and SVP ($J_H > 0$) polarization, that we find around $|\nu| = 2$, we obtain, respectively, a non-unitary triplet and its Hund’s partner—a singlet-unitary-triplet admixed phase [27]. For the superconducting phases (SC II) close to but not in the flavor-polarized region, we find triplet pairing to dominate for $J_H < 0$, while the state will either be a singlet or admixed singlet-triplet phases for $J_H > 0$.

The behavior of these superconducting phases in in-plane magnetic fields, B_{\parallel} , follows from Ref. 27 where their respective phase diagrams in the temperature- B_{\parallel} -plane have been worked out: in the presence of SP polarization ($J_H < 0$), the critical temperature, T_c , of the non-unitary triplet is not affected by the Zeemann coupling and suppressed in quadratic order in the in-plane orbital coupling. This naturally explains the strong violation of the Pauli limit [32]. In the other case of SVP polarization ($J_H > 0$), the behavior of T_c of the associated singlet-triplet phase is the same, with the only difference that also the singlet-triplet admixture will gradually transform into a non-unitary triplet with increasing magnetic field. Understanding the reentrant superconducting behavior seen at even larger magnetic fields [32] will require understanding the fate of the correlated parent states, e.g., in Fig. 7, in large magnetic fields, which we leave for future work. Irrespective of whether the order parameter of SC II is a triplet or a singlet-triplet admixed phase, it will continuously transform into a non-unitary triplet [27] upon applying B_{\parallel} , while being eventually suppressed by the orbital coupling. We emphasize that all of the superconducting states we find, including the triplets, are protected [54] against non-magnetic impurities on the moiré scale, i.e., exhibit an analogue of the “Anderson theorem”, which is typically only expected

for singlet superconductors [68].

For completeness, we also investigated superconductivity close to the charge neutrality point (SC III), although not prominently seen in current experiments. Here we predict singlet pairing to dominate for the same sign of the Hund’s coupling, $J_H > 0$, we found triplet pairing near $\nu = 2$.

Because the band structure of MSTG also exhibits Dirac cones, a subset of which persist in most of our dominant particle-hole instabilities [see, e.g., IVC₋ in Fig. 3 and Fig. 7(c,d); related to the preserved $C_{2z}\Theta$ symmetry], we have also studied the possible WZW terms between superconducting and insulating orders; as discussed in [25, 55] for TBG, these topological terms are associated with and sensitive to the chirality of the Dirac cones in the normal-state band structure. We have shown here that the set of possibilities in MSTG is greatly reduced as compared to TBG [25], resulting from the reduced flavor degeneracy at $|\nu| \gtrsim 2$ and the impact of the displacement field at $\nu = 0$. As compiled in Table IV, only *s*-wave singlet pairing is consistent with WZW terms close to $\nu = 0$ (SC I), while both singlet and triplet pairing can have WZW terms in the flavor-polarized region (SC III). The corresponding particle-hole order parameters that form a mutual WZW term with these superconductors feature the sublattice-polarized states that we find to dominate at finite D_0 (in particular, the SSLP₋ state) and intervalley-coherent moiré density wave (MDW) phases.

Unlike all the states studied in the HF analyses of this paper, the MDW phases break translational symmetry at the scale of the moiré period. The MDW states were previously studied [25] in the context of WZW terms in TBG, break the moiré translational symmetry and the valley $U(1)_v$ symmetry. A closely related ‘incommensurate Kekulé spiral’ (IKS) appeared in a recent HF numerics study [69] of TBG. (We also note that Kekulé states have been observed in single layer graphene on a Cu substrate [70], and in the zeroth Landau level [71].) Motivated by these results, and our study of WZW terms here, we believe that these types of states are very promising possible additional instabilities in MSTG as well, in models which include the breaking of the $SU(2)_+ \times SU(2)_-$ symmetry to the physical $SU(2)_s$ spin rotation symmetry. We leave a detailed energetic study of these states for future work.

ACKNOWLEDGMENTS

M.S.S. thanks R. Samajdar for discussions and previous collaborations [27, 54] on unconventional superconductivity in moiré systems, and acknowledges discussions with T. Lang, A. Läuchli, and R. Fernandes. This research was supported by the National Science Foundation under Grant No. DMR-2002850. This work was also supported by the Simons Collaboration on Ultra-Quantum Matter, which is a grant from the Simons Foundation (651440, S.S.).

-
- [1] A. H. MacDonald, “Bilayer Graphene’s Wicked, Twisted Road,” *Physics* **12**, 12 (2019).
 - [2] E. Y. Andrei and A. H. MacDonald, “Graphene bilayers with a twist,” *Nat. Mater.* **19**, 1265 (2020).
 - [3] D. M. Kennes, M. Claassen, L. Xian, A. Georges, A. J. Millis, J. Hone, C. R. Dean, D. N. Basov, A. N. Pasupathy, and A. Rubio, “Moiré heterostructures as a condensed-matter quantum simulator,” *Nat. Phys.* **17**, 155 (2021).
 - [4] L. Balents, C. R. Dean, D. K. Efetov, and A. F. Young, “Superconductivity and strong correlations in moiré flat bands,” *Nat. Phys.* **16**, 725 (2020).
 - [5] M. S. Scheurer, “Spectroscopy of graphene with a magic twist,” *Nature* **572**, 40 (2019).
 - [6] T. Zaletel, “Stronger-correlated superconductivity in magic-angle twisted trilayer graphene,” *Journal Club in Condensed Matter Physics* (2021).
 - [7] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, “Correlated insulator behaviour at half-filling in magic-angle graphene superlattices,” *Nature* **556**, 80 (2018).
 - [8] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, “Unconventional superconductivity in magic-angle graphene superlattices,” *Nature* **556**, 43 (2018).
 - [9] C. Shen, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, Y. Chu, K. Watanabe, T. Taniguchi, R. Yang, Z. Y. Meng, D. Shi, and G. Zhang, “Correlated states in twisted double bilayer graphene,” *Nat. Phys.* **16**, 520 (2020).
 - [10] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, Y. Ronen, H. Yoo, D. H. Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, “Tunable spin-polarized correlated states in twisted double bilayer graphene,” *Nature* **583**, 221 (2020).
 - [11] Y. Cao, D. Rodan-Legrain, O. Rubies-Bigorda, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, “Tunable correlated states and spin-polarized phases in twisted bilayer–bilayer graphene,” *Nature* **583**, 215 (2020).
 - [12] G. W. Burg, J. Zhu, T. Taniguchi, K. Watanabe, A. H. MacDonald, and E. Tutuc, “Correlated Insulating States in Twisted Double Bilayer Graphene,” *Phys. Rev. Lett.* **123**, 197702 (2019).
 - [13] G. Chen, L. Jiang, S. Wu, B. Lyu, H. Li, B. L. Chittari, K. Watanabe, T. Taniguchi, Z. Shi, J. Jung, Y. Zhang,

- and F. Wang, “Evidence of a gate-tunable Mott insulator in a trilayer graphene moiré superlattice,” *Nat. Phys.* **15**, 237 (2019), [arXiv:1803.01985 \[cond-mat.mes-hall\]](#).
- [14] G. Chen, A. L. Sharpe, P. Gallagher, I. T. Rosen, E. J. Fox, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, J. Jung, Z. Shi, D. Goldhaber-Gordon, Y. Zhang, and F. Wang, “Signatures of tunable superconductivity in a trilayer graphene moiré superlattice,” *Nature* **572**, 215 (2019).
- [15] G. Chen, A. L. Sharpe, E. J. Fox, Y.-H. Zhang, S. Wang, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, Z. Shi, T. Senthil, D. Goldhaber-Gordon, Y. Zhang, and F. Wang, “Tunable correlated Chern insulator and ferromagnetism in a moiré superlattice,” *Nature* **579**, 56 (2020).
- [16] J. M. B. L. Dos Santos, N. M. R. Peres, and A. H. C. Neto, “Graphene bilayer with a twist: electronic structure,” *Phys. Rev. Lett.* **99**, 256802 (2007).
- [17] R. Bistritzer and A. H. MacDonald, “Moiré bands in twisted double-layer graphene,” *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12233 (2011).
- [18] J. M. B. L. Dos Santos, N. M. R. Peres, and A. H. C. Neto, “Continuum model of the twisted graphene bilayer,” *Phys. Rev. B* **86**, 155449 (2012).
- [19] A. Kerelsky, L. J. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, “Maximized electron interactions at the magic angle in twisted bilayer graphene,” *Nature* **572**, 95 (2019).
- [20] Y. Cao, D. Rodan-Legrain, J. M. Park, F. Noah Yuan, K. Watanabe, T. Taniguchi, R. M. Fernandes, L. Fu, and P. Jarillo-Herrero, “Nematicity and Competing Orders in Superconducting Magic-Angle Graphene,” *arXiv e-prints* (2020), [arXiv:2004.04148 \[cond-mat.mes-hall\]](#).
- [21] C. Rubio-Verdú, S. Turkel, L. Song, L. Klebl, R. Samajdar, M. S. Scheurer, J. W. F. Venderbos, K. Watanabe, T. Taniguchi, H. Ochoa, L. Xian, D. Kennes, R. M. Fernandes, Ángel Rubio, and A. N. Pasupathy, “Universal moiré nematic phase in twisted graphitic systems,” (2020), [arXiv:2009.11645 \[cond-mat.str-el\]](#).
- [22] R. Samajdar, M. Scheurer, S. Turkel, C. Rubio-Verdú, A. Pasupathy, J. Venderbos, and R. M. Fernandes, “Electric-field-tunable electronic nematic order in twisted double-bilayer graphene,” *2D Materials* **8** (2021).
- [23] U. Zondiner, A. Rozen, D. Rodan-Legrain, Y. Cao, R. Queiroz, T. Taniguchi, K. Watanabe, Y. Oreg, F. von Oppen, A. Stern, E. Berg, P. Jarillo-Herrero, and S. Ilani, “Cascade of phase transitions and dirac revivals in magic-angle graphene,” *Nature* **582**, 203 (2020).
- [24] D. Wong, K. P. Nuckolls, M. Oh, B. Lian, Y. Xie, S. Jeon, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, “Cascade of electronic transitions in magic-angle twisted bilayer graphene,” *Nature* **582**, 198 (2020).
- [25] M. Christos, S. Sachdev, and M. 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).
- [26] J. Kang, B. A. Bernevig, and O. Vafek, “Cascades between light and heavy fermions in the normal state of magic angle twisted bilayer graphene,” *arXiv e-prints*, [arXiv:2104.01145](#) (2021), [arXiv:2104.01145 \[cond-mat.str-el\]](#).
- [27] M. S. Scheurer and R. Samajdar, “Pairing in graphene-based moiré superlattices,” *Phys. Rev. Research* **2**, 033062 (2020).
- [28] C. Zhang, T. Zhu, S. Kahn, S. Li, B. Yang, C. Herbig, X. Wu, H. Li, K. Watanabe, T. Taniguchi, S. Cabrini, A. Zettl, M. P. Zaletel, F. Wang, and M. F. Crommie, “Visualizing delocalized correlated electronic states in twisted double bilayer graphene,” *Nature Communications* **12**, 2516 (2021).
- [29] X. Liu, C.-L. Chiu, J. Y. Lee, G. Farahi, K. Watanabe, T. Taniguchi, A. Vishwanath, and A. Yazdani, “Spectroscopy of a tunable moiré system with a correlated and topological flat band,” *Nature Communications* **12**, 2732 (2021).
- [30] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, “Tunable strongly coupled superconductivity in magic-angle twisted trilayer graphene,” *Nature* **590**, 249–255 (2021).
- [31] Z. Hao, A. M. Zimmerman, P. Ledwith, E. Khalaf, D. H. Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, “Electric field-tunable superconductivity in alternating-twist magic-angle trilayer graphene,” *Science* **371**, 1133–1138 (2021).
- [32] Y. Cao, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, “Large Pauli Limit Violation and Reentrant Superconductivity in Magic-Angle Twisted Trilayer Graphene,” *arXiv e-prints*, [arXiv:2103.12083](#) (2021), [arXiv:2103.12083 \[cond-mat.mes-hall\]](#).
- [33] E. Khalaf, A. J. Kruchkov, G. Tarnopolsky, and A. Vishwanath, “Magic angle hierarchy in twisted graphene multilayers,” *Physical Review B* **100** (2019), [10.1103/physrevb.100.085109](#).
- [34] S. Carr, C. Li, Z. Zhu, E. Kaxiras, S. Sachdev, and A. Kruchkov, “Ultraheavy and ultrarelativistic dirac quasiparticles in sandwiched graphenes,” *Nano Letters* **20**, 3030–3038 (2020).
- [35] C. Mora, N. Regnault, and B. A. Bernevig, “Flatbands and perfect metal in trilayer moiré graphene,” *Phys. Rev. Lett.* **123**, 026402 (2019).
- [36] D. Călugăru, F. Xie, Z.-D. Song, B. Lian, N. Regnault, and B. A. Bernevig, “Twisted symmetric trilayer graphene: Single-particle and many-body Hamiltonians and hidden nonlocal symmetries of trilayer moiré systems with and without displacement field,” *Phys. Rev. B* **103**, 195411 (2021), [arXiv:2102.06201 \[cond-mat.str-el\]](#).
- [37] J. Shin, B. L. Chittari, and J. Jung, “Stacking and gate tunable topological flat bands, gaps and anisotropic strip patterns in twisted trilayer graphene,” (2021), [arXiv:2104.01570 \[cond-mat.mes-hall\]](#).
- [38] C. Lei, L. Linhart, W. Qin, F. Libisch, and A. H. MacDonald, “Mirror symmetry breaking and lateral stacking shifts in twisted trilayer graphene,” (2021), [arXiv:2010.05787 \[cond-mat.mes-hall\]](#).
- [39] A. Ramires and J. L. Lado, “Emulating heavy fermions in twisted trilayer graphene,” *arXiv e-prints*, [arXiv:2102.03312](#) (2021), [arXiv:2102.03312 \[cond-mat.mes-hall\]](#).
- [40] Y. W. Choi and H. J. Choi, “Dichotomy of Electron-Phonon Coupling in Graphene Moire Flat Bands,” *arXiv e-prints*, [arXiv:2103.16132](#) (2021), [arXiv:2103.16132](#)

- [cond-mat.mes-hall].
- [41] E. Lake and T. Senthil, “Re-entrant Superconductivity Through a Quantum Lifshitz Transition in Twisted Trilayer Graphene,” arXiv e-prints , arXiv:2104.13920 (2021), arXiv:2104.13920 [cond-mat.supr-con].
- [42] W. Qin and A. H. MacDonald, “In-plane critical magnetic fields in magic-angle twisted trilayer graphene,” arXiv e-prints , arXiv:2104.14026 (2021), arXiv:2104.14026 [cond-mat.mes-hall].
- [43] A. Fischer, Z. A. H. Goodwin, A. A. Mostofi, J. Lischner, D. M. Kennes, and L. Klebl, “Unconventional Superconductivity in Magic-Angle Twisted Trilayer Graphene,” arXiv e-prints , arXiv:2104.10176 (2021), arXiv:2104.10176 [cond-mat.supr-con].
- [44] Y.-Z. Chou, F. Wu, J. D. Sau, and S. D. Sarma, “Correlation-induced triplet pairing superconductivity in graphene-based moiré systems,” (2021), arXiv:2105.00561 [cond-mat.supr-con].
- [45] M. Xie and A. MacDonald, “Nature of the correlated insulator states in twisted bilayer graphene,” *Physical Review Letters* **124** (2020), 10.1103/physrevlett.124.097601.
- [46] N. Bultinck, E. Khalaf, S. Liu, S. Chatterjee, A. Vishwanath, and M. P. Zaletel, “Ground state and hidden symmetry of magic angle graphene at even integer filling,” (2019), arXiv:1911.02045 [cond-mat.str-el].
- [47] S. Liu, E. Khalaf, J. Y. Lee, and A. Vishwanath, “Nematic topological semimetal and insulator in magic-angle bilayer graphene at charge neutrality,” *Physical Review Research* **3** (2021), 10.1103/physrevresearch.3.013033.
- [48] Y.-D. Liao, X.-Y. Xu, Z.-Y. Meng, and J. Kang, “Correlated insulating phases in the twisted bilayer graphene*,” *Chinese Physics B* **30**, 017305 (2021).
- [49] N. Bultinck, S. Chatterjee, and M. P. Zaletel, “Mechanism for anomalous hall ferromagnetism in twisted bilayer graphene,” *Physical Review Letters* **124** (2020), 10.1103/physrevlett.124.166601.
- [50] Y. Zhang, K. Jiang, Z. Wang, and F. Zhang, “Correlated insulating phases of twisted bilayer graphene at commensurate filling fractions: A hartree-fock study,” *Physical Review B* **102** (2020), 10.1103/physrevb.102.035136.
- [51] P. Stepanov, I. Das, X. Lu, A. Fahimniya, K. Watanabe, T. Taniguchi, F. H. L. Koppens, J. Lischner, L. Levitov, and D. K. Efetov, “Untying the insulating and superconducting orders in magic-angle graphene,” *Nature* **583**, 375 (2020).
- [52] X. Liu, Z. Wang, K. Watanabe, T. Taniguchi, O. Vafek, and J. I. A. Li, “Tuning electron correlation in magic-angle twisted bilayer graphene using coulomb screening,” *Science* **371**, 1261 (2021).
- [53] Y. Saito, J. Ge, K. Watanabe, T. Taniguchi, and A. F. Young, “Independent superconductors and correlated insulators in twisted bilayer graphene,” *Nature Physics* **16**, 926 (2020).
- [54] R. Samajdar and M. S. Scheurer, “Microscopic pairing mechanism, order parameter, and disorder sensitivity in moiré superlattices: Applications to twisted double-bilayer graphene,” *Phys. Rev. B* **102**, 064501 (2020).
- [55] E. Khalaf, S. Chatterjee, N. Bultinck, M. P. Zaletel, and A. Vishwanath, “Charged skyrmions and topological origin of superconductivity in magic-angle graphene,” *Science Advances* **7** (2021), 10.1126/sciadv.abf5299.
- [56] N. N. T. Nam and M. Koshino, “Lattice relaxation and energy band modulation in twisted bilayer graphene,” *Phys. Rev. B* **96**, 075311 (2017).
- [57] S. Carr, S. Fang, Z. Zhu, and E. Kaxiras, “Exact continuum model for low-energy electronic states of twisted bilayer graphene,” *Physical Review Research* **1**, 013001 (2019), arXiv:1901.03420 [cond-mat.mes-hall].
- [58] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, “Origin of magic angles in twisted bilayer graphene,” *Physical Review Letters* **122** (2019), 10.1103/physrevlett.122.106405.
- [59] B. A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, “Tbg iii: Interacting hamiltonian and exact symmetries of twisted bilayer graphene,” (2020), arXiv:2009.12376 [cond-mat.str-el].
- [60] B. Lian, Z.-D. Song, N. Regnault, D. K. Efetov, A. Yazdani, and B. A. Bernevig, “TBG IV: Exact Insulator Ground States and Phase Diagram of Twisted Bilayer Graphene,” arXiv e-prints , arXiv:2009.13530 (2020), arXiv:2009.13530 [cond-mat.str-el].
- [61] J. Kang and O. Vafek, “Strong coupling phases of partially filled twisted bilayer graphene narrow bands,” *Phys. Rev. Lett.* **122**, 246401 (2019).
- [62] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, “The electronic properties of graphene,” *Rev. Mod. Phys.* **81**, 109 (2009).
- [63] Note that the term c_2 , which comes from the Hartree term, turns out to be numerically very small.
- [64] S. Chatterjee, N. Bultinck, and M. P. Zaletel, “Symmetry breaking and skyrmionic transport in twisted bilayer graphene,” *Phys. Rev. B* **101**, 165141 (2020).
- [65] P. M. R. Brydon, S. Das Sarma, H.-Y. Hui, and J. D. Sau, “Odd-parity superconductivity from phonon-mediated pairing: Application to $\text{Cu}_x\text{Bi}_2\text{Se}_3$,” *Phys. Rev. B* **90**, 184512 (2014).
- [66] M. S. Scheurer, “Mechanism, time-reversal symmetry, and topology of superconductivity in noncentrosymmetric systems,” *Phys. Rev. B* **93**, 174509 (2016).
- [67] S. Chatterjee, M. Ippoliti, and M. P. Zaletel, “Skyrmion Superconductivity: DMRG evidence for a topological route to superconductivity,” arXiv e-prints , arXiv:2010.01144 (2020), arXiv:2010.01144 [cond-mat.str-el].
- [68] A. P. Mackenzie and Y. Maeno, “The superconductivity of Sr_2RuO_4 and the physics of spin-triplet pairing,” *Rev. Mod. Phys.* **75**, 657 (2003).
- [69] Y. H. Kwan, G. Wagner, T. Soejima, M. P. Zaletel, S. H. Simon, S. A. Parameswaran, and N. Bultinck, “Kekulé spiral order at all nonzero integer fillings in twisted bilayer graphene,” arXiv e-prints , arXiv:2105.05857 (2021), arXiv:2105.05857 [cond-mat.str-el].
- [70] C. Gutiérrez, C.-J. Kim, L. Brown, T. Schiros, D. Nordlund, E. B. Lochocki, K. M. Shen, J. Park, and A. N. Papaty, “Imaging chiral symmetry breaking from Kekulé bond order in graphene,” *Nature Physics* **12**, 950 (2016).
- [71] S.-Y. Li, Y. Zhang, L.-J. Yin, and L. He, “Scanning tunneling microscope study of quantum Hall isospin ferromagnetic states in the zero Landau level in a graphene monolayer,” *Phys. Rev. B* **100**, 085437 (2019), arXiv:1904.06902 [cond-mat.mes-hall].

[72] B. A. Bernevig, B. Lian, A. Cowsik, F. Xie, N. Regnault, and Z.-D. Song, “TBG V: Exact Analytic Many-Body Excitations In Twisted Bilayer Graphene Coulomb Hamiltonians: Charge Gap, Goldstone Modes and Absence of Cooper Pairing,” arXiv e-prints, arXiv:2009.14200 (2020), [arXiv:2009.14200](https://arxiv.org/abs/2009.14200) [[cond-mat.str-el](https://arxiv.org/abs/2009.14200)].

[73] This means that its adjacency graph is strongly connected. In our case, it corresponds to the situation that one can scatter between any pair of momenta via some number of virtual states; at least for $D_0 \neq 0$, this is generically expected to be the case.

Appendix A: Hamiltonian for the system

Here we will state the detailed form of the interacting Hamiltonian of MSTG used in our work, discuss the relevant symmetries and their representation in the basis employed in this paper, and their consequences for the form factors.

1. Continuum model and low-energy degrees of freedom

To describe the non-interacting bands, we use a continuum model, which is just the three-layer version of the continuum model of TBG [16–18]. Let $c_{\mathbf{k};\rho,l,\eta,s,\mathbf{G}}$ denote the operator annihilating an electron at crystalline momentum \mathbf{k} in the moiré Brillouin zone (MBZ), in sublattice $\rho = A, B$ and valley $\eta = \pm$ of the microscopic graphene sheets, of spin $s = \uparrow, \downarrow$, and with reciprocal lattice (RL) vector \mathbf{G} of the moiré lattice, thus, forming a triangular lattice; $\mathbf{G} = \sum_j n_j \mathbf{G}_j^M$, $n_j \in \mathbb{Z}$, with $\mathbf{G}_1^M = -\sqrt{3}k_\theta(1, \sqrt{3})^T/2$ and $\mathbf{G}_2^M = \sqrt{3}k_\theta(1, 0)^T$, where k_θ represents the magnitude of the vector connecting the K and K' points in the MBZ.

Before writing down the Hamiltonian, we follow [33] and perform a unitary transformation in layer space,

$$c_{\mathbf{k};\rho,l,\eta,s,\mathbf{G}} = V_{l,\ell} \psi_{\mathbf{k};\rho,\ell,\eta,s,\mathbf{G}}, \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad (\text{A1})$$

that decomposes the system into mirror-even ($\ell = 1, 2$) and mirror-odd ($\ell = 3$) subspaces (for $D_0 = 0$), as will become apparent below. The full continuum model reads as

$$H_0^{\text{Full}} = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\rho, \rho' = A, B} \sum_{\ell, \ell' = 1, 2, 3} \sum_{\eta = \pm} \sum_{s = \uparrow, \downarrow} \sum_{\mathbf{G}, \mathbf{G}' \in \text{RL}} \psi_{\mathbf{k};\rho,\ell,\eta,s,\mathbf{G}}^\dagger (h_{\mathbf{k},\eta})_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} \psi_{\mathbf{k};\rho',\ell',\eta,s,\mathbf{G}'}, \quad (\text{A2})$$

where $h_{\mathbf{k},\eta} = h_{\mathbf{k},\eta}^{(g)} + h_{\mathbf{k},\eta}^{(t)} + h_{\mathbf{k}}^{(D)}$ with the decoupled-layer graphene Hamiltonian

$$\left(h_{\mathbf{k},+}^{(g)} \right)_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} = \delta_{\ell,\ell'} \delta_{\mathbf{G},\mathbf{G}'} v_F (\rho_{\theta_\ell})_{\rho,\rho'} (\mathbf{k} + \mathbf{G} - (-1)^\ell \mathbf{q}_1/2), \quad (\text{A3})$$

$$\left(h_{\mathbf{k},-}^{(g)} \right)_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} = \left(h_{-\mathbf{k},+}^{(g)} \right)_{\rho,\ell,-\mathbf{G};\rho',\ell',-\mathbf{G}'}, \quad (\text{A4})$$

$\rho_\theta = e^{i\theta\rho_3/2} \rho e^{-i\theta\rho_3/2}$, \mathbf{q}_1 connecting the K and K' points in the MBZ. Throughout the appendix, $\mathbf{k} \in \text{MBZ}$ is measured relative to the M point of the moiré Brillouin zone. The contribution

$$\left(h_{\mathbf{k},+}^{(t)} \right)_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} = \sqrt{2} \begin{pmatrix} 0 & (T_{\mathbf{G}-\mathbf{G}'})_{\rho,\rho'} & 0 \\ (T_{\mathbf{G}'-\mathbf{G}}^*)_{\rho',\rho} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \left(h_{\mathbf{k},-}^{(t)} \right)_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} = \left(h_{-\mathbf{k},+}^{(t)} \right)_{\rho,\ell,-\mathbf{G};\rho',\ell',-\mathbf{G}'} \quad (\text{A5})$$

to the Hamiltonian describes the tunnelling that is modulated on the moiré lattice. We use the usual BM form [17],

$$T_{\delta\mathbf{G}} = \sum_{j=0,1,2} \delta_{\delta\mathbf{G}+\mathbf{A}_j,0} \left[w_0 \rho_0 + w_1 \begin{pmatrix} 0 & \omega^j \\ \omega^{-j} & 0 \end{pmatrix} \right], \quad (\text{A6})$$

$$\omega = e^{-i\frac{2\pi}{3}}, \quad \mathbf{A}_0 = 0, \quad \mathbf{A}_1 = \mathbf{G}_1, \quad \mathbf{A}_2 = \mathbf{G}_1 + \mathbf{G}_2. \quad (\text{A7})$$

Note that it holds $T_{\delta\mathbf{G}}^\dagger = T_{\delta\mathbf{G}}$ and $\rho_x T_{\delta\mathbf{G}} \rho_x = T_{\delta\mathbf{G}}^*$. So we see very clearly in this basis that the Hamiltonian is given by the sum of the TBG continuum model (with rescaled interlayer couplings) and a single graphene cone. This is different once a displacement field is applied, $D_0 \neq 0$; its contribution to the Hamiltonian reads as

$$\left(h_{\mathbf{k}}^{(D)} \right)_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} = -D_0 \delta_{\rho,\rho'} \delta_{\mathbf{G},\mathbf{G}'} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}_{\ell,\ell'}, \quad (\text{A8})$$

which is seen to couple the different mirror-eigenvalue sectors, as it breaks the mirror symmetry.

Here and in the following we use the same symbol with subscript $j = 0, 1, 2, 3$ for Pauli matrices and the associated quantum numbers: ρ_j are Pauli matrices in sublattice, s_j in spin, and η_j in valley space.

The Hamiltonian H_0^{Full} can be diagonalized by solving the eigenvalue problem

$$h_{\mathbf{k},\eta} u_{n,\eta}(\mathbf{k}) = \epsilon_n(\mathbf{k}) u_{n,\eta}(\mathbf{k}) \quad (\text{A9})$$

at every $\mathbf{k} \in \text{MBZ}$ and defining new fermionic operators $f_{\mathbf{k};n,\eta,s} = \sum_{\rho,\ell,\mathbf{G}} [u_{n,\eta}(\mathbf{k})]_{\rho,\ell,\mathbf{G}} \psi_{\mathbf{k};\rho,\ell,\eta,s,\mathbf{G}}$.

In this work, we will mostly focus on the four bands (for each spin and valley flavor) that are closest to the charge-neutrality point. Let us label them by the multi-index $n = (t, p)$ with $p = \pm$, $t = g, b$ indicating whether these bands are *graphene-like* or *TBG-like*. To make the association unique, we use the conventions

$$p \epsilon_{(t,p),\eta}(\mathbf{k}) \geq 0, \quad |\epsilon_{(b,p),\eta}(\mathbf{k})| \leq |\epsilon_{(g,p),\eta}(\mathbf{k})|. \quad (\text{A10})$$

In this way, the graphene-like (TBG-like) bands transition to the single-layer graphene (TBG) band structure for $D_0 \rightarrow 0$. But Eq. (A10) allows us to generalize this notion to non-zero D_0 . We will further introduce the notation

$$b_{\mathbf{k};p,\eta,s} := f_{\mathbf{k};(b,p),\eta,s}, \quad g_{\mathbf{k};p,\eta,s} := f_{\mathbf{k};(g,p),\eta,s} \quad (\text{A11})$$

on the level of electronic field operators. Our non-interacting, low-energy Hamiltonian then reads as

$$H_0 = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\eta=\pm} \sum_{s=\uparrow,\downarrow} \sum_{p=\pm} \left[W_{\text{TBG}} \epsilon_{(b,p),\eta}(\mathbf{k}) b_{\mathbf{k};p,\eta,s}^\dagger b_{\mathbf{k};p,\eta,s} + \epsilon_{(g,p),\eta}(\mathbf{k}) g_{\mathbf{k};p,\eta,s}^\dagger g_{\mathbf{k};p,\eta,s} \right] \quad (\text{A12})$$

and in total contains $2^4 = 16$ bands, associated with valley, spin, Dirac, and even-odd mirror space. In Eq. (A12), we introduced the dimensionless parameter $W_{\text{TBG}} \in [0, 1]$. In the real system, we have $W_{\text{TBG}} = 1$ and we define $W_{\text{TBG}} = 0$ as the *flat-band limit*. In Appendix F 2 below, we will compute the effect of the dispersion to second order in W_{TBG} .

2. Single-particle symmetries

As it will play an important role for our analysis of the correlated phases in the system, we will here list the relevant symmetries of the continuum-model Hamiltonian (A2) and their representation on our low-energy degrees of freedom $b_{\mathbf{k};p,\eta,s}$ and $g_{\mathbf{k};p,\eta,s}$ in Eq. (A12).

We distinguish four different classes of symmetries:

1. moiré translations, which we will assume to be preserved throughout and, hence, will not have to be considered explicitly here.
2. spin and charge conservation symmetries: since we do not take into account spin-orbit coupling nor the for small twist angles suppressed intervalley tunneling processes, the Hamiltonian is invariant under spin and $U(1)$ phase rotations in each valley $\eta = \pm$ separately.
3. point symmetries, forming the point group C_6 (for $D_0 \neq 0$) or C_{6h} (when $D_0 = 0$).
4. additional internal symmetries: the anti-unitary time-reversal symmetry Θ , the chiral symmetry C , and the unitary particle-hole symmetry P .

While all other symmetries leave H_0 invariant, the ‘‘anti-commuting symmetries’’ C and P send $H_0 \rightarrow -H_0$. Note that C is only an anti-commuting symmetry if $w_0 = 0$ in Eq. (A6) (but persists for $D_0 \neq 0$), while P is only present if $D_0 = 0$ (but persists for $w_0 \neq 0$). So the removal of the unitary particle hole symmetry in the TBG-like subspace upon turning on D_0 is a novel aspect of the trilayer system as it is always a symmetry in TBG (at least within the analogous continuum-model description). Among the point symmetries, the mirror symmetry σ_h is only present if $D_0 = 0$, similar to C_{2x} in TBG.

In Table V, we list the action of point and internal symmetries on the microscopic operators $\psi_{\mathbf{k}}$ in Eq. (A2) and on the low-energy operators $b_{\mathbf{k};p,\eta,s}$ and $g_{\mathbf{k};p,\eta,s}$. The latter representations have to be understood as a gauge-fixing condition; the gauge-fixing conventions we use here, closely follow [59]. For instance, we choose the phase of the wavefunctions (modulo remaining \pm sign) in Eq. (A9) such that $C_{2z}\Theta u_{n,\eta}(\mathbf{k}) = \rho_1[u_{n,\eta}(\mathbf{k})]^* = u_{n,\eta}(\mathbf{k})$.

To illustrate the gauge fixing and for later reference, we discuss it explicitly for $D_0 = 0$ in the mirror-odd sector bands ($\ell = 3$). As detailed in Appendix A 1, the Hamiltonian in this sector is simply that of single-layer graphene and, in our notation of Eq. (A3), given by

$$\left(h_{\mathbf{k},+}^{(e)}\right)_{\rho,\ell,\mathbf{G};\rho',\ell',\mathbf{G}'} = \delta_{\ell,3}\delta_{\ell',3}\delta_{\mathbf{G},\mathbf{G}'}v_F(\rho_{\theta_\ell})_{\rho,\rho'}(\mathbf{k} + \mathbf{G} + \mathbf{q}_1/2) \quad (\text{A13})$$

in valley $\eta = +$ (the wavefunctions in the other valley will simply follow from C_{2z} and Θ). Since Eq. (A13) is already diagonal in the \mathbf{G} indices, the wavefunctions simply follow from diagonalization of the structure in sublattice space. Focusing, as above, on the lowest bands, we have

$$\left(u_{(g,p),+}(\mathbf{k})\right)_{\rho,\ell,\mathbf{G}} = \delta_{\mathbf{G},-\mathbf{G}_{\mathbf{k}+\mathbf{q}_1/2}}\delta_{\ell,3}\frac{i^{(1-p)/2}}{\sqrt{2}}\begin{pmatrix} pe^{-i(\xi\delta\mathbf{k}-\theta)/2} \\ e^{i(\xi\delta\mathbf{k}-\theta)/2} \end{pmatrix}, \quad p = \pm 1, \quad e^{i\xi\mathbf{k}} = \frac{k_x + ik_y}{|\mathbf{k}|}, \quad \delta\mathbf{k} = \text{MBZ}(\mathbf{k} + \mathbf{q}_1/2), \quad (\text{A14})$$

where we defined the notation $\mathbf{G}_{\mathbf{k}}$ to indicate the reciprocal lattice vector that folds \mathbf{k} back to the MBZ, i.e., $\mathbf{k} - \mathbf{G}_{\mathbf{k}} \in \text{MBZ}$. Furthermore, we write $\text{MBZ}(\mathbf{k}) := \mathbf{k} - \mathbf{G}_{\mathbf{k}}$ for the momentum folded back into the MBZ.

Modulo a (in general \mathbf{k} and p dependent) minus sign to be discussed shortly, the phase of the wavefunctions has been chosen such that (‘‘real gauge’’)

$$C_{2z}\Theta u_{(g,p),+}(\mathbf{k}) \equiv \rho_1[u_{(g,p),+}(\mathbf{k})]^* = u_{n,+}(\mathbf{k}). \quad (\text{A15})$$

The relative minus sign between $p = +$ and $p = -$ has been fixed in Eq. (A14) to ensure the behavior,

$$Cu_{(g,p),+}(\mathbf{k}) \equiv -\rho_3Cu_{(g,p),+}(\mathbf{k}) = -p i u_{(g,-p),+}(\mathbf{k}), \quad (\text{A16})$$

under C given in Table V. Note that the relative minus sign at different \mathbf{k} is simply determined by continuity of $u_{(g,p),+}(\mathbf{k})$ as a function of \mathbf{k} (away from K/K').

Finally, we furthermore fixed the relative signs between wavefunctions at \mathbf{k} and $\mathbf{k} + \mathbf{G}_1$, $\mathbf{G}_1 \in \text{RL}$, to be $+1$ (‘‘periodic gauge’’),

$$\left(u_{(g,p),+}(\mathbf{k} + \mathbf{G}_1)\right)_{\rho,\ell,\mathbf{G}} = \left(u_{(g,p),+}(\mathbf{k})\right)_{\rho,\ell,\mathbf{G}+\mathbf{G}_1}, \quad (\text{A17})$$

which follows readily from Eq. (A14) by noting that $\mathbf{G}_{\mathbf{k}+\mathbf{G}_1} = \mathbf{G}_{\mathbf{k}} + \mathbf{G}_1$, for all $\mathbf{G}_1 \in \text{RL}$.

With these symmetry representations at hand, it is very simple to see that the band structure of the TBG-like band in Eq. (A12) has the form,

$$\epsilon_{(b,p),\eta}(\mathbf{k}) = p E_0^b(\mathbf{k}) + \eta E_1^b(\mathbf{k}) + p \eta E_2^b(\mathbf{k}) + E_3^b(\mathbf{k}), \quad (\text{A18})$$

where E_1^b (E_2^b) can only be non-zero if the same holds for w_0 (D_0) such that C (P) is broken; $E_3^b \neq 0$ requires $w_0, D_0 \neq 0$ simultaneously. Furthermore, it holds $E_j^b(\mathbf{k}) = E_j^b(-\mathbf{k})$, for $j = 0, 3$, and $E_j^b(\mathbf{k}) = -E_j^b(-\mathbf{k})$, for $j = 1, 2$, due to Θ . Clearly, the band structure for finite D_0 shown in Fig. 2(b) can only be consistent with Eq. (A18) if $E_2^b(\mathbf{k}) \neq 0$, which is consistent with our observation that P must be broken when $D_0 \neq 0$.

For later reference we introduce the following notation: for any quantity $q(\{x_k\})$ (such as E_j^b here) that depends on a set of parameters x_k (here D_0 or w_0), we introduce \bar{q} via

$$q(\{x_k\}) = \bar{q}(\{x_k\}) \prod_k (x_k)^{n_k}, \quad \text{with maximal } n_k \text{ such that we still have } \lim_{x_k \rightarrow 0} |\bar{q}(\{x_k\})| < \infty. \quad (\text{A19})$$

TABLE V: Action of the symmetries of the continuum theory on the microscopic field operators ($\psi_{\mathbf{k}}$) and in the band basis ($b_{\mathbf{k}}$ and $g_{\mathbf{k}}$). Although redundant, we have added the relevant \mathbf{k} -local combinations $C_{2z}\Theta$ and $C_{2z}P$ of the symmetries for convenience of the reader. Note that our choice of action of P in the graphene-like subspace is not independent from the other symmetries (it is related to C , C_{2z} , combined with a valley rotation symmetry). As indicated, P (and, hence, $C_{2z}P$) only anti-commutes with the Hamiltonian if we neglect the rotation of the Pauli matrices ρ_{θ} in Eq. (A3), which becomes asymptotically valid for small twist angles. As in the main text, η_j act in valley, σ_j in band, s_j in spin, and ρ_j in sublattice space.

Symmetry S	unitary	$H_0 \rightarrow \pm H_0$	$S\psi_{\mathbf{k};\ell,\mathbf{G}}S^\dagger$	$Sb_{\mathbf{k}}S^\dagger$	$Sg_{\mathbf{k}}S^\dagger$	condition
C_{2z}	✓	+	$\eta_1\rho_1\psi_{-\mathbf{k};\ell,-\mathbf{G}}$	$\eta_1\sigma_0b_{-\mathbf{k}}$	$\eta_1\sigma_0g_{-\mathbf{k}}$	—
C_{3z}	✓	+	$e^{i\frac{2\pi}{3}}\rho_3\eta_3\psi_{C_{3z}\mathbf{k};\ell,C_{3z}\mathbf{G}}$	$\eta_0\sigma_0b_{C_{3z}\mathbf{k}}$	$\eta_0\sigma_0g_{C_{3z}\mathbf{k}}$	—
σ_h	✓	+	$(1, 1, -1)_\ell\psi_{\mathbf{k};\ell,\mathbf{G}}$	$\eta_0\sigma_0b_{\mathbf{k}}$	$-\eta_0\sigma_0g_{\mathbf{k}}$	$D_0 = 0$
Θ	✗	+	$\eta_1\psi_{-\mathbf{k};\ell,-\mathbf{G}}$	$\eta_1\sigma_0b_{-\mathbf{k}}$	$\eta_1\sigma_0g_{-\mathbf{k}}$	—
C	✓	—	$(1, 1, -1)_\ell\rho_3\psi_{\mathbf{k};\ell,\mathbf{G}}$	$\eta_3\sigma_2b_{\mathbf{k}}$	$\eta_3\sigma_2g_{\mathbf{k}}$	$w_0 = 0$
P	✓	—	$\begin{pmatrix} 0 & \eta_3 & 0 \\ -\eta_3 & 0 & 0 \\ 0 & 0 & i\rho_2\eta_2 \end{pmatrix}_{\ell,\ell'}\psi_{-\mathbf{k};\ell',-\mathbf{G}}$	$-i\eta_3\sigma_2b_{-\mathbf{k}}$	$i\eta_1\sigma_2g_{-\mathbf{k}}$	$D_0 = 0, \theta \rightarrow 0$
$C_{2z}\Theta$	✗	+	$\rho_1\psi_{\mathbf{k};\ell,\mathbf{G}}$	$\eta_0\sigma_0b_{\mathbf{k}}$	$\eta_0\sigma_0g_{\mathbf{k}}$	—
$C_{2z}P$	✓	—	$\begin{pmatrix} 0 & -i\eta_2\rho_1 & 0 \\ i\eta_2\rho_1 & 0 & 0 \\ 0 & 0 & -i\rho_3\eta_3 \end{pmatrix}_{\ell,\ell'}\psi_{\mathbf{k};\ell',\mathbf{G}}$	$\eta_2\sigma_2b_{\mathbf{k}}$	$i\sigma_2g_{\mathbf{k}}$	$D_0 = 0, \theta \rightarrow 0$

This allows us to organize our perturbation theory. For instance, Eq. (A18) can then be written in the more informative way

$$\epsilon_{(b,p),\eta}(\mathbf{k}) = p\bar{E}_0^b(\mathbf{k}) + \eta\bar{E}_1^b(\mathbf{k})w_0 + p\eta\bar{E}_2^b(\mathbf{k})D_0 + \bar{E}_3^b(\mathbf{k})w_0D_0. \quad (\text{A20})$$

Similarly, we can analyze the properties of the graphene-like bands. As anticipated above, P , when present, does not lead to additional constraints and one readily concludes that

$$\epsilon_{(g,p),\eta}(\mathbf{k}) = p\bar{E}_0^g(\mathbf{k};\eta) + \bar{E}_1^g(\mathbf{k};\eta)w_0D_0, \quad (\text{A21})$$

where we noticed that the graphene-like bands still have a chiral symmetry for $D_0 = 0$ even if $w_0 \neq 0$. Θ dictates that $\bar{E}_j^g(\mathbf{k};\eta) = \bar{E}_j^g(-\mathbf{k};-\eta)$. Note that, throughout this work (cf. Fig. 2), we will label the valleys such that $|\epsilon_{(g,p),+}(\mathbf{k})| < |\epsilon_{(g,p),-}(\mathbf{k})|$ for \mathbf{k} near the K point.

3. Interaction and form factors

We follow a variety of previous works on moiré graphene systems [45–50, 60] and study a density-density interaction

$$H_1^{\text{Full}} = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) : \rho_{\mathbf{q}}^{\text{Full}} \rho_{-\mathbf{q}}^{\text{Full}} :, \quad (\text{A22})$$

with $: \dots :$ indicating normal ordering and

$$\rho_{\mathbf{q}}^{\text{Full}} = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\rho,\ell,\eta,s,\mathbf{G}} \psi_{\text{MBZ}(\mathbf{k}+\mathbf{q});\rho,\ell,\eta,s,\mathbf{G}+\mathbf{G}_{\mathbf{k}+\mathbf{q}}}^\dagger \psi_{\mathbf{k};\rho,\ell,\eta,s,\mathbf{G}} \quad (\text{A23})$$

denoting the Fourier transform of the electron density, where, as above, $\mathbf{q} = \text{MBZ}(\mathbf{q}) + \mathbf{G}_{\mathbf{q}}$ with $\text{MBZ}(\mathbf{q}) \in \text{MBZ}$. Furthermore, $V(\mathbf{q})$ in Eq. (A22) is the momentum-space interaction potential. For our analytical analysis, we will not have to specify the precise form of $V(\mathbf{q})$ and will only use that $V(\mathbf{q}) = V(-\mathbf{q}) > 0$. For our numerical computations we use $V(\mathbf{q})$ as given in Eq. (2.9).

Similar to our discussion of the non-interacting theory, we transform to the eigenbasis of H_0 and only keep the lowest four bands per spin and valley; the resulting projected interaction then reads as

$$H_1 = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) : \rho_{\mathbf{q}} \rho_{-\mathbf{q}} :, \quad \rho_{\mathbf{q}} = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\eta=\pm} \sum_{s=\uparrow, \downarrow} \sum_{t, t'=b, g} \sum_{p, p'=\pm} f_{\text{MBZ}(\mathbf{k}+\mathbf{q});(t,p),\eta,s}^\dagger \left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{(t,p),(t',p')} f_{\mathbf{k};(t',p'),\eta,s}, \quad (\text{A24})$$

where all the microscopic details are encoded in the *form factors*

$$\left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{n,n'} := \sum_{\rho,\ell,\mathbf{G}} [u_{n,\eta}(\text{MBZ}(\mathbf{k}+\mathbf{q}))]_{\rho,\ell,\mathbf{G}+\mathbf{G}_{\mathbf{k}+\mathbf{q}}}^* [u_{n',\eta}(\mathbf{k})]_{\rho,\ell,\mathbf{G}}. \quad (\text{A25})$$

In order to make sure that the projected interaction respects (many-body) particle-hole symmetry, we follow [59] and first rewrite the unprojected interaction (A22) as

$$\tilde{H}_1^{\text{Full}} = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta \rho_{\mathbf{q}}^{\text{Full}} \delta \rho_{-\mathbf{q}}^{\text{Full}} + \dots, \quad \delta \rho_{\mathbf{q}}^{\text{Full}} = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\rho,\ell,\eta,s,\mathbf{G}} \left(\psi_{\text{MBZ}(\mathbf{k}+\mathbf{q});\rho,\ell,\eta,s,\mathbf{G}+\mathbf{G}_{\mathbf{k}+\mathbf{q}}}^\dagger \psi_{\mathbf{k};\rho,\ell,\eta,s,\mathbf{G}} - \frac{1}{2} \delta_{\mathbf{q},0} \right), \quad (\text{A26})$$

where \dots just represent irrelevant constants and chemical potential terms. The wavefunctions in Eq. (A9) obey the completeness relation

$$\sum_n [u_{n,\eta}(\mathbf{k})]_{\rho_1,\ell_1,\mathbf{G}_1} [u_{n,\eta}(\mathbf{k})]_{\rho_2,\ell_2,\mathbf{G}_2}^* = \delta_{\rho_1,\rho_2} \delta_{\ell_1,\ell_2} \delta_{\mathbf{G}_1,\mathbf{G}_2}. \quad (\text{A27})$$

Setting $\rho_1 = \rho_2 = \rho$, $\ell_1 = \ell_2 = \ell$, $\mathbf{G}_1 = \mathbf{G}$, and $\mathbf{G}_2 = \mathbf{G} + \mathbf{G}'$ and summing both sides of Eq. (A27) over ρ , ℓ , and \mathbf{G} , we get

$$\sum_n \left(F_{\mathbf{k},\mathbf{G}'}^\eta \right)_{n,n} = \sum_{\rho,\ell,\mathbf{G}} \delta_{\mathbf{G}',0}, \quad (\text{A28})$$

where the sum over n is over *all* bands. Using this property we rewrite the full density in Eq. (A26) as

$$\delta \rho_{\mathbf{q}}^{\text{Full}} = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{n,n',\eta,s} \left(f_{\text{MBZ}(\mathbf{k}+\mathbf{q});n,\eta,s}^\dagger \left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{n,n'} f_{\mathbf{k};n',\eta,s} - \frac{1}{2} \delta_{n,n'} \sum_{\mathbf{G} \in \text{RL}} \left(F_{\mathbf{k},\mathbf{G}}^\eta \right)_{n,n'} \delta_{\mathbf{q},\mathbf{G}} \right). \quad (\text{A29})$$

So far, this is still an exact rewriting of the full interaction. However, in this form, the projection of the interaction to the low-energy bands (graphene-like and TBG-like) in a way that respects particle-hole symmetry is particularly natural: one simply has to restrict the sum over n, n' in Eq. (A29) to those bands. Defining for notational convenience

$$(F_{\mathbf{k},\mathbf{q}})_{((t,p),\eta,s),((t',p'),\eta',s')} = \delta_{s,s'} \delta_{\eta,\eta'} \left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{(t,p),(t',p')}, \quad (\text{A30})$$

we then arrive at the final form of the projected interaction

$$\tilde{H}_1 = \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta \rho_{\mathbf{q}} \delta \rho_{-\mathbf{q}}, \quad \delta \rho_{\mathbf{q}} = \sum_{\mathbf{k} \in \text{MBZ}} \left(f_{\text{MBZ}(\mathbf{k}+\mathbf{q})}^\dagger F_{\mathbf{k},\mathbf{q}} f_{\mathbf{k}} - \frac{1}{2} \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} \text{tr}(F_{\mathbf{k},\mathbf{G}}) \right). \quad (\text{A31})$$

Let us next discuss properties of these form factors. While we have computed these form factors numerically using the continuum model, we can make analytical progress by exploiting constraints on them. First, Hermiticity, $\rho_{\mathbf{q}}^\dagger = \rho_{-\mathbf{q}}$, implies

$$\left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{n',n}^* = \left(F_{\text{MBZ}(\mathbf{k}+\mathbf{q}),-\mathbf{q}}^\eta \right)_{n,n'} \quad \text{and, specifically,} \quad \left(F_{\mathbf{k},\mathbf{G}}^\eta \right)_{n',n}^* = \left(F_{\mathbf{k},-\mathbf{G}}^\eta \right)_{n,n'}. \quad (\text{A32})$$

Furthermore, since the density operator $\rho_{\mathbf{q}}^{\text{Full}}$ and, by design, its projection $\rho_{\mathbf{q}}$ transforms as a scalar function of the vector \mathbf{q} under all regular and anti-commuting symmetries of H_0 in Table V, we obtain further constraints on the form factors. Using all \mathbf{k} -local operators listed in Table V one finds after straightforward algebra:

$$\begin{aligned} \left(F_{\mathbf{k},\mathbf{q}}^\eta \right)_{(b,p),(b,p')} &= \bar{F}_1^{bb}(\mathbf{k},\mathbf{q})(\sigma_0)_{p,p'} + i \bar{F}_2^{bb}(\mathbf{k},\mathbf{q})(\sigma_2)_{p,p'} \\ &+ [\bar{F}_3^{bb}(\mathbf{k},\mathbf{q})\eta(\sigma_0)_{p,p'} + i \bar{F}_4^{bb}(\mathbf{k},\mathbf{q})\eta(\sigma_2)_{p,p'}] D_0^2 \\ &+ [\bar{F}_5^{bb}(\mathbf{k},\mathbf{q})\eta(\sigma_1)_{p,p'} + \bar{F}_6^{bb}(\mathbf{k},\mathbf{q})\eta(\sigma_3)_{p,p'}] w_0 \\ &+ [\bar{F}_7^{bb}(\mathbf{k},\mathbf{q})(\sigma_1)_{p,p'} + \bar{F}_8^{bb}(\mathbf{k},\mathbf{q})(\sigma_3)_{p,p'}] w_0 D_0^2, \end{aligned} \quad (\text{A33a})$$

$$\begin{aligned} \left(F_{\mathbf{k},\mathbf{q}}^\eta\right)_{(g,p),(g,p')} &= \bar{F}_1^{gg}(\mathbf{k},\mathbf{q})(\sigma_0)_{p,p'} + i\bar{F}_2^{gg}(\mathbf{k},\mathbf{q})(\sigma_2)_{p,p'} + \bar{F}_3^{gg}(\mathbf{k},\mathbf{q})\eta(\sigma_0)_{p,p'} + i\bar{F}_4^{gg}(\mathbf{k},\mathbf{q})\eta(\sigma_2)_{p,p'} \\ &\quad + \left[\bar{F}_5^{gg}(\mathbf{k},\mathbf{q})\eta(\sigma_1)_{p,p'} + \bar{F}_6^{gg}(\mathbf{k},\mathbf{q})\eta(\sigma_3)_{p,p'} + \bar{F}_7^{gg}(\mathbf{k},\mathbf{q})(\sigma_1)_{p,p'} + \bar{F}_8^{gg}(\mathbf{k},\mathbf{q})(\sigma_3)_{p,p'}\right] w_0 D_0^2, \end{aligned} \quad (\text{A33b})$$

$$\begin{aligned} \left(F_{\mathbf{k},\mathbf{q}}^\eta\right)_{(g,p),(b,p')} &= \left[\bar{F}_1^{gb}(\mathbf{k},\mathbf{q})(\sigma_0)_{p,p'} + i\bar{F}_2^{gb}(\mathbf{k},\mathbf{q})(\sigma_2)_{p,p'} + \bar{F}_3^{gb}(\mathbf{k},\mathbf{q})\eta(\sigma_0)_{p,p'} + i\bar{F}_4^{gb}(\mathbf{k},\mathbf{q})\eta(\sigma_2)_{p,p'}\right] D_0 \\ &\quad + \left[\bar{F}_5^{gb}(\mathbf{k},\mathbf{q})\eta(\sigma_1)_{p,p'} + \bar{F}_6^{gb}(\mathbf{k},\mathbf{q})\eta(\sigma_3)_{p,p'} + \bar{F}_7^{gb}(\mathbf{k},\mathbf{q})(\sigma_1)_{p,p'} + \bar{F}_8^{gb}(\mathbf{k},\mathbf{q})(\sigma_3)_{p,p'}\right] w_0 D_0, \end{aligned} \quad (\text{A33c})$$

while $(F_{\mathbf{k},\mathbf{q}}^\eta)_{(b,p),(g,p')}$ follows from Eq. (A32), i.e.,

$$\bar{F}_{j,\eta}^{gb}(\mathbf{k},\mathbf{q}) = \bar{F}_{j,\eta}^{bg}(\text{MBS}(\mathbf{k}+\mathbf{q}),-\mathbf{q}). \quad (\text{A34})$$

In Eq. (A33), all expansion co-efficients are real, $F_j^{tt'}(\mathbf{k},\mathbf{q}) \in \mathbb{R}$, and we again employed the convenient notation defined in Eq. (A19) to make their leading parameter dependence visible. While, by construction, all of the terms in $F_{(g,p),(g,p')}^\eta$ and $F_{(g,p),(b,p')}^\eta$ are absent in the previously studied TBG system, also half of the terms (F_j^{bb} , $j = 3, 4, 7, 8$) entering $F_{(g,p),(g,p')}^\eta$ in Eq. (A33a) are unique to MSTG, as they can only be non-zero if $D_0 \neq 0$.

Note that F_j^{bb} , $j = 3, 4, 7, 8$, have to be even functions of D_0 at generic momenta and, hence, start at order D_0^2 , which can be seen by noting that σ_h transforms the continuum model Hamiltonian $h_{\mathbf{k},\eta}(D_0)$ to $h_{\mathbf{k},\eta}(-D_0)$; the eigenstates, $u_{n,\eta}(\mathbf{k}; D_0)$, in Eq. (A9) thus obey $\sigma_h u_{n,\eta}(\mathbf{k}; D_0) = \pm u_{n,\eta}(\mathbf{k}; -D_0)$. Alternatively, it is also clear from the structure of (non-degenerate) perturbation theory in D_0 : corrections to matrix elements of states that have the same eigenvalue under σ_h can only appear in even orders of perturbation theory. The same line of reasoning shows that all F_j^{gb} have to be odd functions of D_0 , as indicated in Eq. (A33c).

In addition, C_{2z} implies that

$$\begin{aligned} F_j^{tt'}(\mathbf{k},\mathbf{q}) &= F_j^{tt'}(-\mathbf{k},-\mathbf{q}), \quad j = 1, 2, 7, 8, \\ F_j^{tt'}(\mathbf{k},\mathbf{q}) &= -F_j^{tt'}(-\mathbf{k},-\mathbf{q}), \quad j = 3, 4, 5, 6. \end{aligned} \quad (\text{A35})$$

The Hermiticity constraint in Eq. (A32) further implies for the form factors with $t = t'$ that

$$\begin{aligned} F_j^{tt}(\mathbf{k},\mathbf{G}) &= F_j^{tt}(\mathbf{k},-\mathbf{G}), \quad j = 1, 3, 5, 6, 7, 8, \\ F_j^{tt}(\mathbf{k},\mathbf{G}) &= -F_j^{tt}(\mathbf{k},-\mathbf{G}), \quad j = 2, 4. \end{aligned} \quad (\text{A36})$$

For later reference, we combine this with Eq. (A35) above to obtain

$$\begin{aligned} F_j^{tt}(\mathbf{k},\mathbf{G}) &= F_j^{tt}(-\mathbf{k},\mathbf{G}), \quad j = 1, 4, 7, 8 \\ F_j^{tt}(\mathbf{k},\mathbf{G}) &= -F_j^{tt}(-\mathbf{k},\mathbf{G}), \quad j = 2, 3, 5, 6. \end{aligned} \quad (\text{A37})$$

Although we will not use C_{3z} constraints on the form factors explicitly in our analytical considerations, we note that it further forces the momentum dependence of all pre-factors to obey $F_j^{tt'}(\mathbf{k},\mathbf{q}) = F_j^{tt'}(C_{3z}\mathbf{k}, C_{3z}\mathbf{q})$.

Finally, we point out that $F_{\mathbf{k},\mathbf{q}=0}^\eta = \mathbb{1}$, for all $\mathbf{k} \in \text{MBZ}$, as readily follows from the definition (A25). As such, we have

$$F_j^{tt'}(\mathbf{k},\mathbf{q}=0) = \delta_{j,1}\delta_{t,t'}. \quad (\text{A38})$$

The form factors of the mirror-odd (graphene) subsystem for $D_0 = 0$ can be readily computed from the wavefunctions defined in Eq. (A14). One finds

$$\left(F_{\mathbf{k},\mathbf{q}}^+\right)_{(g,p),(g,p')} = \delta_{\mathbf{G}_{\mathbf{k}+\mathbf{q}+\frac{\mathbf{q}_1}{2}}, \mathbf{G}_{\mathbf{k}+\frac{\mathbf{q}_1}{2}}} \left[\delta_{p,p'} \cos\left(\frac{\xi_\delta(\mathbf{k}+\mathbf{q}) - \xi_\delta\mathbf{k}}{2}\right) + i(\sigma_y)_{p,p'} \sin\left(\frac{\xi_\delta(\mathbf{k}+\mathbf{q}) - \xi_\delta\mathbf{k}}{2}\right) \right]. \quad (\text{A39})$$

Note that this agrees with the symmetry-based structure given in Eq. (A33b). On top of this, we observe that the matrix elements obey

$$\left(F_{\mathbf{k},\mathbf{G}}^\eta\right)_{(g,p),(g,p')} = \delta_{\mathbf{G},0}\delta_{p,p'}. \quad (\text{A40})$$

We will use both the explicit form in Eq. (A39) as well as the special property (A40) of the graphene matrix elements below.

Appendix B: Exact statements about the interacting groundstates

In this appendix, we will discuss limits where exact many-body groundstates of MSTG can be derived. To this end, we will focus on $D_0 = 0$, where the *non-interacting* part of the Hamiltonian, H_0 , is just given by the sum of the TBG continuum model, H_0^b , and that of single-layer graphene, H_0^g . Importantly, the presence of mirror symmetry σ_h also leads to $F_j^{gb}(\mathbf{k}, \mathbf{q}) = F_j^{bg}(\mathbf{k}, \mathbf{q}) = 0$, see Eq. (A33c). Consequently, the density operator $\rho_{\mathbf{q}}$ in Eq. (A24) and its symmetrized cousin $\delta\rho_{\mathbf{q}}$ in Eq. (A31) are just given by the sum of the contributions from the graphene (g) and TBG (b) subsystem. Focusing on $\delta\rho$, we have $\delta\rho_{\mathbf{q}} = \delta\rho_{\mathbf{q}}^g + \delta\rho_{\mathbf{q}}^b$ with

$$\delta\rho_{\mathbf{q}}^g = \sum_{\mathbf{k} \in \text{MBZ}} \left(\sum_{\mu, \mu'} g_{\text{MBZ}(\mathbf{k}+\mathbf{q});\mu}^\dagger (F_{\mathbf{k},\mathbf{q}}^{gg})_{\mu, \mu'} g_{\mathbf{k};\mu'} - \frac{1}{2} \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} \text{tr}(F_{\mathbf{k}, \mathbf{G}}^{gg}) \right), \quad (\text{B1})$$

$$\delta\rho_{\mathbf{q}}^b = \sum_{\mathbf{k} \in \text{MBZ}} \left(\sum_{\mu, \mu'} b_{\text{MBZ}(\mathbf{k}+\mathbf{q});\mu}^\dagger (F_{\mathbf{k},\mathbf{q}}^{bb})_{\mu, \mu'} b_{\mathbf{k};\mu'} - \frac{1}{2} \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} \text{tr}(F_{\mathbf{k}, \mathbf{G}}^{bb}) \right), \quad (\text{B2})$$

where we introduced the multi-indices comprising band, valley, and spin, $\mu = (p, \eta, s)$, and the submatrix form factors in the respective subsystem, $(F_{\mathbf{k},\mathbf{q}}^{tt})_{(p,\eta,s),(p',\eta',s')} := \delta_{s,s'} \delta_{\eta,\eta'} (F_{\mathbf{k},\mathbf{q}}^{\eta})_{(t,p),(t,p')}$. Notwithstanding this simplification, the interaction in Eq. (A31) still couples the graphene and TBG subsystem and it is not clear how this will affect the groundstates. In particular, it is unclear whether one can still write down exact eigenstates (or even groundstates) of the combined system, as has been done in TBG [60], and whether the system will also exhibit order in the graphene subspace.

1. Turning on the coupling between the subsystems

To address these questions, we split the projected Hamiltonian of the system into three parts,

$$H_0 + \tilde{H}_1 = H^g + H^b + \lambda H^{\text{gb}}, \quad (\text{B3})$$

where

$$H^g = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\eta, p = \pm} \sum_{s = \uparrow, \downarrow} p E_0^g(\mathbf{k}; \eta) g_{\mathbf{k};p,\eta,s}^\dagger g_{\mathbf{k};p,\eta,s} + \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta\rho_{\mathbf{q}}^g \delta\rho_{-\mathbf{q}}^g \quad (\text{B4})$$

is just the usual Hamiltonian of graphene with Coulomb interaction, while

$$H^b = W_{\text{TBG}} \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\eta, p = \pm} \sum_{s = \uparrow, \downarrow} (p \bar{E}_0^b(\mathbf{k}) + \eta \bar{E}_1^b(\mathbf{k}) w_0) b_{\mathbf{k};p,\eta,s}^\dagger b_{\mathbf{k};p,\eta,s} + \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta\rho_{\mathbf{q}}^b \delta\rho_{-\mathbf{q}}^b \quad (\text{B5})$$

is the TBG Hamiltonian, consisting of the quasi-flat bands interacting via a projected density-density interaction. We further introduced the dimensionless parameter λ that allows us to adiabatically turn on the density-density interaction,

$$\lambda H^{\text{gb}} = \frac{\lambda}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta\rho_{-\mathbf{q}}^g \delta\rho_{\mathbf{q}}^b \quad (\text{B6})$$

between the two subsystems, with $\lambda = 1$ corresponding to the physical system.

Let us start with $\lambda = 0$, where the Hamiltonian is given by $H^g + H^b$. Being the sum of the two commuting subsystem Hamiltonians, its eigenstates are just given by all combinations of the individual eigenstates of graphene and TBG. As discussed in more detail in the main text, the ground state of the graphene subspace is a symmetry-unbroken correlated semimetal. We denote its ground state by $|\Psi_0^g(\nu_b)\rangle$ and its (gapless) excited states by $|\Psi_j^g(\nu_b)\rangle$, $j > 0$, where ν_b indicates that the filling fraction of the graphene sector is a function of the filling ν_b in the TBG sector. We will here focus on $\nu_b = 0, \pm 1, \pm 2, \pm 3, \pm 4$.

The exact ground states of certain limits of H^b in Eq. (B5) at integer filling fractions ν_b have been discussed in detail in [60] and we next use these insights, within our notation, and transfer them to MSTG. First, one can see from Eq. (A33a) that the TBG form factors become

$$\bar{F}_{\mathbf{k},\mathbf{q}}^{bb} = \bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}) \sigma_0 + i \bar{F}_2^{bb}(\mathbf{k}, \mathbf{q}) \sigma_2 \quad (\text{B7})$$

in the chiral-decoupled limit ($w_0 = D_0 = 0$). We can, hence, bring the form factors and density operators to a diagonal form by introducing another set of field-operators,

$$\delta\rho_{\mathbf{q}}^{\mathbf{b}} = \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\eta, c = \pm, s} \bar{F}_c^{bb}(\mathbf{k}, \mathbf{q}) \left(\tilde{b}_{\text{MBZ}(\mathbf{k}+\mathbf{q}); c, \eta, s}^{\dagger} \tilde{b}_{\mathbf{k}; c, \eta, s} - \frac{1}{2} \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} \right), \quad \tilde{b}_{\mathbf{k}; c, \eta, s} = U_{c, p} b_{\mathbf{k}; p, \eta, s}, \quad U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}, \quad (\text{B8})$$

where $\bar{F}_{\pm}^{bb}(\mathbf{k}, \mathbf{q}) = \bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}) \pm i\bar{F}_2^{bb}(\mathbf{k}, \mathbf{q})$. As is readily verified, the following set of states [60]

$$|\Psi_0^{\mathbf{b}}(\nu_b)\rangle = \prod_{\mathbf{k} \in \text{MBZ}} \prod_{c = \pm} \prod_{j_c = 1}^{\nu_c} \tilde{b}_{\mathbf{k}; c, \eta_{j_c}^c, s_{j_c}^c}^{\dagger} |0^{\mathbf{b}}\rangle \quad (\text{B9})$$

with $\nu_+ + \nu_- = \nu_b$ obey

$$\delta\rho_{\mathbf{q}}^{\mathbf{b}} |\Psi_0^{\mathbf{b}}(\nu_b)\rangle = \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} R_{\mathbf{G}} |\Psi_0^{\mathbf{b}}(\nu_b)\rangle, \quad R_{\mathbf{G}} = \nu_b \sum_{\mathbf{k}} F_1^{bb}(\mathbf{k}, \mathbf{G}), \quad (\text{B10})$$

in the chiral limit, $w_0 = 0$, for arbitrary combinations of occupied spin/valley-flavors $\{\eta_j^{\pm}, s_j^{\pm}\}$. Consequently, all of these states are eigenstates of $H^{\mathbf{b}}$ in the flat-limit, i.e., when $W_{\text{TBG}} = 0$ (or right at the magic angle for $w_0 = 0$ where the bands become perfectly flat). It can further be shown [60] that all $|\Psi_0^{\mathbf{b}}(\nu_b)\rangle$ in Eq. (B9) are groundstates of $H^{\mathbf{b}}$ for $\nu_b = 0$ and that the same holds for all integer $\nu_b \neq 0$, if the flat-metric condition,

$$F_{\mathbf{k}, \mathbf{G}}^{bb} = \mathbb{1}f(\mathbf{G}), \quad \forall \mathbf{k}, \mathbf{G} \quad (\text{B11})$$

applies. As argued in Ref. 60, Eq. (B11) holds approximately for realistic parameters of TBG and the $|\Psi_0^{\mathbf{b}}(\nu_b)\rangle$ in Eq. (B9) will remain *exact* ground states if it is only weakly violated. The fact that there are several exact ground states is associated with an emergent enhanced continuous symmetry $[(\text{U}(4) \times \text{U}(4))_{\text{b,cf}}]$ in the chiral-flat limit [46, 59, 61].

Once $w_0 \neq 0$, i.e., in the nonchiral-flat limit, this symmetry group is reduced [to $\text{U}(4)$] and only the subset of states in Eq. (B9) where $\eta_j^+ = \eta_j^-$ and $s_j^+ = s_j^-$, which can only be defined for even ν_b , are exact eigenstates of $\delta\rho_{\mathbf{q}}^{\mathbf{b}}$ and obey Eq. (B10). They are always the exact ground states at $\nu_b = 0$, while this only holds if Eq. (B11) is not sufficiently violated for $\nu_b = \pm 2$. We refer to Ref. 60 for the derivations of these statements.

Having established the properties of the spectrum of the graphene subsystem and the exact groundstates of the TBG Hamiltonian, $H^{\mathbf{b}}$, we can next address what happens when the coupling λH^{gb} is turned on. As a first step, we rewrite the coupling Hamiltonian in Eq. (B6) as

$$H^{\text{gb}} = \frac{1}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\delta\rho_{-\mathbf{q}}^{\text{g}} \left(\delta\rho_{\mathbf{q}}^{\mathbf{b}} - \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} R_{\mathbf{G}} \right) + \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} R_{\mathbf{G}} \delta\rho_{-\mathbf{G}}^{\text{g}} \right]. \quad (\text{B12})$$

Recall from Eq. (A40) that $F_{\mathbf{k}, \mathbf{G}}^{gg} = \delta_{\mathbf{G}, 0} \mathbb{1}$, i.e., the graphene form factors obey the flat-metric condition (B11) exactly (with $f(\mathbf{G}) = \delta_{\mathbf{G}, 0}$) in the limit $D_0 = 0$, we have

$$\delta\rho_{\mathbf{G}}^{\text{g}} = N\delta_{\mathbf{G}, 0}\hat{\nu}_g, \quad \hat{\nu}_g = \frac{1}{N} \sum_{\mathbf{k}} \sum_{\mu} \left(g_{\mathbf{k}; \mu}^{\dagger} g_{\mathbf{k}; \mu} - \frac{1}{2} \right) \quad (\text{B13})$$

and

$$H^{\text{gb}} = H^{\text{gb}, 0} + V(0)\hat{\nu}_g, \quad H^{\text{gb}, 0} = \frac{1}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \delta\rho_{-\mathbf{q}}^{\text{g}} \left(\delta\rho_{\mathbf{q}}^{\mathbf{b}} - \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q}, \mathbf{G}} R_{\mathbf{G}} \right). \quad (\text{B14})$$

Note that the last term proportional to the filling fraction $\hat{\nu}_g$ of the graphene subspace is only a constant in the low-energy subspace of the system since the TBG ground states in Eq. (B9) are known to be insulators and, hence, exhibit a charge gap (in fact, also the two-particle spectrum was shown to be gapped [72]; more generally, this is also expected based on the experimental absence of superconductivity at integer filling in TBG in the vicinity of the magic angle, which implies that potential additional two-particle bound states cannot occur below the gapped particle-hole continuum).

Using the form of the coupling in Eq. (B14) and recalling that the exact groundstates of the TBG sector obey Eq. (B10), we immediately see that

$$H^{\text{gb},0} |\Psi_j^{\text{g}}\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle = 0 \quad (\text{B15})$$

for both the graphene ground state, $j = 0$, and its excited states, $j > 0$, and all exact TBG groundstates discussed above. As such, all of these states remain *exact* eigenstates for arbitrary λ .

Next, we address the question whether they remain the ground states for a finite range of $\lambda > 0$. First, as we have already discussed, the TBG sector has a finite charge gap and, hence, states which involve wavefunctions where one (or more) electrons of the TBG sector are transferred to the graphene sector can only become ground states at sufficiently large λ . Second, since the TBG ground states break symmetries, there are gapless Goldstone modes which require a more careful analysis. This is presented in the next subsection.

2. Behavior of Goldstone modes

As follows from C_{3z} rotational symmetry (and has been demonstrated in Ref. 72), the Goldstone spectrum of TBG must be of the form $E_G(\mathbf{p}) = \frac{1}{2}\rho\mathbf{p}^2 + \mathcal{O}(\mathbf{p}^4)$, with (isotropic) stiffness $\rho > 0$. Due to the arbitrarily small energies of the TBG Goldstone modes for $\mathbf{p} \rightarrow 0$ at $\lambda = 0$, we have to make sure that non-zero λ will not immediately lower the energy of these states below that of $|\Psi_0^{\text{g}}\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$ in MSTG. For that reason, we have to compute the evolution of the $\mathbf{p} \rightarrow 0$ Goldstone modes of the TBG sector upon turning on λ . Using momentum \mathbf{p} to label those states, we have to compute the behavior of their energy $\Delta E(\lambda, \mathbf{p})$ (relative to $|\Psi_0^{\text{g}}\rangle |\Psi_0^{\text{b}}(\nu_b)\rangle$). Assuming that $\Delta E(\lambda, \mathbf{p})$ is analytic in \mathbf{p} , C_{3z} again implies that

$$\Delta E(\lambda, \mathbf{p}) = \frac{1}{2}\rho\mathbf{p}^2 + m(\lambda) + \delta\rho(\lambda)\mathbf{p}^2 + \mathcal{O}(\mathbf{p}^4) \quad (\text{B16})$$

and our goal will be to determine $m(\lambda)$ and $\delta\rho(\lambda)$.

To get started, let us investigate the wavefunctions of the Goldstone modes of TBG for $\lambda = 0$. Fortunately, these can be computed exactly in the chiral-flat limit [72], as the property in Eq. (B10) reduces their computation to a one-particle problem. The mathematical reason for this is that $[H^{\text{b}}, \tilde{b}_{\mathbf{k}+\mathbf{p};c_2,\eta_2,s_2}^\dagger \tilde{b}_{\mathbf{k};c_1,\eta_1,s_1}] |\Psi_0^{\text{b}}(\nu_b)\rangle$ is a superposition of states of the form $\tilde{b}_{\mathbf{k}'+\mathbf{p};c_a,\eta_a,s_a}^\dagger \tilde{b}_{\mathbf{k}';c_b,\eta_b,s_b} |\Psi_0^{\text{b}}(\nu_b)\rangle$. Further noting that the density operator in Eq. (B8) does not scatter states between different quantum numbers c, η, s and is independent of η, s , we see that the Goldstone modes with momentum \mathbf{p} must be of the form

$$|\Psi_{\mathbf{p}}^{\text{b}}(\nu_b; c_1, \eta_1, s_1; c_2, \eta_2, s_2)\rangle = \sum_{\mathbf{k}} \alpha_{\mathbf{p},c_1,c_2}(\mathbf{k}) \tilde{b}_{\mathbf{k}+\mathbf{p};c_2,\eta_2,s_2}^\dagger \tilde{b}_{\mathbf{k};c_1,\eta_1,s_1} |\Psi_0^{\text{b}}(\nu_b)\rangle, \quad (\text{B17})$$

where the (c_1, η_1, s_1) flavor is occupied and (c_2, η_2, s_2) is unoccupied in the specific groundstate $|\Psi_0^{\text{b}}(\nu_b)\rangle$ we start from. In the limit $\mathbf{p} \rightarrow 0$, Ref. 72 has shown that one Goldstone zero mode has $c_1 = c_2$ and $\alpha_{\mathbf{p}=0,c,c}(\mathbf{k}) = \alpha$, i.e.,

$$|\Psi_{\mathbf{p} \rightarrow 0}^{\text{b}}(\nu_b; c, \eta_1, s_1; c, \eta_2, s_2)\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \tilde{b}_{\mathbf{k};c,\eta_2,s_2}^\dagger \tilde{b}_{\mathbf{k};c,\eta_1,s_1} |\Psi_0^{\text{b}}(\nu_b)\rangle. \quad (\text{B18})$$

All other zero-energy Goldstone modes follow from $(\text{U}(4) \times \text{U}(4))_{\text{b,cf}}$ rotations.

It is a matter of straightforward algebra to show that

$$\delta\rho_{\mathbf{q}}^{\text{b}} |\Psi_{\mathbf{p} \rightarrow 0}^{\text{b}}(\nu_b; c, \eta_1, s_1; c, \eta_2, s_2)\rangle = \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} R_{\mathbf{G}} |\Psi_{\mathbf{p} \rightarrow 0}^{\text{b}}(\nu_b; c, \eta_1, s_1; c, \eta_2, s_2)\rangle, \quad R_{\mathbf{G}} = \nu_b \sum_{\mathbf{k}} F_1^{bb}(\mathbf{k}, \mathbf{G}) \in \mathbb{R}. \quad (\text{B19})$$

Since $\delta\rho_{\mathbf{q}}^{\text{b}}$ is invariant under $(\text{U}(4) \times \text{U}(4))_{\text{b,cf}}$ rotations, this also holds for all other zero-energy Goldstone modes. In fact, this behavior can be understood more generally and intuitively by rewriting the H_b in Eq. (B5) (recall we always focus on the flat-band limit, $W_{\text{TBG}} = 0$) as

$$\begin{aligned} H^{\text{b}} &= \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \left(\delta\rho_{\mathbf{q}}^{\text{b}} - \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} R_{\mathbf{G}} \right) \left(\delta\rho_{-\mathbf{q}}^{\text{b}} - \sum_{\mathbf{G} \in \text{RL}} \delta_{-\mathbf{q},\mathbf{G}} R_{\mathbf{G}} \right) \\ &\quad + \frac{1}{N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) R_{\mathbf{G}} \delta\rho_{\mathbf{G}}^{\text{b}} - \frac{1}{2N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) R_{\mathbf{G}} R_{-\mathbf{G}}. \end{aligned} \quad (\text{B20})$$

While the last term in Eq. (B20) is just a constant, it holds

$$\delta\rho_{\mathbf{G}}^{\text{b}} = Nf(\mathbf{G})\hat{\nu}_b, \quad \hat{\nu}_b = \frac{1}{N} \sum_{\mathbf{k}} \sum_{\mu} \left(b_{\mathbf{k};\mu}^{\dagger} b_{\mathbf{k};\mu} - \frac{1}{2} \right), \quad (\text{B21})$$

if the flat-metric condition (B11) applies. Since $R_{\mathbf{G}} = R_{-\mathbf{G}}$, following from Eq. (A36), and as ν_b is constant in the low-energy subspace of TBG, we can write in this subspace [61]

$$H^{\text{b}} = \frac{1}{2N} \sum_{\mathbf{q}} \mathcal{O}_{\mathbf{q}}^{\dagger} \mathcal{O}_{\mathbf{q}} + \text{const.}, \quad \mathcal{O}_{\mathbf{q}} := \sqrt{V(\mathbf{q})} \left(\delta\rho_{\mathbf{q}}^{\text{b}} - \sum_{\mathbf{G} \in \text{RL}} \delta_{\mathbf{q},\mathbf{G}} R_{\mathbf{G}} \right). \quad (\text{B22})$$

Since all exact TBG groundstates discussed above obey Eq. (B10), or $\mathcal{O}_{\mathbf{q}} |\Psi_0^{\text{b}}(\nu_b)\rangle = 0$, all Goldstone modes with vanishing momentum, $\mathbf{p} \rightarrow 0$, must be annihilated by all $\mathcal{O}_{\mathbf{q}}$ as their energy must approach that of the groundstates. This is why Eq. (B19) must hold. Note that, for $\nu_b = 0$, we did not have to assume the flat-metric condition since $R_{\mathbf{G}} = 0$.

Having established Eq. (B19), we also immediately see that

$$H^{\text{gb},0} |\Psi_j^{\text{g}}\rangle |\Psi_{\mathbf{p} \rightarrow 0}^{\text{b}}(\nu_b; c_1, \eta_1, s_1; c_2, \eta_2, s_2)\rangle = 0. \quad (\text{B23})$$

This shows that also $|\Psi_j^{\text{g}}\rangle |\Psi_{\mathbf{p} \rightarrow 0}^{\text{b}}(\nu_b; c_1, \eta_1, s_1; c_2, \eta_2, s_2)\rangle$ remain exact eigenstates of the Hamiltonian for any $\lambda \neq 0$ and that $m(\lambda) = 0$ in Eq. (B16). Making the natural assumption that $\delta\rho(\lambda)$ is analytic in λ , we can expand

$$\delta\rho(\lambda) \sim \sum_{n \geq 0} c_n^{\rho} \lambda^n. \quad (\text{B24})$$

Therefore, it will require a finite value of λ before $\Delta E(\lambda, \mathbf{p})$ can lose its positive semi-definite nature and the groundstates discussed above have to remain the groundstates in a finite region of $\lambda > 0$.

As we discuss in Sec. VIII of the main text, one possibility is that at sufficiently large λ a state with finite momentum \mathbf{p} resulting from a mixing between the graphene and TBG degrees of freedom becomes the ground state. We leave a quantitative analysis of such a MDW state [25] for future work.

Appendix C: Hartree-Fock functional

1. General form

Let us assume that the ground state is the Slater determinant $|\psi[P_{\mathbf{k}}]\rangle$ characterized by the correlator

$$(P_{\mathbf{k}})_{\alpha, \alpha'} = \langle \psi[P_{\mathbf{k}}] | f_{\mathbf{k};\alpha}^{\dagger} f_{\mathbf{k};\alpha'} | \psi[P_{\mathbf{k}}] \rangle, \quad P_{\mathbf{k}}^{\dagger} = P_{\mathbf{k}} = (P_{\mathbf{k}})^2, \quad (\text{C1})$$

where we use the multi-index $\alpha = ((t, p), \eta, s)$ to keep the expressions more compact. The HF mean-field Hamiltonian H^{MF} associated with $H_0 + H_1$, defined in Eqs. (A12) and (A24) contains four terms,

$$H^{\text{MF}}[P_{\mathbf{k}}] = H_{\mathbf{k}}^{\text{MF}} + H_{\text{H}}^{\text{MF}}[P_{\mathbf{k}}] + H_{\text{F}}^{\text{MF}}[P_{\mathbf{k}}] + E_0[P_{\mathbf{k}}]. \quad (\text{C2})$$

Postponing the definition of the energetic off-set, E_0 , which is proportional to the identity operator, these terms are the kinetic term

$$H_{\mathbf{k}}^{\text{MF}} = H_0 = \sum_{\mathbf{k} \in \text{MBZ}} f_{\mathbf{k};\alpha}^{\dagger} h_{\alpha, \alpha'}^{\mathbf{k}}(\mathbf{k}) f_{\mathbf{k};\alpha'} \equiv \sum_{\mathbf{k} \in \text{MBZ}} f_{\mathbf{k}}^{\dagger} h^{\mathbf{k}}(\mathbf{k}) f_{\mathbf{k}}, \quad (\text{C3})$$

where $h^{\mathbf{k}}$ is the diagonal matrix

$$h_{((t,p),\eta,s),((t',p'),\eta',s')}^{\mathbf{k}}(\mathbf{k}) = \delta_{t,t'} \delta_{p,p'} \delta_{\eta,\eta'} \delta_{s,s'} \epsilon_{(t,p),\eta}(\mathbf{k}) \quad (\text{C4})$$

with $\epsilon_{(t,p),\eta}$ given in Eqs. (A20) and (A21), the Hartree term,

$$H_{\text{H}}^{\text{MF}}[P_{\mathbf{k}}] = \frac{1}{N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) \Gamma_{\mathbf{G}} \sum_{\mathbf{k} \in \text{MBZ}} f_{\mathbf{k}}^{\dagger} F_{\mathbf{k},\mathbf{G}} f_{\mathbf{k}}, \quad \Gamma_{\mathbf{G}} = \sum_{\mathbf{k} \in \text{MBZ}} \text{tr} [P_{\mathbf{k}} F_{\mathbf{k},\mathbf{G}}^*], \quad (\text{C5})$$

and the Fock contribution

$$H_{\text{F}}^{\text{MF}}[P_{\mathbf{k}}] = -\frac{1}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{\mathbf{k} \in \text{MBZ}} f_{\mathbf{k}}^{\dagger} F_{\mathbf{k},\mathbf{q}}^{\dagger} P_{\text{MBZ}(\mathbf{k}+\mathbf{q})}^T F_{\mathbf{k},\mathbf{q}} f_{\mathbf{k}}. \quad (\text{C6})$$

The associated ground state energy can be decomposed into three contributions,

$$E^{\text{HF}}[P_{\mathbf{k}}] = \langle \psi[P_{\mathbf{k}}] | (H_0 + H_1) | \psi[P_{\mathbf{k}}] \rangle = E_{\mathbf{k}}[P_{\mathbf{k}}] + E_{\text{H}}[P_{\mathbf{k}}] + E_{\text{F}}[P_{\mathbf{k}}], \quad (\text{C7a})$$

which read as

$$E_{\mathbf{k}}[P_{\mathbf{k}}] = \sum_{\mathbf{k} \in \text{MBZ}} \text{tr} [P_{\mathbf{k}}^T h^{\mathbf{k}}(\mathbf{k})], \quad (\text{C7b})$$

$$E_{\text{H}}[P_{\mathbf{k}}] = \frac{1}{2N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) \left| \sum_{\mathbf{k} \in \text{MBZ}} \text{tr} [P_{\mathbf{k}} F_{\mathbf{k},\mathbf{G}}^T] \right|^2, \quad (\text{C7c})$$

$$E_{\text{F}}[P_{\mathbf{k}}] = -\frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{\mathbf{k} \in \text{MBZ}} \text{tr} [P_{\mathbf{k}} F_{\mathbf{k},\mathbf{q}}^T P_{\text{MBZ}(\mathbf{k}+\mathbf{q})} F_{\mathbf{k},\mathbf{q}}^*]. \quad (\text{C7d})$$

Having established this notation, the energetic off-set E_0 in Eq. (C2) can now be conveniently stated as $E_0 = -E_{\text{H}} - E_{\text{F}}$.

As has been discussed in previous works on TBG [45–48], it is important to note that the continuum model defined in Sec. A 1 already takes into account some correlation effects of the system. So using the continuum model dispersion in $h^{\mathbf{k}}(\mathbf{k})$ in the HF mean-field Hamiltonian as defined in Eq. (C2) would constitute a double counting of these effects. To avoid this, we replace (ignoring the irrelevant constant $E_0[P_{\mathbf{k}}^0]$)

$$H_{\mathbf{k}}^{\text{MF}}[P_{\mathbf{k}}] \longrightarrow H_{\mathbf{k}}^{\text{MF}}[P_{\mathbf{k}}] - (H_{\text{H}}^{\text{MF}}[P_{\mathbf{k}}^0] + H_{\text{F}}^{\text{MF}}[P_{\mathbf{k}}^0]) \quad (\text{C8})$$

in Eq. (C2), where $P_{\mathbf{k}}^0$ is a reference density matrix for which the continuum model dispersion is expected to be valid. The redefinition in Eq. (C8) is equivalent to replacing

$$h^{\mathbf{k}}(\mathbf{k}) \longrightarrow \tilde{h}^{\mathbf{k}}(\mathbf{k}) = h^{\mathbf{k}}(\mathbf{k}) - h_{\text{HF}}[P_{\mathbf{k}}^0](\mathbf{k}), \quad (\text{C9})$$

in Eq. (C3), where we defined $h_{\text{HF}}[P_{\mathbf{k}}]$ such that $H_{\text{H}}^{\text{MF}}[P_{\mathbf{k}}] + H_{\text{F}}^{\text{MF}}[P_{\mathbf{k}}] = \sum_{\mathbf{k}} f_{\mathbf{k}}^{\dagger} h_{\text{HF}}[P_{\mathbf{k}}](\mathbf{k}) f_{\mathbf{k}}$.

To connect to our analytical discussion and with other work on TBG [59, 60, 72], we will use a subtraction point such that the starting Hamiltonian of our HF analysis is equal to (modulo a constant) the non-normal ordered but manifestly particle-hole symmetric Hamiltonian $H = H_0 + \tilde{H}_1$ with \tilde{H}_1 given in Eq. (A31). It is a matter of straightforward algebra to show that this corresponds to using $P_{\mathbf{k}}^0 = \mathbb{1}/2$ in Eq. (C9).

2. Numerical procedure

a. How we fix the phases in numerics

In this appendix, we describe how we fix the phases of the wavefunctions in the TBG sector in our numerics, using symmetries. As above, we denote the wavefunction in valley $\eta = \pm$, at momentum \mathbf{k} , and of band $n = \pm$ by $u_{n,\eta,\mathbf{k}}$; these are vectors in layer, sublattice, and \mathbf{G} space.

To fix the phase of the wavefunctions (modulo \pm) at every \mathbf{k} point, η , and n , we use that the Hamiltonian commutes with $C_{2z}\Theta$ and enforce that

$$C_{2z}\Theta u_{n,\eta,\mathbf{k}} = u_{n,\eta,\mathbf{k}}. \quad (\text{C10})$$

All that is left to do is remove the relevant parts of the remaining \mathbf{k} , η , and n dependent sign ambiguity. As a first step, we fix the relative sign of $u_{n,\eta,\mathbf{k}}$ and $u_{n,-\eta,-\mathbf{k}}$ by choosing

$$\Theta u_{n,\eta,\mathbf{k}} = u_{n,-\eta,-\mathbf{k}}. \quad (\text{C11})$$

Because of Eq. (C10), this is equivalent to $C_{2z}u_{n,\eta,\mathbf{k}} = u_{n,-\eta,-\mathbf{k}}$.

Next we fix the relative sign of the wavefunctions in the two bands $n = +$ and $n = -$ at each \mathbf{k} and η . If we are in the chiral limit and the system has the chiral symmetry C (which anti-commutes with the Hamiltonian at each \mathbf{k}), we can just enforce

$$Cu_{\pm,\eta,\mathbf{k}} = \pm i\eta u_{\mp,\eta,\mathbf{k}}. \quad (\text{C12})$$

Note it is easy to see that the prefactor has to have the form $\pm i\eta$ since $C^2 = \mathbb{1}$, $\{C_{2z}\Theta, C\} = 0$, and $[\Theta, C] = 0$. In our numerics, we can readily implement Eq. (C12) by

$$u_{+,\eta,\mathbf{k}} \rightarrow u_{+,\eta,\mathbf{k}}, \quad u_{-,\eta,\mathbf{k}} \rightarrow -i\eta \langle u_{-,\eta,\mathbf{k}} | C | u_{+,\eta,\mathbf{k}} \rangle u_{-,\eta,\mathbf{k}} \quad (\text{C13})$$

at every \mathbf{k} point.

When $w_0 \neq 0$ but not extremely large, we can expect $Cu_{\pm,\eta,\mathbf{k}} \simeq \text{const.} \times u_{\mp,\eta,\mathbf{k}}$. So we replace Eq. (C13) by the generalized condition

$$u_{+,\eta,\mathbf{k}} \rightarrow u_{+,\eta,\mathbf{k}}, \quad u_{-,\eta,\mathbf{k}} \rightarrow \eta \text{sign} [\text{Im} \langle u_{-,\eta,\mathbf{k}} | C | u_{+,\eta,\mathbf{k}} \rangle] u_{-,\eta,\mathbf{k}}. \quad (\text{C14})$$

Note that Eq. (C10) actually implies $\langle u_{-,\eta,\mathbf{k}} | C | u_{+,\eta,\mathbf{k}} \rangle \in i\mathbb{R}$ even away from the chiral limit.

Finally, the only remaining relevant relative sign is that of $u_{n,\eta,\mathbf{k}}$ and $u_{n,-\eta,\mathbf{k}}$. We fix this one by making sure that

$$C_{2z}Pu_{n,\eta,\mathbf{k}} = n\eta u_{n,-\eta,\mathbf{k}}. \quad (\text{C15})$$

In practice, we make sure that Eq. (C15) holds by adjusting the sign of the wavefunctions of *both* bands in the $\eta = -$ valley according to

$$u_{+,+,\mathbf{k}} \rightarrow u_{+,+,\mathbf{k}} \langle u_{+,+,\mathbf{k}} | C_{2z}P | u_{-,-,\mathbf{k}} \rangle \quad u_{+,-,\mathbf{k}} \rightarrow -\langle u_{+,-,\mathbf{k}} | C_{2z}P | u_{+,-,\mathbf{k}} \rangle u_{+,-,\mathbf{k}}, \quad n = \pm. \quad (\text{C16})$$

b. Iterative procedure

We will here give the details of how we numerically solve the HF equations in MSTG and find self-consistent solutions $P_{\mathbf{k}}$ in Eq. (C1). Our iterative procedure has the following steps:

1. Guess an initial form of the projector $P_{\mathbf{k}}$. We choose the initial ansatz for $P_{\mathbf{k}}$ for a given symmetry breaking state to be those given in the band basis in Table III.
2. $P_{\mathbf{k}}$ is then substituted into the HF functional expression (4.1). The HF functional is then diagonalized at each point via a unitary transformation $U_{\mathbf{k}}$.
3. $P_{\mathbf{k}}$ is then recomputed from the HF functional as $P_{\mathbf{k}} = U_{\mathbf{k}}^* D U_{\mathbf{k}}^T$ where D is the density matrix in the diagonal basis at zero temperature, with 1's on the diagonal elements corresponding to filled bands of the HF Hamiltonian and zeros elsewhere.
4. We then check if $P_{\mathbf{k}}$ has converged from the previous iteration. If it has, the procedure is finished and we have found a self-consistent solution characterized by $P_{\mathbf{k}}$. If $P_{\mathbf{k}}$ has not converged, we return to step 2.

Unless otherwise specified, we include 3 shells of moiré Brillouin zones in our construction of the trilayer model (37 unit moiré unit cells total) of which we include out to 2 shells of moiré Brillouin zones in our numerical calculations and a 10×10 grid of \mathbf{k} -points per moiré Brillouin zone (243 \mathbf{k} -points total per moiré Brillouin zone). Unless otherwise specified, we take $\epsilon = 7$ and screening length $d_s = 40$ nm.

c. Ground State dependence on grid size

Here, we make note of an effect we observe for smaller grid size for a 6×6 \mathbf{k} grid (75 \mathbf{k} -points per moiré unit cell). We find for the smaller grid, a region of spin or valley polarized ground state emerges for intermediate, nonzero D_0 and intermediate to large w_0 which became disfavored relative to the IVC_- state as we increased the grid size to better resolve the Dirac cones of the graphene-like bands. We attribute this difference to momentum dependent mixing between remote bands in the Fock term.

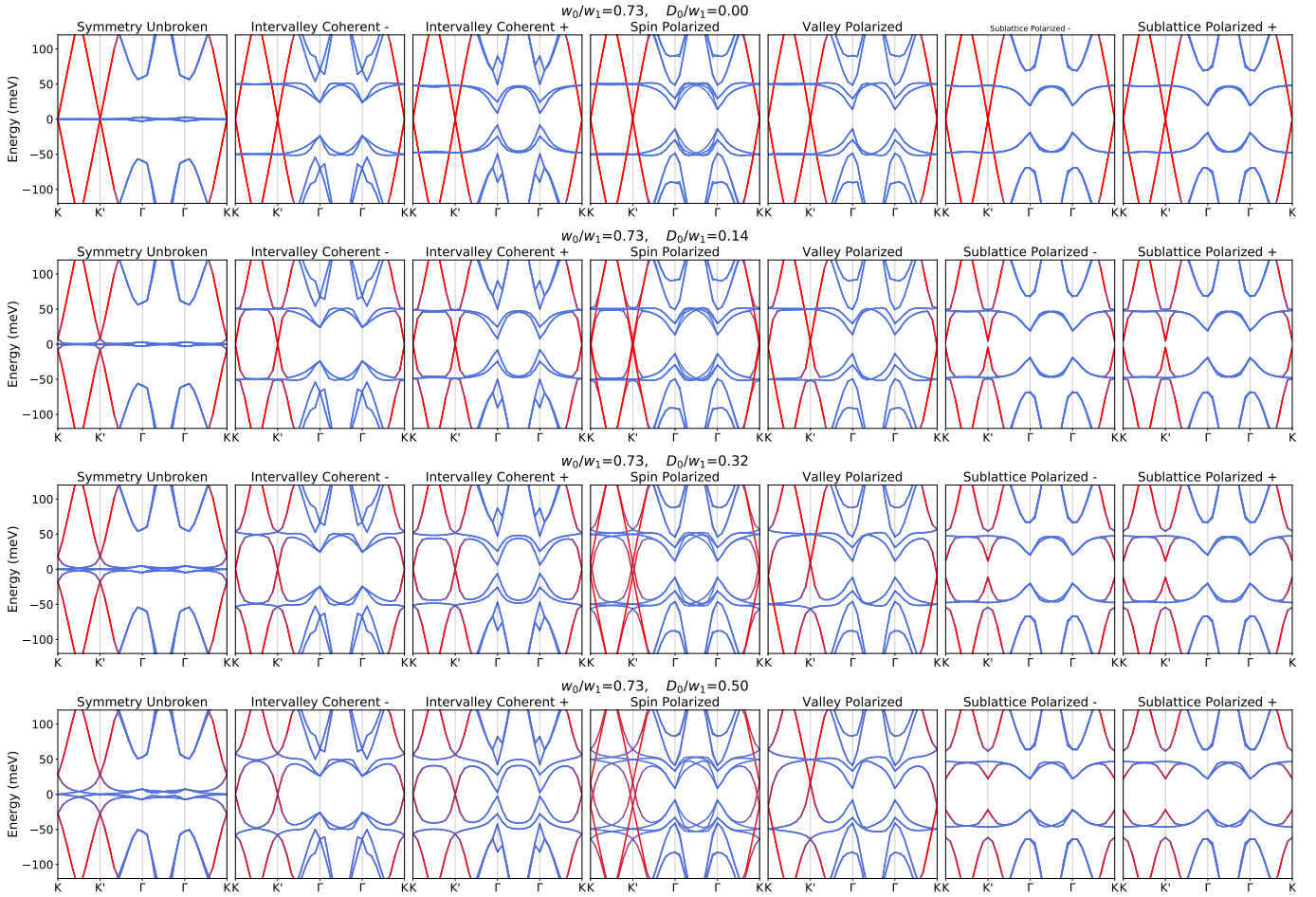


FIG. 8: We show band structures of converged solutions at $\nu = 0$ corresponding to the ansätze in Table III, including those which we do not appear in the phase diagram for dielectric constant $\epsilon = 7$ at zero and nonzero D_0 .

3. Varying Parameters

In Fig. 8, we show the band structures obtained from our self-consistent calculation for all ansätze we attempt, excluding the spin Hall state which is quantitatively and qualitatively similar to our SLP₋ state which is shown. The contour taken through the moiré Brillouin zone to produce the band structures is shown in Fig. 2.

We verify our results are relatively independent of parameters w_0/w_1 and D_0/w_1 may vary in experiment. We first consider additional values of the dielectric constant ϵ , which controls the relative contributions to the Hartree-Fock energy of the kinetic and interaction terms. We find the phase boundary between the IVC₋ state and sublattice polarized states shifts slightly but the overall features do not depend on the values of ϵ we test in Fig. 9. We also vary the screening length d_s from 10 nm to 80 nm and find while the relative energies shift, the phase diagram does not change when d_s is varied. We also verify our results are stable to increasing the number of bands we keep in our self-consistent calculation, by repeating the same calculation with the closest 8 bands per valley and spin to the Fermi level at charge neutrality. The resulting band structures as D_0 is varied are shown in Fig. 11. We note the IVC₊ which we found a converged solution for in the calculation keeping only 4 bands per spin and valley no longer converges when more bands are added and is not shown in Fig. 11. Given the large energy difference separating the IVC₊ from the other lower energy symmetry breaking states, the absence of the IVC₊ on doubling the number of bands is not relevant to our main results. We also increase our \mathbf{k} -grid size up to a 14×14 grid to verify convergence.

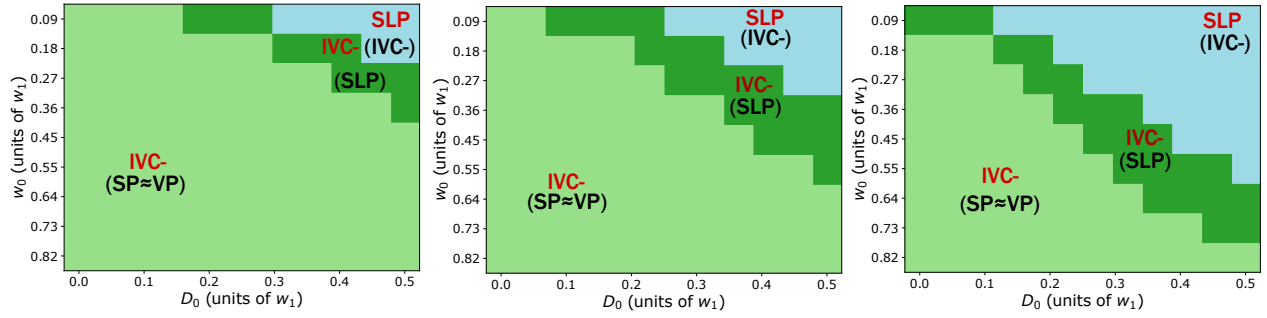


FIG. 9: Phases as we vary the dielectric constant in the Coulomb potential in Eq. (2.9) from $\epsilon = 4$ (left), $\epsilon = 7$ (center), and $\epsilon = 15$ (right). As the ϵ is varied, the phase boundaries shift but the presences of a time-reversal-odd IVC phase and sublattice polarized phase are robust against changes in the potential. As ϵ increases, the bandwidth of the TBG-like bands relative to the Coulomb potential $V(\mathbf{q})$ increases; as predicted in Table III, the D_0 -dependent increases in bandwidth favors a sublattice polarized state over the IVC...

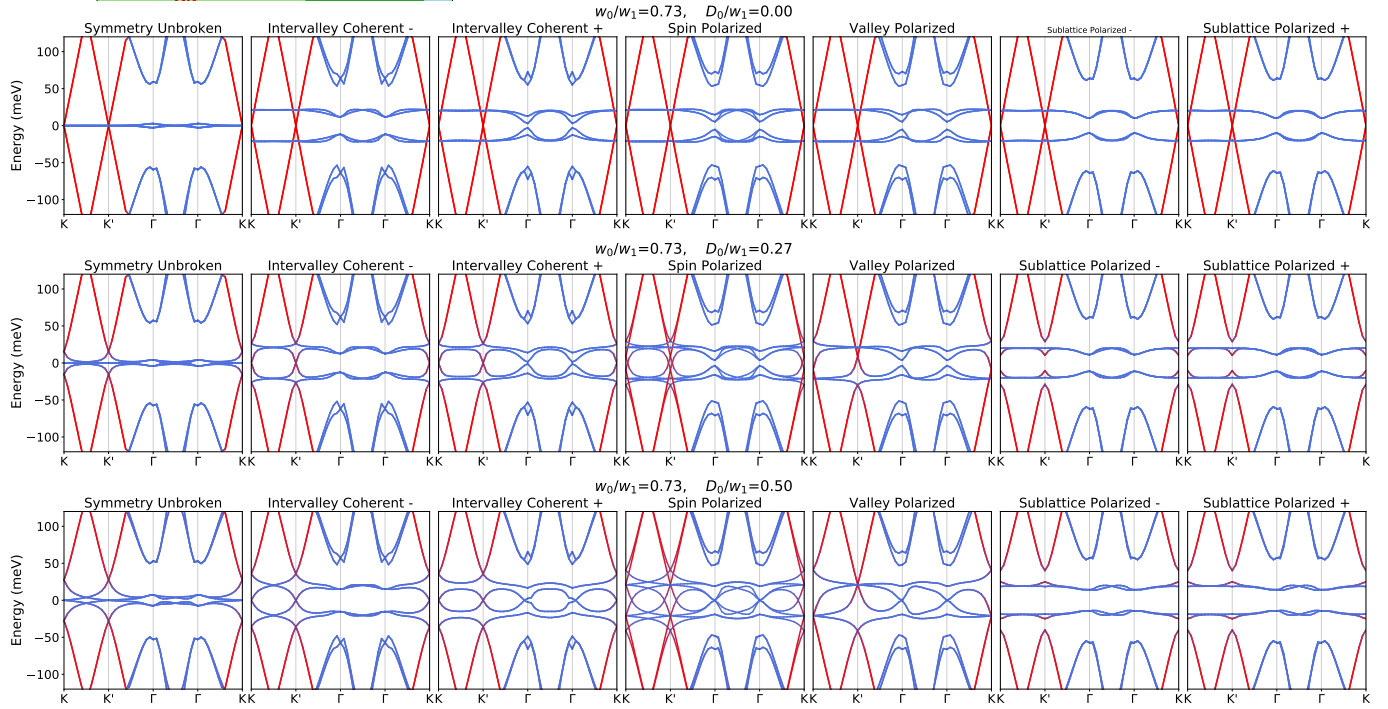


FIG. 10: We show band structures for a HF calculation with a 14×14 grid for extreme value $\epsilon = 15$, $d_s = 40$ nm.

Appendix D: Numerics at $\nu = 2$

1. Stable Phases at $\nu = 2$

In this appendix we will briefly discuss our numerical results at $\nu = 2$. We readjust our numerical procedure such that the chemical potential is recomputed at each iteration to fix the filling at $\nu = +2$ electrons per moiré unit cell. In the projected space of the TBG-like and graphene-like bands, this means the trace of $Q_{\mathbf{k}}$ satisfies $\frac{1}{N} \sum_{\mathbf{k}} \text{Tr}[Q_{\mathbf{k}}] = 10$. We also adjust the projectors P corresponding to our starting ansätze to be compatible with half-filling by setting $P_{b,\nu=2} = \frac{1}{2}(1 + s_z)P_{b,\nu=0}$ in the TBG-bands at $\nu = 2$. The structure of $P_{g,\nu=2} = \frac{1}{2}(1 - \sigma_z)$ is the same as at charge neutrality. The energies resulting from our HF calculations are shown in Fig. 12. We observe no qualitative change from the phases at charge neutrality, but note the spin-polarized IVC₋ and spin-valley polarized state are closer in energy at $\nu = 2$. We show the band structures obtained for the full set of ansätze we check in Fig. 13. We note that where the IVC₋ at charge neutrality was a semimetal, at $\nu = 2$ the spin polarized IVC₋ is a metal with a small

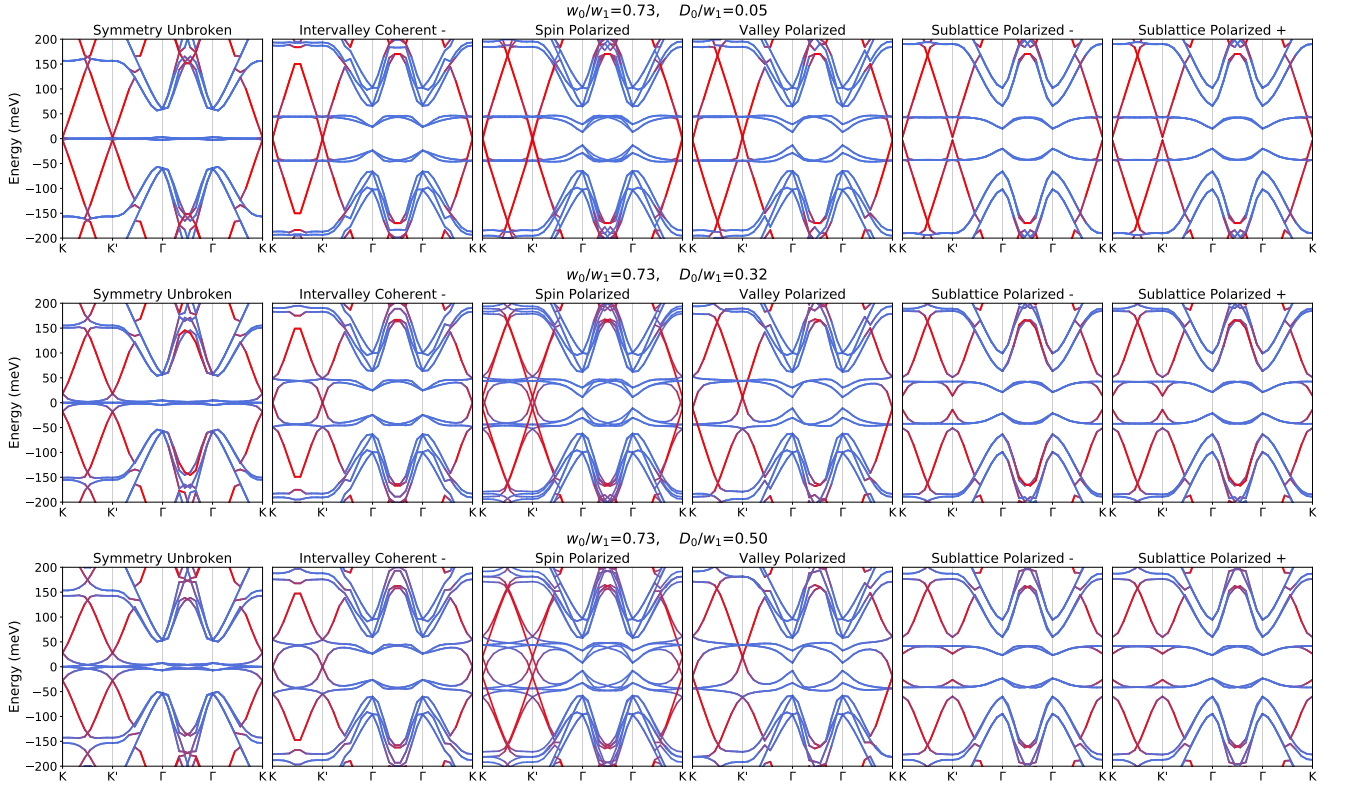


FIG. 11: We show band structures for a self-consistent calculation keeping 8 rather than 4 bands per spin-valley flavor, for $\epsilon = 7$, $d_s = 40$ nm.

Fermi surface. We also note that the spin polarized versions of the sublattice polarized states now are either a metal or semimetal. The change of the SLP states to semimetals at $|\nu| = 2$ can be understood as state with an insulating SLP order in the TBG-like bands of one spin flavor and an order which fills both TBG-like bands in the other spin flavor, resulting in a Dirac crossing where the TBG-like and graphene-like bands connect at the Fermi level.

2. Connection to Experimental Phase Diagram

We can connect the band structures at $\nu = 0$ and $\nu = 2$ to the experimental phase diagram; at $\nu = 0$, the leading instability is either a semimetallic IVC₋ state or an insulating sublattice polarized state. Either of these states can be related to the region of high resistivity at $\nu = 0$ which persists both at $D_0 = 0$ and $|D_0| > 0$. At $\nu = 2$, the spin polarized IVC₋ is still the leading phase at $D_0 = 0$ and for a finite range of D_0 . However, unlike at $\nu = 0$, the IVC₋ is metallic at $\nu = 2$. We note the lack of a strong high-resistivity state at $\nu = 2$ when $D_0 = 0$ in the experimental phase diagram. For large enough D_0 , the IVC₋ state transitions to a semimetallic spin-polarized SLP state. This transition may be related to the appearance of a state with higher resistivity at $\nu = 2$ for a finite value of D_0 in the experimental phase diagram.

Appendix E: Band Structure and solution for $D_0 \sim w_1$

In Sec. IV, we studied phases varying D_0 up to half the value of w_1 . We will here discuss the behavior of the phases we find varying D_0 up to and past w_1 . We find our numerics begin not to converge for $D_0/w_1 > 1$ for some values of w_0 we study. We will focus here on the behavior of the valley polarized state at $\nu = 0$ and the spin polarized version of this state at $\nu = 2$ but note the similarities between the behavior of this state and the sublattice polarized states and IVC states where they converge in the HF numerics.

At $\nu = 0$, as stated in the main text, the valley polarized solution becomes metallic as D_0 increases and the bands

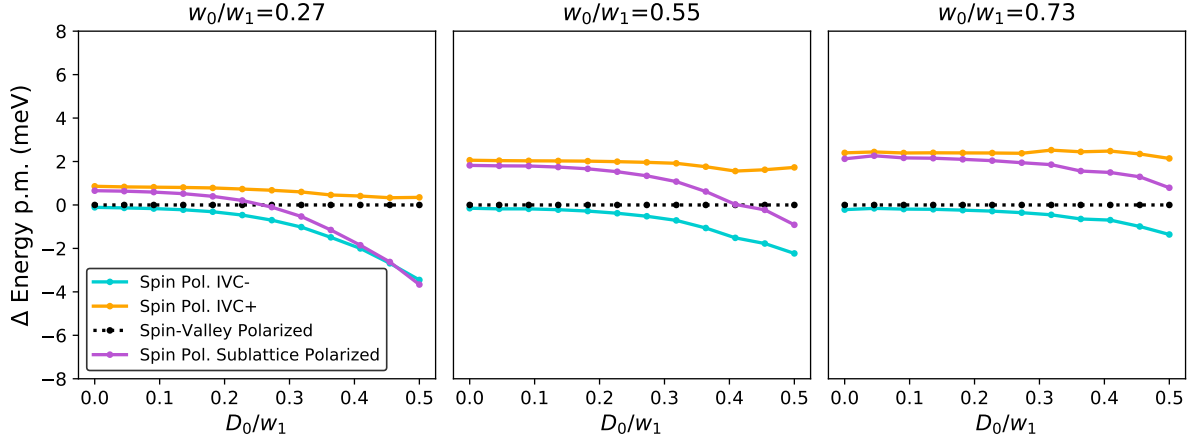


FIG. 12: We show the energies resulting from our HF calculations at $\nu = 2$. As at $\nu = 0$, the spin-polarized IVC₋ is dominant relative to a spin-valley polarized state, except for a region of smaller w_0 and large D_0 where a spin-polarized sublattice polarized state dominates. We note that the spin-valley polarized state is closer in energy to the spin-polarized IVC₋ state as we vary D_0 in our $\nu = 2$ calculation than at $\nu = 0$.

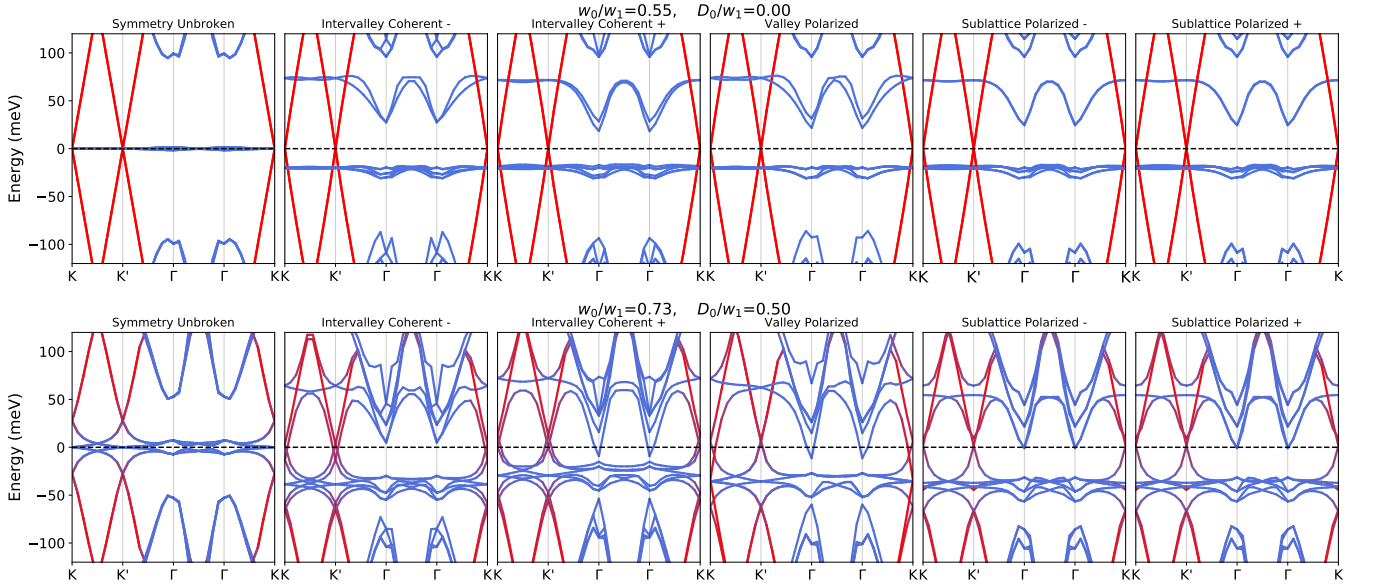


FIG. 13: Band structures obtained at $\nu = 2$ with $\epsilon = 7$, $d_s = 40$ nm for the full set of spin-polarized versions of ansätze in Table III.

in either minivalley are pushed away from the Fermi level. For D_0/w_1 relatively small, the bands in each valley flavor cross the Fermi level only near the K and K' points. However as $D_0/w_1 \rightarrow 1$, the bandwidth of the valley-polarized TBG-like bands increases, resulting in additional Fermi surfaces near the Γ point. For large D_0 , the lower TBG-like band which begins above the Fermi level at small D_0 is pushed completely below the Fermi level and, for large enough D_0 , the band structure reverts to a filled lower bands solution.

At $\nu = 2$, the band structure of the solutions we obtain via HF for the spin-valley polarized ansatz also begin with separated Fermi-surfaces around the K and K' points and no Fermi surfaces near Γ . Similar to the $\nu = 0$ case, as D_0 increases, the bandwidth of the polarized TBG-like bands increases, leading to changes in the Fermi surface, with additional hole-like Fermi surfaces appearing near the Γ point at intermediate values of D_0 until the lower polarized band which begins above the Fermi level is pushed below the Fermi level for large D_0 . Unlike at $\nu = 0$, the solution at large D_0 retains a finite valley polarization. Experimental Hall densities measured in MSTG show a change in sign of the Hall density at half filling for large D_0 at $|\nu| = 2$. In the main text we argued this effect could be accounted for by the merging of Fermi surfaces of the IVC₋ bands (or a similar transition of the SLP₋ phase) near the K and

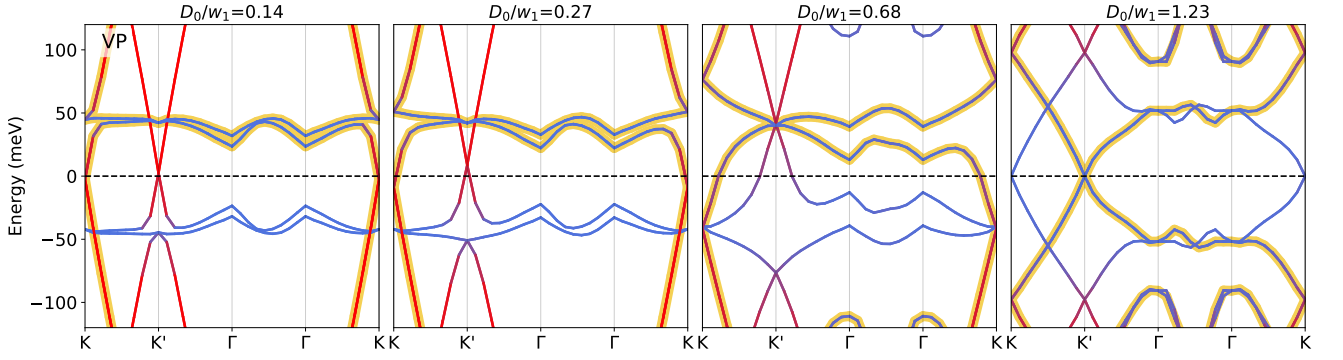


FIG. 14: We show the band structure of the VP state at $\nu = 0$ as we vary D_0 for $w_0/w_1 = 0.55$. At charge neutrality, the VP state at D_0 comparable to w_1 acquires additional hole-like Fermi surfaces. We hi-light a single valley-spin flavor band in yellow to show the evolution of a single flavor at large D_0 .

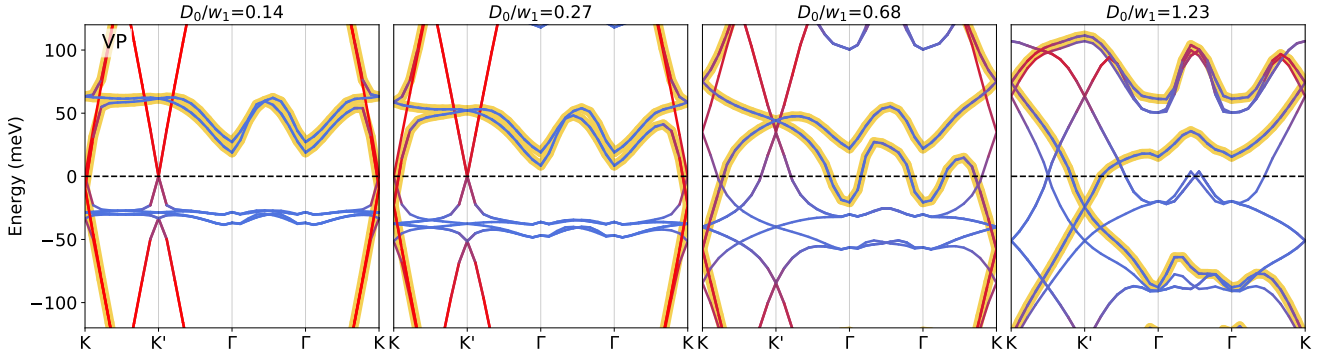


FIG. 15: We show the band structure of the spin-VP phase at $\nu = 2$ as we vary D_0 for $w_0/w_1 = 0.33$. We note at half-filling both the spin-valley polarized and intervalley coherent states at D_0 comparable to w_1 acquire additional hole-like Fermi surfaces. We hi-light a single valley-spin flavor band in yellow to show where the additional Fermi surfaces appear as a function of D_0 .

K' point doping slightly away from $\nu = 2$ when $D_0 > 0$. Should the experimental value of the displacement field at which the sign change appears exceed the values of the displacement field discussed in the main text, it is possible that instead the sign change could be related to the dramatic changes in Fermi surfaces we observe and discuss here. We show how the band structure of the valley polarized state evolves at large D_0 in Fig. 14 for $\nu = 0$ and in Fig. 15 for $|\nu| = 2$. The IVC₋ and sublattice polarized states display similar behavior at charge-neutrality and $\nu = 2$.

Another general feature we note is a tendency after a critical value of $D_0/w_1 > 1$ for all of our orders to converge to solutions which are nearly degenerate in energy. While some retain the symmetry breaking in their initial ansatz, the solutions generally have a preference to fill the lower bands up to the chemical potential at $\nu = 0$ and $\nu = 2$ and exhibit additional Fermi surfaces at large D_0 . Our numerics do not always converge in this region, so we do not include it in our phase diagrams or energies which are limited to $D_0/w_1 < 0.5$.

Appendix F: Different contributions to the energy

In this appendix, we detail how the different energetic contributions for particle-hole instabilities, discussed briefly in Sec. V of the main text, are derived and what their analytical form is. We use these expressions to compute their respective scaling with system parameters, summarized in Table III.

Our starting point are Slater-determinant states with correlation matrix $P_{\mathbf{k}}^0 = (\mathbb{1} + Q_{\mathbf{k}}^0)/2$, see Eq. (3.11) for its definition, where $Q_{\mathbf{k}}^0$ is block diagonal in the graphene-like-TBG-like space and of the form

$$Q_{\mathbf{k}}^0 = \begin{pmatrix} Q_{\mathbf{k}}^{g,0} & 0 \\ 0 & Q_{\mathbf{k}}^{b,0} \end{pmatrix}, \quad Q_{\mathbf{k}}^{g,0} = -\sigma_z, \quad (\text{F1})$$

i.e., further assume that the graphene-like bands do not develop any order while we take several different candidate orders, $Q_{\mathbf{k}}^{b,0}$, in the TBG-like band to be discussed shortly. Equation (F1) is motivated by the following observations: first, at $D_0 = 0$, all order parameters can be classified as mirror-even (diagonal in graphene-TBG space) and mirror-odd (off-diagonal in graphene-TBG space). Based on the band structure, it is natural to expect that the former are energetically favored. In fact, we will see in Appendix F 4 explicitly and in our HF numerics [see second row in Fig. 17] as well as in our analysis of exact eigenstates in Appendix B that off-diagonal components are not favored at $D_0 = 0$. Second, $Q_{\mathbf{k}}^{g,0} = -\sigma_z$ is motivated by the observations of Appendix B that the manifold of exact ground states of the MSTG Hamiltonian has this form for $D_0 = 0$ in a finite range of coupling parameters [λ in Eq. (3.4)] between the graphene cones and the TBG bands. Again, this will be confirmed by our HF numerics, where the lowest-energy solutions are of the form of Eq. (F1) in the flat-decoupled limit ($D_0 = W_{\text{TBG}} = 0$), and our discussion of order in the graphene-like bands in Appendix F 3 below.

In Sec. III C of the main text, we have constructed the different possible ansätze, $Q_{\mathbf{k}}^{b,0}$, using the $(U(4) \times U(4))_{b,cf}$ symmetry. We will complement this here by deriving the exact same set of states purely within HF. As a first step, let us focus on minimizing the Fock energy in Eq. (C7d) at $D_0 = w_0 = 0$:

$$E_{\text{F}}[P_{\mathbf{k}}^0] = E_{\text{F,g}} - \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{\mathbf{k} \in \text{MBZ}} \text{tr} \left[P_{\mathbf{k}}^b (F_{\mathbf{k},\mathbf{q}}^{bb})^T P_{\text{MBZ}(\mathbf{k}+\mathbf{q})}^b (F_{\mathbf{k},\mathbf{q}}^{bb})^* \right], \quad P_{\mathbf{k}}^b = \frac{1}{2}(\mathbb{1} + Q_{\mathbf{k}}^{b,0}), \quad (\text{F2})$$

where $E_{\text{F,g}}$ is the Fock energy of the graphene subsystem being in a semimetallic state and F^{bb} the form factors in TBG subspace, which obey [cf. Eq. (A33a)]

$$\bar{F}^{bb}(\mathbf{k}, \mathbf{q}) = \bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}) \eta_0 \sigma_0 s_0 + i \bar{F}_2^{bb}(\mathbf{k}, \mathbf{q}) \eta_0 \sigma_2 s_0 + \mathcal{O}(w_0, D_0^2). \quad (\text{F3})$$

Since $\langle A, B \rangle = \text{tr}[AB]$ defines an inner product on Hermitian matrices, the Cauchy-Schwarz inequality holds, $\langle A, B \rangle \leq \sqrt{\langle A, A \rangle} \sqrt{\langle B, B \rangle}$ with maximum reached when $A = cB$ with $c > 0$. Applying this to the second term in Eq. (F2), we see that this term reaches its minimum when [47]

$$P_{\mathbf{k}}^b = C_{\mathbf{k},\mathbf{q}} (F_{\mathbf{k},\mathbf{q}}^{bb})^T P_{\text{MBZ}(\mathbf{k}+\mathbf{q})}^b (F_{\mathbf{k},\mathbf{q}}^{bb})^*, \quad C_{\mathbf{k},\mathbf{q}} > 0, \quad \forall \mathbf{k}, \mathbf{q}. \quad (\text{F4})$$

In general, solving Eq. (F4) can be difficult (or impossible) and the minimum of $E_{\text{F}}[P_{\mathbf{k}}^0]$ requires a momentum-dependent correlator. However, for $D_0 = w_0 = 0$, it holds $(F_{\mathbf{k},\mathbf{q}}^{bb})^T (F_{\mathbf{k},\mathbf{q}}^{bb})^* = \gamma_{\mathbf{k},\mathbf{q}} \mathbb{1}$ with $\gamma_{\mathbf{k},\mathbf{q}} = \sum_{j=1,2} (\bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}))^2 > 0$. Then, Eq. (F4) is obeyed as long as

$$\left[Q_{\mathbf{k}}^{b,0}, \eta_0 \sigma_2 s_0 \right] = 0. \quad (\text{F5})$$

This is the same commutator relation as in Eq. (3.15). All momentum independent states that obey this property, as well as their symmetries are summarized in Table II. Furthermore, we have also checked all of the states in Table II are also exactly degenerate in the Hartree term (C7c) and in the additional subtraction point contribution in Eq. (C9), as long as $w_0 = D_0 = 0$.

In the following subsections, we will discuss the various energetic corrections to $Q_{\mathbf{k}}^0$ in Eq. (F1) once we turn on D_0 , w_0 , and W_{TBG} , and allow the graphene-like bands to develop order and mix with the TBG-like bands in the vicinity of the Dirac cones. To simplify the discussion, we will first focus on the $SU(2)_+ \times SU(2)_-$ -symmetric model defined in Sec. II. In that case, certain pairs of states, referred to Hund's partners in [27], are guaranteed to be degenerate. Therefore, we will focus, without loss of generality, on only one member of each pair of Hund's partners; for concreteness, we choose those above the vertical line in Table II. In Sec. F 5, we will study how these degeneracies are lifted for $J_H \neq 0$ in Eq. (3.16).

1. Deformation of the form factors

Let us first analyze the energetic impact for these different phases due to the deformation of the form factors, see Eq. (A33), as a consequence of D_0 and w_0 . To this end, we compute the change of the HF energy E^{HF} in Eq. (C7) for $Q_{\mathbf{k}}^0$ in Eq. (F1) for the different $Q_{\mathbf{k}}^b$ in Table II, setting $W_{\text{TBG}} = 0$. Using the structure of the form factors in Eq. (A33) along with the important constraints in Eqs. (A35) and (A37), we obtain after straightforward algebra the resultant change of energy $\Delta E_{bb}(W_{\text{TBG}} = 0)$ listed in Table III. The prefactors are given by

$$c_1 = \frac{2}{N} \sum_{\mathbf{q}, \mathbf{k}} V(\mathbf{q}) \left\{ [\bar{F}_3^{bb}(\mathbf{k}, \mathbf{q})]^2 + [\bar{F}_4^{bb}(\mathbf{k}, \mathbf{q})]^2 \right\}, \quad (\text{F6})$$

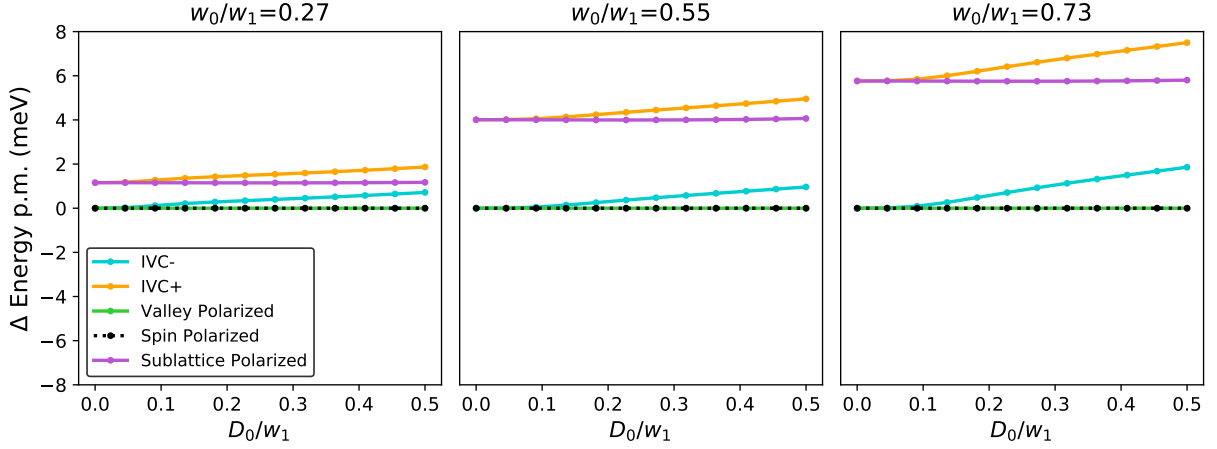


FIG. 16: HF energies relative to the SP state obtained by plugging in the ansätze in Table II in the HF energy expression. The behavior agrees with the analytical prediction in the column $\Delta E_{bb}(W_{\text{TBG}} = 0)$ in Table III.

$$c_2 = \frac{8}{N} \sum_{\mathbf{G}} V(\mathbf{G}) \left[\sum_{\mathbf{k}} \bar{F}_4^{bb}(\mathbf{k}, \mathbf{G}) \right]^2, \quad (\text{F7})$$

$$c_3 = \frac{2}{N} \sum_{\mathbf{q}, \mathbf{k}} V(\mathbf{q}) \{ [\bar{F}_5^{bb}(\mathbf{k}, \mathbf{q})]^2 + [\bar{F}_6^{bb}(\mathbf{k}, \mathbf{q})]^2 \}, \quad (\text{F8})$$

$$c_4 = \frac{2}{N} \sum_{\mathbf{q}, \mathbf{k}} V(\mathbf{q}) \{ [\bar{F}_7^{bb}(\mathbf{k}, \mathbf{q})]^2 + [\bar{F}_8^{bb}(\mathbf{k}, \mathbf{q})]^2 \}. \quad (\text{F9})$$

Due to Eq. (A38), we expect the Hartree contribution, c_2 , to be numerically smaller than c_1 . We have verified the form of these energy contributions within our HF numerics, see Fig. 16.

2. Finite bandwidth of the TBG-like bands

Next, we allow for $W_{\text{TBG}} \neq 0$ in Eq. (A12) by treating it as a perturbation to the different states in Table II.

a. First order perturbation theory

The first-order contribution is readily found by evaluating the expectation value of $\Delta H_0 := H_0 - H_0|_{W_{\text{TBG}}=0}$ with H_0 given in Eq. (A12) with respect to the states $|\psi[P_{\mathbf{k}}]\rangle$. Within the matrix notation introduced in Appendix C, the correction to the energy can also be written as

$$\Delta E^{(1)}[P_{\mathbf{k}}] = \sum_{\mathbf{k} \in \text{MBZ}} \text{tr} [P_{\mathbf{k}}^T (h^{\mathbf{k}}(\mathbf{k}) - h^{\mathbf{k}}(\mathbf{k})|_{W_{\text{TBG}}=0})]. \quad (\text{F10})$$

Using the form of the kinetic energy of the TBG-like bands given in Eq. (A20) and that $\bar{E}_1^b(\mathbf{k}) = -\bar{E}_1^b(-\mathbf{k})$, one finds

$$\Delta E^{(1)} = \frac{1}{2} \sum_{\mathbf{k} \in \text{MBZ}} \bar{E}_3^b(\mathbf{k}) w_0 D_0 \quad (\text{F11})$$

for all states of Table II. Consequently, we have to go to second order in W_{TBG} to find a contribution that favors one of the orders over the others. This is what we will discuss next.

b. Second order perturbation theory: superexchange

For the discussion of second order perturbation theory, it will most convenient to switch back to the second quantization description. The unperturbed Hamiltonian is the full Hamiltonian in the flat-band limit, $H_1 = H_0|_{W_{\text{TBG}}=0} + \tilde{H}_1$,

with \tilde{H}_1 given in Eq. (A31), and the perturbation is the dispersion in the TBG-like bands, i.e.,

$$H_{\text{II}} = W_{\text{TBG}} \sum_{\mathbf{k} \in \text{MBZ}} \sum_{\eta=\pm} \sum_{s=\uparrow, \downarrow} \sum_{p=\pm} \epsilon_{(b,p),\eta}(\mathbf{k}) b_{\mathbf{k};p,\eta,s}^\dagger b_{\mathbf{k};p,\eta,s} = W_{\text{TBG}} \sum_{j=0}^3 H_{\text{II}}^{(j)}, \quad H_{\text{II}}^{(j)} = \sum_{\mathbf{k} \in \text{MBZ}} E_j^b(\mathbf{k}) b_{\mathbf{k}}^\dagger M_j b_{\mathbf{k}}. \quad (\text{F12})$$

Here we used the parameterization in Eq. (A20) and introduced the matrices

$$M_0 = \sigma_3 \eta_0 s_0, \quad M_1 = \sigma_0 \eta_3 s_0, \quad M_2 = \sigma_3 \eta_3 s_0, \quad M_3 = \sigma_0 \eta_0 s_0 \quad (\text{F13})$$

acting in the TBG-like subspace. As will become useful shortly, this decomposition of the kinetic energy has the property that all M_j defined in Eq. (F13) either commute or anticommute with all of the Q^b in Table II separately.

Let us assume that (one of) the groundstate(s) of H_I with energy $E_{Q^b}^0$ is the product state $|\psi[Q^b]\rangle$, characterized by the correlation matrix $P_{\mathbf{k}} = (\mathbb{1} + \text{diag}(-\sigma_z, Q^b))/2$ in Eq. (C1). Here, Q^b can be any of the states listed in Table II. The energy correction to second order in W_{TBG} has the form

$$\Delta E^{(2)} = - \sum_{|n\rangle \neq |\psi[Q^b]\rangle} \frac{|\langle n | H_{\text{II}} |\psi[Q^b]\rangle|^2}{E_n^0 - E_{Q^b}^0}, \quad (\text{F14})$$

where the sum involves all eigenstates $|n\rangle$ of H_I (excluding $|\psi[Q^b]\rangle$) with energy E_n^0 . Assuming that the system is in an insulating phase in the TBG-like sector, we will have a finite gap, $E_n^0 > E_{Q^b}^0$, such that perturbation theory is controlled in bandwidth over gap and $\Delta E^{(2)} \leq 0$, i.e., the ‘‘superexchange’’-like processes discussed here, if finite, will lower the energy.

We will next show that for a given Q^b , all terms j in Eq. (F12) with $[M_j, Q^b] = 0$ will not contribute to Eq. (F14) and, hence, only those that obey $\{M_j, Q^b\} = 0$ can lower the energy to order W_{TBG}^2 . To see this, first note that $[M_{j_0}, Q^b] = 0$ implies that there exists a unitary matrix V such that both $V^\dagger Q^b V$ and $V^\dagger M_{j_0} V$ are diagonal. Since $\text{tr} Q^b = 0$, we can take $V^\dagger Q^b V = \text{diag}(1, 1, 1, 1, -1, -1, -1, -1)$ without loss of generality at $\nu = 0$. Consequently, the associated many-body state assumes the form

$$|\psi[Q^b]\rangle = \prod_{\mathbf{k} \in \text{MBZ}} \prod_{v=1}^4 \bar{b}_{\mathbf{k};v}^\dagger |\psi_{0,g}\rangle, \quad \bar{b}_{\mathbf{k};v} := \sum_{\alpha} b_{\mathbf{k};\alpha} V_{\alpha,v}, \quad (\text{F15})$$

where $|\psi_{0,g}\rangle$ is the state of empty TBG-like bands and lower-filled graphene-like bands. Upon noting that $M_j^T = M_j$ and $M_j^2 = M_j$, we also see

$$H_{\text{II}}^{(j_0)} = \sum_{\mathbf{k} \in \text{MBZ}} E_{j_0}^b(\mathbf{k}) \sum_{v=1}^8 r_v \bar{b}_{\mathbf{k};v}^\dagger \bar{b}_{\mathbf{k};v}, \quad r_v = \pm 1. \quad (\text{F16})$$

Consequently, we get

$$\langle n | H_{\text{II}}^{(j_0)} |\psi[Q^b]\rangle = \sum_{\mathbf{k} \in \text{MBZ}} E_{j_0}^b(\mathbf{k}) \sum_{v=1}^4 r_v \langle n | \psi[Q^b]\rangle = 0 \quad (\text{F17})$$

since $|\psi[Q^b]\rangle$ and $|n\rangle$ are orthogonal for all terms in the sum of Eq. (F14).

From this algebraic criterion, we can immediately read off that the SP and VP states have $\Delta E^{(2)} = 0$, i.e., cannot gain energy from this ‘‘superexchange’’ mechanism. The SLP_\pm and SSLP_\pm order parameters, however, anticommute with M_0 and M_2 . As such, they can benefit both from the superexchange processes associated with $E_0^b = \bar{E}_0^b$ (already present when $w_0 = D_0 = 0$, but θ tuned away from the magic angle) and $E_2^b = \bar{E}_2^b D_0$ (unique to the trilayer system and only non-zero for $D_0 \neq 0$). Furthermore, while both IVCs can benefit from $E_1^b = \bar{E}_1^b w_0$, the IVC_- (IVC_+) also benefits from E_0^b ($E_2^b = \bar{E}_2^b D_0$). This leads to the results shown in Table III.

3. Ordering in the graphene-like bands

Let us next also take into account that the coupling between the graphene-like and TBG-like bands can modify the ansatz in Eq. (F1). Note that a perturbative treatment of bandwidth of the graphene-like bands, similar to our

analysis in Sec. F 2, is not controlled as these bands are highly dispersive and their bandwidth and energy, $\epsilon_{(g,p),\eta}(\mathbf{k})$, at generic momentum points are larger than the interaction energy scale. At the same time, $\epsilon_{(g,p),\eta}(\mathbf{k})$ is small around the K/K' points such that taking the interaction as a small parameter—as employed in the full quantum mechanical analysis in Appendix B—is not necessarily valid either for realistic system parameters. Therefore, we here use a different approach that takes advantage of these observations: while the diagonal structure of Eq. (F1) and $Q_{\mathbf{k}}^g = -\sigma_z$ are likely to be a good approximation far from the K/K' points, it is expected to be modified in the vicinity of these points.

To formalize this, we define $\mathcal{A} = \mathcal{A}_+ \cup \mathcal{A}_-$ where \mathcal{A}_+ (\mathcal{A}_-) is a region of momentum points around the K (K') point. The precise form does not matter, we only need that its area, $V_{\mathcal{A}}$, is small compared to the overall size of the MBZ. We now take

$$Q_{\mathbf{k}} = \begin{cases} Q_{\mathbf{k}}^0, & \mathbf{k} \in \bar{\mathcal{A}}, \\ Q_{\mathbf{k}}^1, & \mathbf{k} \in \mathcal{A}, \end{cases} \quad \bar{\mathcal{A}} := \text{MBZ} \setminus \mathcal{A} \quad (\text{F18})$$

with $Q_{\mathbf{k}}^0$ given in Eq. (F1), and accordingly for $P_{\mathbf{k}}$. Expanding the change of the HF energy, $\Delta E^{\text{HF}} = E^{\text{HF}}[P_{\mathbf{k}}] - E^{\text{HF}}[P_{\mathbf{k}}^0]$, to leading order in $V_{\mathcal{A}}$, one finds after straightforward algebra

$$\begin{aligned} \Delta E^{\text{HF}} &= \frac{1}{2} \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}}^T h^k(\mathbf{k})] + \frac{1}{N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) \text{Re} \left\{ \left(\sum_{\mathbf{k}' \in \bar{\mathcal{A}}} \text{tr}[P_{\mathbf{k}'}^0 F_{\mathbf{k}',\mathbf{G}}^T] \right) \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}} F_{\mathbf{k},\mathbf{G}}^T] \right\} \\ &\quad - \frac{1}{N} \sum_{\mathbf{G} \in \text{RL}} \sum_{\mathbf{k} \in \mathcal{A}} \sum_{\mathbf{k}' \in \bar{\mathcal{A}}} V(\mathbf{k}' - \mathbf{k} + \mathbf{G}) \text{tr}[\delta Q_{\mathbf{k}} F_{\mathbf{k},\mathbf{k}'-\mathbf{k}+\mathbf{G}}^T P_{\mathbf{k}'}^0 F_{\mathbf{k},\mathbf{k}'-\mathbf{k}+\mathbf{G}}^*] \\ &\quad - \frac{1}{2N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) \Gamma_{\mathbf{G}}^0 \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}} F_{\mathbf{k},\mathbf{G}}^T] + \frac{1}{2N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}}^T F_{\mathbf{k},\mathbf{q}}^\dagger F_{\mathbf{k},\mathbf{q}}] + \mathcal{O}(V_{\mathcal{A}}^2), \end{aligned} \quad (\text{F19})$$

where we defined the shorthand $\delta Q_{\mathbf{k}} = Q_{\mathbf{k}}^1 - Q_{\mathbf{k}}^0$ and $\Gamma_{\mathbf{G}}^0 = 8 \sum_{\mathbf{k}} (F_1^{bb}(\mathbf{k}, \mathbf{G}) + F_1^{gg}(\mathbf{k}, \mathbf{G}))$. Here, the first, second, and third terms come from the dispersion, Hartree, and Fock term, while the last line is due to the subtraction point. To simplify further, note that we can extend the sums over \mathbf{k}' in Eq. (F19) to the entire MBZ without changing the result at linear order in $V_{\mathcal{A}}$. Focusing on the relevant case of $P_{\mathbf{k}}^0 = P^0$, we can write

$$\begin{aligned} \Delta E^{\text{HF}} &= \frac{1}{2} \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}}^T h^k(\mathbf{k})] \\ &\quad + \frac{1}{2N} \sum_{\mathbf{G} \in \text{RL}} V(\mathbf{G}) \text{Re} \left(\sum_{\mathbf{k}'} \text{tr}[Q_{\mathbf{k}'}^0 F_{\mathbf{k}',\mathbf{G}}^T] \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}} F_{\mathbf{k},\mathbf{G}}^T] \right) \\ &\quad - \frac{1}{N} \sum_{\mathbf{q}} \sum_{\mathbf{k} \in \mathcal{A}} V(\mathbf{q}) \left(\text{tr}[\delta Q_{\mathbf{k}} F_{\mathbf{k},\mathbf{q}}^T P_{\mathbf{k},\mathbf{q}}^0 F_{\mathbf{k},\mathbf{q}}^*] - \frac{1}{2} \text{tr}[\delta Q_{\mathbf{k}} F_{\mathbf{k},\mathbf{q}}^T F_{\mathbf{k},\mathbf{q}}^*] \right) + \mathcal{O}(V_{\mathcal{A}}^2). \end{aligned} \quad (\text{F20})$$

Postponing the discussion of $Q_{\mathbf{k}}$ with non-zero off-diagonal components that mix the graphene-like and TBG-like bands to the next subsection, we here focus on block diagonal

$$Q_{\mathbf{k}}^1 = \begin{pmatrix} Q_{\mathbf{k}}^g & 0 \\ 0 & Q_{\mathbf{k}}^b \end{pmatrix}, \quad (\text{F21})$$

where $Q_{\mathbf{k}}^g$ will be chosen to describe the same order as in the TBG-like bands, i.e., to break/keep exactly the same symmetries as the considered order parameter in the TBG-like bands.

Let us begin our discussion with the *spin-polarized* state, i.e., take

$$(Q_{\mathbf{k}}^g)_{\eta,\eta'} = \delta_{\eta,\eta'} \begin{cases} \sigma_0 s_3, & \mathbf{k} \in \mathcal{A}_\eta \\ -\sigma_3 s_0, & \text{otherwise} \end{cases} \quad (\text{F22})$$

in Eq. (F18). Using Eq. (F20), it is a matter of straightforward algebra to show that the associated change of the

energy is given by

$$\begin{aligned}
\Delta E_{\text{SpinPol}}^{\text{HF}} &\sim 2 \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} \bar{E}_0^g(\mathbf{k}; \eta) \\
&- \frac{2}{N} \sum_{\mathbf{q}} \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} V(\mathbf{q}) \left[(\bar{F}_{2,\eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{1,\eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] \\
&- \frac{2}{N} \sum_{\mathbf{q}} \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} V(\mathbf{q}) \left[(\bar{F}_{1,\eta}^{bg}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,\eta}^{bg}(\mathbf{k}, \mathbf{q}))^2 \right] D_0^2 \\
&- \frac{2}{N} \sum_{\mathbf{q}} \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} V(\mathbf{q}) \left[(\bar{F}_{7,\eta}^{bg}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{8,\eta}^{bg}(\mathbf{k}, \mathbf{q}))^2 \right] w_0^2 D_0^2 \\
&- \frac{2}{N} \left\{ \sum_{\mathbf{q}} \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} V(\mathbf{q}) \left[(\bar{F}_{7,\eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{8,\eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] + 2 \sum_{\mathbf{G}} V(\mathbf{G}) \sum_{\eta', \mathbf{k}'} \bar{F}_{8,\eta'}^{gg}(\mathbf{k}', \mathbf{G}) \sum_{\eta, \mathbf{k} \in \mathcal{A}_{\eta}} \bar{F}_{8,\eta}^{gg}(\mathbf{k}, \mathbf{G}) \right\} w_0^2 D_0^4
\end{aligned} \tag{F23}$$

where $\bar{E}_0^g(\mathbf{k}; \eta)$ is (the part of) the dispersion of the graphene-like bands defined in Eq. (A21). To keep the notation more compact, we further defined the η -dependent form-factor components as

$$F_{j,\eta}^{tt'}(\mathbf{k}, \mathbf{q}) = F_j^{tt'}(\mathbf{k}, \mathbf{q}) + \eta F_{j+2}^{tt'}(\mathbf{k}, \mathbf{q}), \quad j = 1, 2, \tag{F24}$$

$$F_{j,\eta}^{tt'}(\mathbf{k}, \mathbf{q}) = F_j^{tt'}(\mathbf{k}, \mathbf{q}) + \eta F_{j-2}^{tt'}(\mathbf{k}, \mathbf{q}), \quad j = 7, 8. \tag{F25}$$

To understand the scaling behavior of the first contribution associated with the kinetic term, let us take $\bar{E}_0^g(\mathbf{k}; \eta) = \alpha |D_0| + v_D |\delta_{\eta} \mathbf{k}|$ where $\delta_{\eta} \mathbf{k}$ is the distance from the K (K') point in valley $\eta = +$ ($\eta = -$) and $\alpha > 0$. As such we expect the scaling behavior

$$\sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} \bar{E}_0^g(\mathbf{k}; \eta) \sim c_1 |D_0| (\Delta k)^2 + c_2 v_D (\Delta k)^3, \quad c_j > 0, \tag{F26}$$

as a function of D_0 and the linear size, Δk , of \mathcal{A} . While we have already indicated the scaling behavior of all remaining terms in Eq. (F23) with D_0 and w_0 , we further note that they scale with the area of \mathcal{A} and, thus, quadratically with Δk .

To discuss the consequences, let us first focus on $D_0 = 0$. We see that the energetic penalty due to the kinetic term scales as Δk^3 while the energetic gain from the second line of Eq. (F23) scales as Δk^2 . One could naively conclude that this implies that the system should lower its energy by polarizing the graphene bands in some vicinity of its Dirac cones. Closer inspection, however, shows the involved energetics is equivalent to that of single-layer graphene: note that, for $D_0 = 0$, the first two lines of Eq. (F23) involve the dispersion and interaction matrix elements only in the graphene subspace. Since we know that graphene is not a spin-polarized insulator, this spin polarization cannot be preferred by the system. In fact, one can show that the two terms in the second line of Eq. (F23) cancel each other to leading order in Δk : To see this, let us focus on the contribution of one valley, say $\eta = +$, and consider $\mathbf{k} \rightarrow -\mathbf{q}_1/2$ (more precisely $\mathbf{k} = k(\cos \phi_0, \sin \phi_0) - \mathbf{q}_1/2$ and $k \rightarrow 0^+$) where we get

$$\sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{2,+}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{1,+}^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] \rightarrow \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sin^2 \left(\frac{\xi_{\text{MBZ}(\mathbf{q})} - \phi_0}{2} \right) - \cos^2 \left(\frac{\xi_{\text{MBZ}(\mathbf{q})} - \phi_0}{2} \right) \right] \delta_{\mathbf{G}_{\mathbf{q}}, 0}, \tag{F27}$$

$$= \sum_{\mathbf{q} \in \text{MBZ}} V(\mathbf{q}) \left[\sin^2 \left(\frac{\xi_{\mathbf{q}} - \phi_0}{2} \right) - \cos^2 \left(\frac{\xi_{\mathbf{q}} - \phi_0}{2} \right) \right]. \tag{F28}$$

Here we inserted the explicit form of the form factors given in Eq. (A39). This expression has to vanish since $V(\mathbf{q}) = V(-\mathbf{q})$, $\xi_{\mathbf{q}} = \xi_{-\mathbf{q}} + \pi$, and $\mathbf{q} \in \text{MBZ}$ implies $-\mathbf{q} \in \text{MBZ}$. The same analysis can be performed for the other valley $\eta = -$. For this reason, also the interaction correction in Eq. (F23) scales (at least) as $(\Delta k)^3$ for $D_0 = 0$ and, hence, does not generically dominate the kinetic contribution of the first line.

Turning on D_0 , we see that while there can be some additional gain starting at order D_0^2 , this gain is always overcompensated for sufficiently small D_0 by the additional cost from the kinetic energy which scales as D_0 .

These results are completely consistent with our HF numerics: as can be seen in Fig. 3, at $D_0 = 0$, the graphene Dirac cones are not split. Furthermore, while there are spin-split Dirac cones around the Fermi level at K/K' for

$D_0 \neq 0$, these are not related to the spin-ordering in the graphene-like bands defined in Eq. (F22). This becomes obvious by noting that both spin flavors of the upper (lower) graphene-like band are unoccupied (occupied), as is clearly visible based on the band connectivity. Furthermore, we can see it in the correlator of the SP around the K/K' points shown in Fig. 17.

All other states in Table II can be analyzed in a similar way. To begin with the *valley polarized* state, the analogue of Eq. (F22) reads as

$$Q_{\mathbf{k}}^g = \begin{cases} \sigma_0 \eta_3 s_0, & \mathbf{k} \in \mathcal{A}, \\ -\sigma_3 \eta_0 s_0, & \text{otherwise,} \end{cases} \quad (\text{F29})$$

which leads to a kinetic energy contribution [from the first line of Eq. (F20)]

$$\Delta E^{\text{kin}} = \frac{1}{2} \sum_{\mathbf{k} \in \mathcal{A}} \text{tr}[\delta Q_{\mathbf{k}}^T h^k(\mathbf{k})] = 2 \sum_{\mathbf{k} \in \mathcal{A}} \sum_{\eta} (\bar{E}_0^g(\mathbf{k}, \eta) + \eta \bar{E}_1^g(\mathbf{k}, \eta) w_0 D_0) = 2 \sum_{\mathbf{k} \in \mathcal{A}} \sum_{\eta} \bar{E}_0^g(\mathbf{k}, \eta), \quad (\text{F30})$$

where (although not crucial for the following conclusion), we made the natural assumption $\mathbf{k} \in \mathcal{A} \Leftrightarrow -\mathbf{k} \in \mathcal{A}$ in the last equality. Most importantly, we see that the kinetic energy contribution now scales as $(\Lambda + D_0)(\Delta k)^2$ to leading order in Δk , where Λ is the large energy of the graphene-like remote band, i.e., in the valley that does not have a Dirac cone at the respective K point \mathbf{k} is close to. As such, it is energetically not favorable to have valley polarization of the graphene-like bands around the K/K' points.

As is intuitively clear and can be derived in the same way, this suppression due to the kinetic energy also applies to the *IVC states* in Table II. One also finds exactly the same expression, $\Delta E^{\text{kin}} = 2 \sum_{\mathbf{k} \in \mathcal{A}} \sum_{\eta} \bar{E}_0^g(\mathbf{k}, \eta) \sim (\Lambda + D_0)(\Delta k)^2$. Consequently, ordering of the graphene-like bands in the vicinity of the K/K' points can also be excluded for the IVC states.

For the *sublattice polarized* states, SLP_{\pm} , the situation is slightly more complicated since the structure in the graphene-like bands is in general of the form

$$(Q_{\mathbf{k}}^g)_{\eta, \eta'} = \delta_{\eta, \eta'} \begin{cases} (-\sigma_3 \cos \theta_{\mathbf{k}} + \sigma_2 \sin \theta_{\mathbf{k}}) s_0, & \mathbf{k} \in \mathcal{A}_{\eta} \\ -\sigma_3 s_0, & \text{otherwise,} \end{cases} \quad (\text{F31})$$

and it is left to determine the optimal $\theta_{\mathbf{k}} \in \mathbb{R}$ subject to the constraint $\theta_{\mathbf{k}} = \mp \theta_{-\mathbf{k}}$ for the SLP_{\pm} state resulting from C_{2z} and Θ . Note that $(Q_{\mathbf{k}}^g)^2 = \mathbb{1}$ and $\text{tr} Q_{\mathbf{k}}^g = 0$, hold for any $\theta_{\mathbf{k}}$ and, hence, does not further constrain it.

Making the natural assumption that $\sin \theta_{\mathbf{k}}$ is significantly non-zero only in the small region \mathcal{A} around the K/K' points, we can still use Eq. (F20) which yields

$$\Delta E_{\text{SLP}_{\pm}}^{\text{HF}}[\theta_{\mathbf{k}}] = \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} [A_{\mathbf{k}, \eta} (1 - \cos \theta_{\mathbf{k}}) + B_{\mathbf{k}, \eta} \sin \theta_{\mathbf{k}}], \quad (\text{F32})$$

where

$$A_{\mathbf{k}, \eta} = 2\bar{E}_0^g(\mathbf{k}, \eta) + \frac{2}{N} \sum_{\mathbf{q}} V(\mathbf{q}) [(\bar{F}_{1, \eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2, \eta}^{gg}(\mathbf{k}, \mathbf{q}))^2] + \mathcal{O}(D_0^4), \quad (\text{F33a})$$

$$B_{\mathbf{k}, \eta} = -\frac{2}{N} \sum_{\mathbf{q}} V(\mathbf{q}) [(\bar{F}_{1, \eta}^{gb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2, \eta}^{gb}(\mathbf{k}, \mathbf{q}))^2] D_0^2 + \frac{2}{N} \sum_{\mathbf{q}} V(\mathbf{q}) [(\bar{F}_{7, \eta}^{gb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{8, \eta}^{gb}(\mathbf{k}, \mathbf{q}))^2] w_0^2 D_0^2. \quad (\text{F33b})$$

First, note that $B_{\mathbf{k}, \eta} = 0$ for $D_0 = 0$. Furthermore, the second term in Eq. (F33a) is suppressed to leading order in Δk , as already discussed above, such that $A_{\mathbf{k}, \eta} > 0$ and, thus, $\theta_{\mathbf{k}} = 0$ for $D_0 = 0$. This agrees with our result of Appendix B and with the HF band structure in Fig. 3 which exhibits a gapless graphene Dirac cone at $D_0 = 0$ for the SLP_- (the same applies for SLP_+ , not shown). More explicitly, this can also be seen in the upper panel of Fig. 17, where the correlators of the SLP_{\pm} states close to the K point at $D_0 = 0$ are shown: the graphene bands are just in the lower-filled-band configuration.

The system behaves differently when $D_0 \neq 0$. We then have $B_{\mathbf{k}, \eta} \neq 0$ and it will become energetically favorable to develop finite SLP_{\pm} order in the graphene-like bands according to $\tan \theta_{\mathbf{k}} = \tilde{B}_{\mathbf{k}}^{\pm} / \tilde{A}_{\mathbf{k}}$, $\mathbf{k} \in \mathcal{A}_{+}$, where we defined

$$\tilde{A}_{\mathbf{k}} = A_{\mathbf{k}, +} + A_{-\mathbf{k}, -}, \quad \tilde{B}_{\mathbf{k}}^{\pm} = B_{\mathbf{k}, +} \mp B_{-\mathbf{k}, -}, \quad (\text{F34})$$

with $+$ ($-$) for the SLP_+ (SLP_-) case. The associated minimized energy change is given by

$$\min_{\theta_{\mathbf{k}}} \Delta E_{\text{SLP}_{\pm}}^{\text{HF}}[\theta_{\mathbf{k}}] = \sum_{\mathbf{k} \in \mathcal{A}_+} \left[\tilde{A}_{\mathbf{k}} - \sqrt{\tilde{A}_{\mathbf{k}}^2 + (\tilde{B}_{\mathbf{k}}^{\pm})^2} \right]. \quad (\text{F35})$$

Taken together, both SLP_{\pm} states will develop order in the graphene-like bands around the K and K' points as long as D_0 is non-zero. As $\tilde{A} \sim D_0$ and $\tilde{B}^{\pm} \sim D_0^2$, the associated energy gain will scale as $-g_{\pm} V_{\mathcal{A}} |D_0|^3$ for small D_0 . As follows from Eq. (F33b), $B_{\mathbf{k},\eta}$ has a definite sign for sufficiently small w_0 , such that $|\tilde{B}_{\mathbf{k}}^-| > |\tilde{B}_{\mathbf{k}}^+|$ and, thus, $g_- > g_+$, i.e., the SLP_- state can gain more energy than the SLP_+ phase from this process for small w_0 . Since $\sin \theta_{\mathbf{k}} \neq 0$ in Eq. (F31) corresponds to a mixing of the upper and lower bands of the graphene-like bands, this will gap out the Dirac cones of the SLP_{\pm} states at $D_0 \neq 0$ and we obtain an insulator rather than a semimetal; this is also seen in our HF numerics (cf. Fig. 3). Furthermore, it is directly visible in the correlators of the SLP_{\pm} states shown in the lower panel of Fig. 17; the additional mixing of the TBG-like and graphene-like band that can also be seen in the HF data will be analyzed in Appendix F 4.

In analogy to these two states, we choose for the *quantum spin Hall state* (SSLP_-)

$$(Q_{\mathbf{k}}^g)_{\eta,\eta'} = \delta_{\eta,\eta'} \begin{cases} -\sigma_3 s_0 \cos \theta_{\mathbf{k}} + \sigma_2 s_3 \sin \theta_{\mathbf{k}}, & \mathbf{k} \in \mathcal{A}_{\eta} \\ -\sigma_3 s_0, & \text{otherwise.} \end{cases} \quad (\text{F36})$$

The corresponding change of the energy has again the form of Eq. (F32), this time with

$$A_{\mathbf{k},\eta} = 2\bar{E}_0^g(\mathbf{k},\eta) + \frac{2}{N} \sum_{\mathbf{q}} V(\mathbf{q}) [(\bar{F}_{1,\eta}^{gg}(\mathbf{k},\mathbf{q}))^2 - (\bar{F}_{2,\eta}^{gg}(\mathbf{k},\mathbf{q}))^2] + \mathcal{O}(w_0^2 D_0^4), \quad B_{\mathbf{k},\eta} = 0. \quad (\text{F37})$$

Note that the vanishing of $B_{\mathbf{k},\eta}$ is consistent with the fact that the energy should not depend on whether $\theta_{\mathbf{k}} = \theta_{-\mathbf{k}}$ (SSLP_-) or $\theta_{\mathbf{k}} = -\theta_{-\mathbf{k}}$ (SSLP_+), since these two states are related by a $\text{SU}(2)_+ \times \text{SU}(2)_-$ transformation. We clearly see that the energy is minimized by $\cos \theta_{\mathbf{k}} = 1$, i.e., no order in the graphene-like bands and no associated energetic gain. We emphasize that the presence of SSLP_- order in the TBG-like bands and, at $D_0 \neq 0$, interactions that couple the two subsystems, also induces a gap in the original graphene Dirac cone; we will see this in Appendix F 4 below where we will find a non-zero mixing of the TBG-like and graphene-like bands for the SSLP states.

4. Mixing between the TBG-like and graphene-like bands

We have not yet taken into account the possibility that the TBG-like and graphene-like bands can mix or, in other words, develop some coherence in the vicinity of the K/K' points; this corresponds to allowing for off-diagonal components in $Q_{\mathbf{k}}$. It is clear by symmetry that such a mixing is only possible for $D_0 \neq 0$.

To describe the mixing between the bands, we write

$$Q_{\mathbf{k}} = U_{\mathbf{k}} Q_{\mathbf{k}}^0 U_{\mathbf{k}}^{\dagger}, \quad Q_{\mathbf{k}}^0 = \begin{pmatrix} -\sigma_z \eta_0 s_0 & 0 \\ 0 & Q^b \end{pmatrix} \quad U_{\mathbf{k}} U_{\mathbf{k}}^{\dagger} = \mathbb{1}, \quad (\text{F38})$$

where Q^b is any of the order parameters in the TBG-like bands listed in Table II and the unitary matrix $U_{\mathbf{k}}$, which mixes the TBG-like and graphene-like bands, is constrained by the symmetries of the state Q^b . Note that unitarity of $U_{\mathbf{k}}$ guarantees $Q_{\mathbf{k}}^2 = \mathbb{1}$, $Q_{\mathbf{k}}^{\dagger} = Q_{\mathbf{k}}$, and $\text{tr} Q_{\mathbf{k}} = 0$.

Since are interested in mixing of the bands, we will focus on $U_{\mathbf{k}}$ of the form

$$U_{\mathbf{k}} = e^{i\varphi \lambda_{\mathbf{k}}} = e^{i\varphi \sum_{j=1,2} \sum_{j'} \hat{n}_{\mathbf{k}}^{(j,j')} M^{(j,j')} \zeta_j}, \quad \left(M^{(j,j')} \right)^{\dagger} = M^{(j,j')}, \quad \hat{n}_{\mathbf{k}}^{(j,j')} \in \mathbb{R} \quad (\text{F39})$$

where ζ_j are Pauli matrices acting between the TBG-like and graphene-like bands and $M^{(j,j')}$ are matrices in spin, band, and valley space. It, thus, holds $[M^{(j_1,j_2)}, \zeta_j] = 0$. Before discussing the different candidate states separately, where the different symmetries will constrain the available generators in Eq. (F39), let us expand Eq. (F38) up to second order in the rotational angle φ . To this end, define $\mathcal{M}_{j,\mathbf{k}} = \sum_{j'} \hat{n}_{\mathbf{k}}^{(j,j')} M^{(j,j')}$ and $\mathcal{M}_{\mathbf{k}} = \mathcal{M}_{1,\mathbf{k}} + i\mathcal{M}_{2,\mathbf{k}}$.

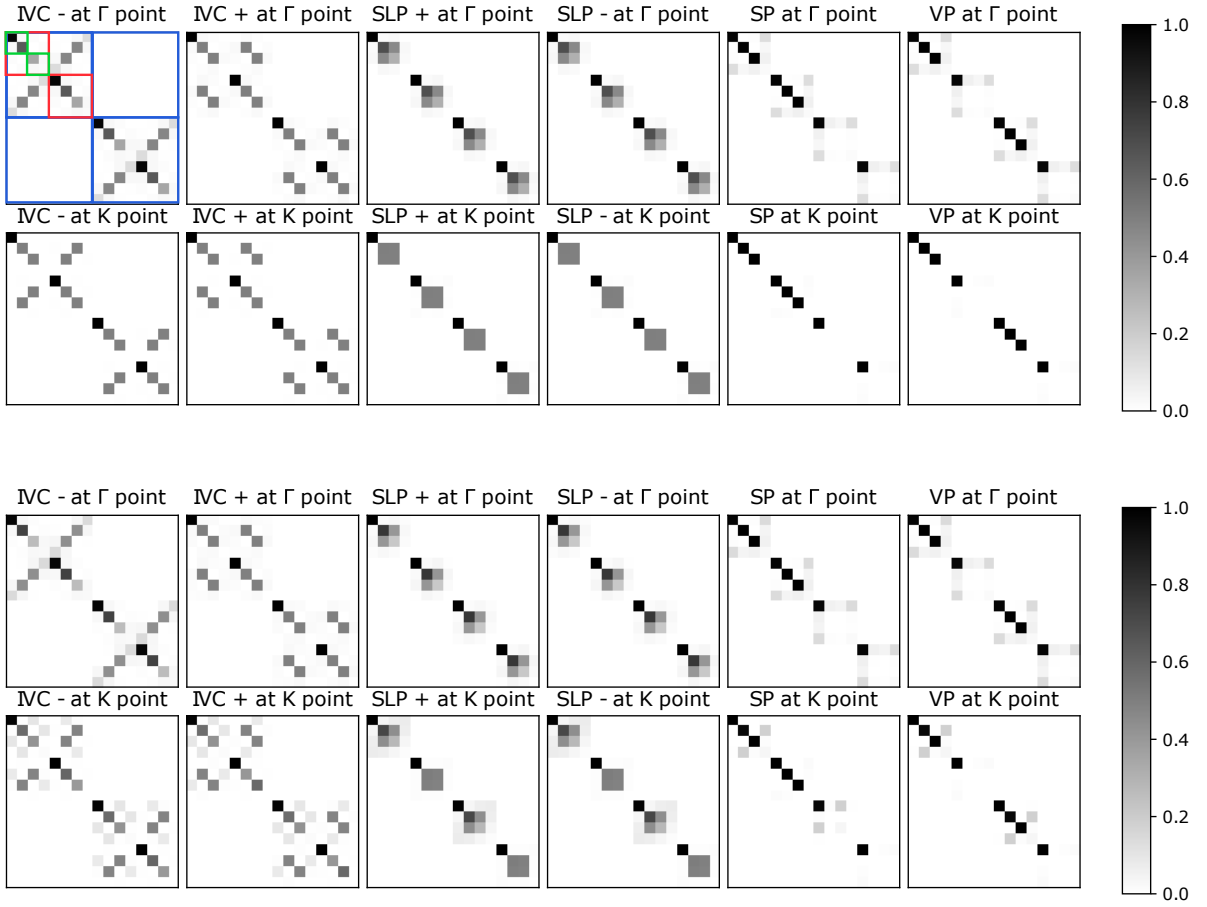


FIG. 17: We show the projector P which characterizes each of our symmetry breaking solution for a point near the Γ point of the Brillouin zone and K point of the Brillouin zone for $w_0/w_1 = 0.73$ and $D_0/w_1 = 0$ (top) and $D_0/w_1 = 0.5$ (bottom). We plot the matrix form of P such that the largest block shown outlined in blue in the first panel of the top row denotes spin flavor, the next smallest blocks outlined in red denote valley flavor, and the next smallest block outlined in green denotes upper and lower bands in the continuum model and final two boxes denoting TBG-like or graphene-like band index in the continuum model, where the final two boxes denote the graphene-like and TBG-like band in the upper-band box and the TBG-like then graphene-like bands in the lower-band box.

Equation (F38) then becomes

$$\begin{aligned}
 Q_{\mathbf{k}} \sim & \begin{pmatrix} -\sigma_z & 0 \\ 0 & Q^b \end{pmatrix} + \begin{pmatrix} 0 & i\mathcal{M}_{\mathbf{k}}^\dagger Q^b + i\sigma_z \mathcal{M}_{\mathbf{k}}^\dagger \\ -iQ^b \mathcal{M}_{\mathbf{k}} - i\mathcal{M}_{\mathbf{k}} \sigma_z & 0 \end{pmatrix} \varphi \\
 & + \left[\begin{pmatrix} \mathcal{M}_{\mathbf{k}}^\dagger Q^b \mathcal{M}_{\mathbf{k}} & 0 \\ 0 & -\mathcal{M}_{\mathbf{k}}^\dagger \sigma_z \mathcal{M}_{\mathbf{k}} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \{\mathcal{M}_{\mathbf{k}}^\dagger \mathcal{M}_{\mathbf{k}}, \sigma_z\} & 0 \\ 0 & \{\mathcal{M}_{\mathbf{k}} \mathcal{M}_{\mathbf{k}}^\dagger, Q^b\} \end{pmatrix} \right] \varphi^2 + \mathcal{O}(\varphi^3).
 \end{aligned} \tag{F40}$$

Below, we will use this expression to interpret the mixing matrices we obtain analytically and to constrain the possible generators $M^{(j,j')}$.

Since the IVC₋ is favored energetically in most of parameter space, let us start with the *intervalley coherent phases*, which can be discussed simultaneously. Due to the valley $U(1)_v$ symmetry, we can choose their “undeformed” correlator in the TBG-like subspace, without loss of generality, to be $Q^b = \eta_x$ and $Q^b = \sigma_y \eta_y$ for the IVC₊ and IVC₋, respectively. As such, the phases still preserve the $C_{2z} \Theta$ symmetry in Table V (for other choices, it has to be combined with a valley $U(1)_v$ rotation). Consequently, $M^{(j,j')}$ in Eq. (F39) must obey

$$(M^{(1,j')})^* = -M^{(1,j')}, \quad (M^{(2,j')})^* = M^{(2,j')}, \quad \Leftrightarrow \quad \mathcal{M}^* = -\mathcal{M}. \tag{F41}$$

Furthermore, the preserved spin-rotation symmetry forces M_j to be trivial in spin-space. This leaves the following

3 + 3 = 6 options for M_1 and $3 \times 3 + 1 = 10$ possibilities for M_2 ,

$$M^{(1,j')} \in \{\sigma_y \eta_{0,x,z}, \sigma_{0,x,z} \eta_y\}, \quad M^{(2,j')} \in \{\sigma_{0,x,z} \eta_{0,x,z}, \sigma_y \eta_y\}. \quad (\text{F42})$$

However, not all (linear combinations) of the 16 generators are important for the IVC states. First, we can neglect all parts of $\lambda_{\mathbf{k}}$ in Eq. (F39) that commute with $Q_{\mathbf{k}}^0$. This can be done formally by setting $\lambda_{\mathbf{k}} \rightarrow (\lambda_{\mathbf{k}} - Q_{\mathbf{k}}^0 \lambda_{\mathbf{k}} Q_{\mathbf{k}}^0)/2$ or, equivalently, by replacing

$$\mathcal{M}_{\mathbf{k}} \rightarrow \frac{1}{2}(\mathcal{M}_{\mathbf{k}} + Q^b \mathcal{M}_{\mathbf{k}} \sigma_z). \quad (\text{F43})$$

For future reference, this means that it holds

$$\mathcal{M}_{\mathbf{k}} \sigma_z = Q^b \mathcal{M}_{\mathbf{k}}. \quad (\text{F44})$$

This reduces the number of independent generators to 8. To begin with the IVC₊, a complete set of associated generators is given by

$$\begin{aligned} & \frac{1}{2} \zeta_y (\sigma_0 \eta_0 + \sigma_z \eta_x), \quad \frac{1}{2} \zeta_y (\sigma_z \eta_0 + \sigma_0 \eta_x), \quad \frac{1}{2} \zeta_y (\sigma_y \eta_y - \sigma_x \eta_z), \quad \frac{1}{2} \zeta_x (\sigma_y \eta_z + \sigma_x \eta_y), \\ & \frac{1}{2} (\zeta_y \sigma_x \eta_0 + \zeta_x \sigma_y \eta_x), \quad \frac{1}{2} (\zeta_y \sigma_0 \eta_z + \zeta_x \sigma_z \eta_y), \quad \frac{1}{2} (\zeta_y \sigma_x \eta_x + \zeta_x \sigma_y \eta_0), \quad \frac{1}{2} (\zeta_y \sigma_z \eta_z + \zeta_x \sigma_0 \eta_y). \end{aligned} \quad (\text{F45})$$

For the IVC₋, they read as

$$\begin{aligned} & \frac{1}{2} \zeta_y (\sigma_z \eta_0 + \sigma_y \eta_y), \quad \frac{1}{2} \zeta_y (\sigma_0 \eta_x + \sigma_x \eta_z), \quad \frac{1}{2} \zeta_y (\sigma_x \eta_x - \sigma_0 \eta_z), \quad \frac{1}{2} \zeta_x (\sigma_y \eta_0 + \sigma_z \eta_y), \\ & \frac{1}{2} (\zeta_x \sigma_y \eta_x - \zeta_y \sigma_z \eta_z), \quad \frac{1}{2} (\zeta_x \sigma_y \eta_z + \zeta_y \sigma_z \eta_x), \quad \frac{1}{2} (\zeta_x \sigma_0 \eta_y + \zeta_y \sigma_x \eta_0), \quad \frac{1}{2} (\zeta_x \sigma_x \eta_y - \zeta_y \sigma_0 \eta_0). \end{aligned} \quad (\text{F46})$$

To further constrain the remaining 8 generators, we note that the graphene-like bands in valley $\eta = -$ (valley $\eta = +$) are at energies far away from the Fermi surface (of scale Λ) for $\mathbf{k} \in \mathcal{A}_+$ ($\mathbf{k} \in \mathcal{A}_-$). In the following, we will neglect all mixing processes that are suppressed in the limit of large Λ . Inspection of the first term in Eq. (F20) shows that this is equivalent to demanding that $U_{\mathbf{k}}$ obey

$$\Delta_{\Lambda}(\eta) = \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} \text{tr}_g [(U_{\mathbf{k}} Q_{\mathbf{k}}^0 U_{\mathbf{k}}^{\dagger} - Q_{\mathbf{k}}^0) \hat{\Lambda}_{\eta}] = 0, \quad \hat{\Lambda}_{\pm} = \Lambda \sigma_z (\eta_0 \mp \eta_3)/2, \quad (\text{F47})$$

where the trace is only over the graphene-like subspace. Using Eq. (F40), we can express this quantity as

$$\Delta_{\Lambda}(\eta) = \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} \left(\text{tr}_g [\mathcal{M}_{\mathbf{k}}^{\dagger} Q^b \mathcal{M}_{\mathbf{k}} \hat{\Lambda}_{\eta}] + \frac{1}{2} \text{tr}_g [\{\mathcal{M}_{\mathbf{k}}^{\dagger} \mathcal{M}_{\mathbf{k}}, \sigma_z\} \hat{\Lambda}_{\eta}] \right) \varphi^2 + \mathcal{O}(\varphi^3). \quad (\text{F48})$$

To simplify further, we use Eq. (F44) and obtain

$$\Delta_{\Lambda}(\eta) = 2 \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} \text{tr}_{g,-\eta} [\mathcal{M}_{\mathbf{k}}^{\dagger} \mathcal{M}_{\mathbf{k}}] \varphi^2 + \mathcal{O}(\varphi^3), \quad (\text{F49})$$

where the trace is only over the graphene-like subspace in valley $-\eta$. As such, $\Delta_{\Lambda}(\eta) = 0$ requires

$$\text{tr}_{g,\mp} [\mathcal{M}_{\mathbf{k}}^{\dagger} \mathcal{M}_{\mathbf{k}}] = \sum_{s,s',p,p',\eta'} |(\mathcal{M}_{\mathbf{k}})_{(p',\eta',s'),(p,\mp,s)}|^2 = 0 \quad \forall \mathbf{k} \in \mathcal{A}_{\pm} \Leftrightarrow (\mathcal{M}_{\mathbf{k}})_{(p',\eta',s'),(p,\mp,s)} = 0 \quad \forall \mathbf{k} \in \mathcal{A}_{\pm}. \quad (\text{F50})$$

It is straightforward to see that there are 4 linear combinations of the 8 generators in Eq. (F45) (in Eq. (F46)) for the IVC₊ (IVC₋) that satisfy this requirement. For $\mathbf{k} \in \mathcal{A}_{\pm}$, these are

$$\frac{1}{2\sqrt{2}} [\zeta_y (\sigma_0 (\eta_0 \pm \eta_z) + \sigma_z \eta_x) \pm \zeta_x \sigma_z \eta_y], \quad \frac{1}{2\sqrt{2}} [\zeta_x (\sigma_y (\eta_0 \pm \eta_z) \pm \sigma_x \eta_y) + \zeta_y \sigma_x \eta_x], \quad (\text{F51a})$$

$$\frac{1}{2\sqrt{2}} [\zeta_y (\sigma_z (\eta_0 \pm \eta_z) + \sigma_0 \eta_x) \pm \zeta_x \sigma_0 \eta_y], \quad \frac{1}{2\sqrt{2}} [\zeta_y (\sigma_x (\eta_0 \pm \eta_z) \mp \sigma_y \eta_y) + \zeta_x \sigma_y \eta_x], \quad (\text{F51b})$$

for the IVC₊ and

$$\frac{1}{2\sqrt{2}} [\zeta_x(\sigma_y(\eta_0 \pm \eta_z) + \sigma_z\eta_y) \pm \zeta_y\sigma_z\eta_x], \quad \frac{1}{2\sqrt{2}} [\zeta_y(\sigma_0(\eta_0 \pm \eta_z) \mp \sigma_x\eta_x) - \zeta_x\sigma_x\eta_y], \quad (\text{F52a})$$

$$\frac{1}{2\sqrt{2}} [\zeta_y(\sigma_z(\eta_0 \pm \eta_z) + \sigma_y\eta_y) \mp \zeta_x\sigma_y\eta_x], \quad \frac{1}{2\sqrt{2}} [\zeta_y(\sigma_x(\eta_0 \pm \eta_z) \mp \sigma_0\eta_x) + \zeta_x\sigma_0\eta_y], \quad (\text{F52b})$$

for the IVC₋.

Taken together, we use $U_{\mathbf{k}} = e^{i\varphi_{\mathbf{k}}\lambda_{\pm}}$ for $\mathbf{k} \in \mathcal{A}_{\pm}$ and $U_{\mathbf{k}} = \mathbb{1}$ otherwise, where λ_{\pm} is any of the four generators in Eq. (F51) for the IVC₊ and any of those in Eq. (F52) for the IVC₋. From Eq. (F20), we can compute the associated change of energy that is found to be of the form

$$\Delta E^{\text{HF}} = \sum_{\mathbf{k} \in \mathcal{A}_+} \alpha_{\mathbf{k}} \sin^2 \varphi_{\mathbf{k}} + \beta_{\mathbf{k}} \sin \varphi_{\mathbf{k}} \cos \varphi_{\mathbf{k}}; \quad (\text{F53})$$

to reduce the summation to momenta in \mathcal{A}_+ only, we have used $\varphi_{\mathbf{k}} = \varphi_{-\mathbf{k}}$. For the IVC₊ with $Q^b = \eta_x$, this follows from the fact that C_{2z} as given in Table V acts as η_x and that $\eta_x\lambda_{\pm}\eta_x = \lambda_{\mp}$ for all generators in Eq. (F51). For the IVC₋ with $Q^b = \sigma_y\eta_y$, C_{2z} has to be combined with the $U(1)_v$ valley rotation $i\eta_z$ to be a symmetry (with action on the spinors given by $\eta_y\sigma_0$). However, as follows from Eq. (F52), it holds $\eta_y\lambda_{\pm}\eta_y = \lambda_{\mp}$ in this case, leading to the same result, $\varphi_{\mathbf{k}} = \varphi_{-\mathbf{k}}$.

For ease of presentation, let us discuss the behavior of ΔE^{HF} separately for the two IVCs and begin with the IVC₊ state. Focusing on the limit $w_0 = 0$ for now, we obtain $\beta_{\mathbf{k}} = 0$ (and $\alpha_{\mathbf{k}} > 0$ at least in a finite range of D_0 around $D_0 = 0$) for the generators in Eq. (F51b) such that ΔE^{HF} reaches its minimum $\Delta E^{\text{HF}} = 0$ when $\sin \varphi_{\mathbf{k}} = 0$ (no mixing). For the first and the second generator in Eq. (F51a), we find

$$\alpha_{\mathbf{k}} = 4 \left(2\epsilon_{(g,+),+}(\mathbf{k}) - \sum_{\eta=\pm} \epsilon_{(b,+),\eta}(\mathbf{k}) \right) + \frac{8}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sum_{j=1}^2 (\bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}))^2 - \sum_{j=1}^4 (-1)^j (\bar{F}_j^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] + \mathcal{O}(D_0^2), \quad (\text{F54a})$$

$$\beta_{\mathbf{k}} = \frac{8\sqrt{2}}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sum_{j=1}^4 \bar{F}_j^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_j^{gb}(\mathbf{k}, \mathbf{q}) - \sum_{j=1}^2 \bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_j^{bg}(\mathbf{k}, \mathbf{q}) \right] D_0 + \mathcal{O}(D_0^3), \quad (\text{F54b})$$

and

$$\alpha_{\mathbf{k}} = 4 \left(2\epsilon_{(g,+),+}(\mathbf{k}) + \sum_{\eta=\pm} \epsilon_{(b,+),\eta}(\mathbf{k}) \right) + \frac{8}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sum_{j=1}^2 (\bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}))^2 - \sum_{j=1}^4 (-1)^j (\bar{F}_j^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] + \mathcal{O}(D_0^2), \quad (\text{F55a})$$

$$\beta_{\mathbf{k}} = \frac{8\sqrt{2}}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_2^{bg}(\mathbf{k}, \mathbf{q}) - \bar{F}_2^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_1^{bg}(\mathbf{k}, \mathbf{q}) + \bar{F}_2^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_1^{gb}(\mathbf{k}, \mathbf{q}) \right. \\ \left. - \bar{F}_1^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_2^{gb}(\mathbf{k}, \mathbf{q}) + \bar{F}_4^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_3^{gb}(\mathbf{k}, \mathbf{q}) - \bar{F}_3^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_4^{gb}(\mathbf{k}, \mathbf{q}) \right] D_0 + \mathcal{O}(D_0^3), \quad (\text{F55b})$$

respectively. To arrive at these expressions, we have used the C_{2z} constraint in Eq. (A35).

We first observe that, in both cases, $\beta_{\mathbf{k}} = 0$ for $D_0 = 0$ and, hence, the energy (F53) is simply minimized when $\sin \varphi_{\mathbf{k}} = 0$, i.e., no mixing between the graphene and TBG bands occurs. This is expected since the presence of σ_h implies that the order parameters have to be either even (intra-system) or odd (inter-system) under σ_h , prohibiting mixing between the bands.

When $D_0 \neq 0$, the system can immediately gain energy by allowing for non-zero $\varphi_{\mathbf{k}}$: ΔE^{HF} in Eq. (F53) is minimized by $\varphi_{\mathbf{k}} = \varphi_{\mathbf{k}}^0$ obeying $\tan 2\varphi_{\mathbf{k}}^0 = -\beta_{\mathbf{k}}/\alpha_{\mathbf{k}}$ with minimal energy given by

$$\Delta E_0^{\text{HF}} = \min_{\varphi_{\mathbf{k}}} \Delta E^{\text{HF}}[\varphi_{\mathbf{k}}] = \frac{1}{2} \sum_{\mathbf{k} \in \mathcal{A}_+} \left(\alpha_{\mathbf{k}} - \sqrt{\alpha_{\mathbf{k}}^2 + \beta_{\mathbf{k}}^2} \right) \leq 0. \quad (\text{F56})$$

As long as $\beta_{\mathbf{k}} \neq 0$, this lowers the energy, which shows that the system will exhibit hybridization between the bands for $D_0 \neq 0$ as described by the two generators in Eq. (F51a), while those in Eq. (F51b) are disfavored energetically

by the interactions. While this only holds for $w_0 = 0$, we still expect the generators in Eq. (F51a) to be dominant in an extended region of finite w_0 . As we will see shortly, this is confirmed by our numerics.

To obtain the scaling of the energy gain with (small) D_0 , we expand Eq. (F56) in $\beta_{\mathbf{k}} \propto D_0$, which yields

$$\Delta E_0^{\text{HF}} \propto - \sum_{\mathbf{k} \in \mathcal{A}_+} \frac{\beta_{\mathbf{k}}^2}{4\alpha_{\mathbf{k}}} \propto -g_3 V_{\mathcal{A}} D_0^2, \quad (\text{F57})$$

where we indicated that the result will be proportional to the area $V_{\mathcal{A}} \propto (\Delta k)^2$ (our small parameter) and the displacement field squared; the constant $g_1 > 0$ contains all the non-universal properties and itself depends on D_0 and w_0 .

Which of these two remaining generators in Eq. (F51a) (or which linear combination) is preferred by the system is the *only* aspect that depends on the microscopic details of the form factors; it is not possible to say which of the $\beta_{\mathbf{k}}$ in Eqs. (F61b) and (F55b) is dominant at a given \mathbf{k} without computing the form factors in the microscopic model. The associated mixing of the bands for these two cases are readily evaluated from Eq. (F40). We get

$$Q^b \mathcal{M}_{\mathbf{k}} + \mathcal{M}_{\mathbf{k}} \sigma_z = 2\mathcal{M}_{\mathbf{k}} \sigma_z = \sqrt{2}i\sigma_z(\eta_0 \pm \eta_z) + \sqrt{2}\sigma_0(i\eta_x \pm \eta_y), \quad \mathbf{k} \in \mathcal{A}_{\pm} \quad (\text{F58})$$

for the first term in Eq. (F51a). This means that the graphene-like band with positive (negative) energy in the valley that is closer to the Fermi level mixes with the TBG-like band of both valleys that are also at positive (negative) energies.

For the second generator in Eq. (F51a), we find a mixing matrix of the form

$$Q^b \mathcal{M}_{\mathbf{k}} + \mathcal{M}_{\mathbf{k}} \sigma_z = \sqrt{2}i\sigma_y(\eta_0 \pm \eta_z) - \sqrt{2}i\sigma_x(i\eta_x \pm \eta_y), \quad \mathbf{k} \in \mathcal{A}_{\pm}. \quad (\text{F59})$$

Instead, we here get that the graphene-like band with positive (negative) energy in the valley that is closer to the Fermi level mixes with the TBG-like bands of both valleys that are at negative (positive) energies. In our numerical HF computations, we find very good agreement with the mixing matrix in Eq. (F59), see Fig. 17, even at the moderately large value of $w_0/w_1 = 0.73$.

The energetic discussion can be done in a similar way for the IVC₋ state and we find quite similar structures: first, we get $\beta_{\mathbf{k}} = 0$ for $w_0 = 0$ for the two generators in Eq. (F52b) and, thus, no energetic gain via mixing. Second, this is different for those in Eq. (F52a): the first one leads to

$$\alpha_{\mathbf{k}} = 4 \left(2\epsilon_{(g,+),+}(\mathbf{k}) + \sum_{\eta=\pm} \eta \epsilon_{(b,+),\eta}(\mathbf{k}) \right) + \frac{8}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sum_{j=1}^2 (\bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}))^2 - \sum_{j=1}^4 (-1)^j (\bar{F}_j^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] + \mathcal{O}(D_0^2), \quad (\text{F60a})$$

$$\beta_{\mathbf{k}} = \frac{8\sqrt{2}}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_2^{bg}(\mathbf{k}, \mathbf{q}) - \bar{F}_2^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_1^{bg}(\mathbf{k}, \mathbf{q}) + \bar{F}_2^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_1^{gb}(\mathbf{k}, \mathbf{q}) \right. \\ \left. - \bar{F}_1^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_2^{gb}(\mathbf{k}, \mathbf{q}) + \bar{F}_4^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_3^{gb}(\mathbf{k}, \mathbf{q}) - \bar{F}_3^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_4^{gb}(\mathbf{k}, \mathbf{q}) \right] D_0 + \mathcal{O}(D_0^3), \quad (\text{F60b})$$

while the second one is associated with the coefficients

$$\alpha_{\mathbf{k}} = 4 \left(2\epsilon_{(g,+),+}(\mathbf{k}) - \sum_{\eta=\pm} \eta \epsilon_{(b,+),\eta}(\mathbf{k}) \right) + \frac{8}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sum_{j=1}^2 (\bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}))^2 - \sum_{j=1}^4 (-1)^j (\bar{F}_j^{gg}(\mathbf{k}, \mathbf{q}))^2 \right] + \mathcal{O}(D_0^2), \quad (\text{F61a})$$

$$\beta_{\mathbf{k}} = \frac{8\sqrt{2}}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[\sum_{j=1}^4 \bar{F}_j^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_j^{gb}(\mathbf{k}, \mathbf{q}) - \sum_{j=1}^2 \bar{F}_j^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_j^{bg}(\mathbf{k}, \mathbf{q}) \right] D_0 + \mathcal{O}(D_0^3) \quad (\text{F61b})$$

in Eq. (F53). As before, while $D_0 = 0$ does not lead to mixing, the bands start to mix at non-zero D_0 ; this mixing is described by the two generators in Eq. (F52a) and which of the two dominates depends on microscopic details. The nature of the mixing depends on the generator. For the first one in Eq. (F52a), the mixing matrix reads as

$$2\mathcal{M}_{\mathbf{k}} \sigma_z = \sqrt{2}i\sigma_x(\eta_0 \pm \eta_z) + \sqrt{2}\sigma_0(\eta_y \pm i\eta_x), \quad \mathbf{k} \in \mathcal{A}_{\pm}. \quad (\text{F62})$$

This means that the lower/upper graphene-like band in the valley that is at low energies hybridizes with the upper/lower (lower/upper) TBG-like band of the same (opposite) valley. For the second generator in Eq. (F52a), we get instead

$$2\mathcal{M}_{\mathbf{k}}\sigma_z = \sqrt{2}i\sigma_z(\eta_0 \pm \eta_z) + \sqrt{2}i\sigma_y(\eta_y \pm i\eta_x), \quad \mathbf{k} \in \mathcal{A}_{\pm}, \quad (\text{F63})$$

i.e., the lower/upper graphene-like band in the valley that is at low energies hybridizes with the upper/lower (lower/upper) TBG-like band of the opposite (same) valley.

Comparison of Eqs. (F55) and (F60) [the same holds for Eqs. (F54) and (F61)] shows that the corresponding prefactors and, hence, the energies are identical to leading order in D_0 (note the subleading corrections due to the kinetic term). Since our numerics has identified the mixing matrix in Eq. (F59) to be realized in the system for the IVC₊ and the parameters in Fig. 17, we thus expect Eq. (F62) to dominate for the IVC₋; this is indeed confirmed by Fig. 17. Another consequence of this observation is that the leading (in D_0 and w_0) correction to the energy of the IVC₋ is identical to that of the IVC₊, given by $-g_1V_{\mathcal{A}}D_0^2$. As we have seen above, further corrections to higher orders in D_0 (e.g., due to the kinetic terms) are different for these two IVC states. These are not included in Table III.

We next continue with the *spin polarized* state, $Q^b = \sigma_0\eta_0s_3$. Since the valley $U(1)_v$ symmetry is preserved, $U_{\mathbf{k}}$ has to be diagonal in valley space. We can write

$$(U_{\mathbf{k}})_{\eta,\eta'} = s_0\sigma_0\delta_{\eta,\eta'}e^{i\sum_{j=1}^3\varphi_{\mathbf{k},\eta}^{(j)}\zeta_j}. \quad (\text{F64})$$

While this form of $U_{\mathbf{k}}$ automatically preserves all unitary symmetries in Table V (except for P , which, however, is broken anyway for $D_0 \neq 0$), it only preserves $C_{2z}\Theta$ if $\varphi^{(1)} = \varphi^{(3)} = 0$ and we are left with

$$(U_{\mathbf{k}})_{\eta,\eta'} = s_0\sigma_0\delta_{\eta,\eta'}e^{i\varphi_{\mathbf{k},\eta}\zeta_2} \quad (\text{F65})$$

with the additional constraint $\varphi_{\mathbf{k},\eta} = \varphi_{-\mathbf{k},-\eta}$, which comes from Θ (or, equivalently, C_{2z}). Our task now it to determine the real-valued function $\varphi_{\mathbf{k},\eta}$ such that the energy is minimized. We expect the mixing to be the strongest (if at all present) around the K and K' points for valley $\eta = +$ and $\eta = -$, respectively, and, thus, assume that $\varphi_{\mathbf{k},\eta} \neq 0$ only in the region \mathcal{A}_{η} and expand in its area leading to Eq. (F20). The energy change due to $U_{\mathbf{k}} \neq \pm 1$ in Eq. (F65) for the spin polarized state is found to be

$$\Delta E[\varphi_{\mathbf{k},\eta}] = \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} (a_{\mathbf{k},\eta} \sin^2 \varphi_{\mathbf{k},\eta} + b_{\mathbf{k},\eta} \sin \varphi_{\mathbf{k},\eta} \cos \varphi_{\mathbf{k},\eta}) \quad (\text{F66})$$

$$= \sum_{\mathbf{k} \in \mathcal{A}_+} (a_{\mathbf{k}} \sin^2 \varphi_{\mathbf{k},+} + b_{\mathbf{k}} \sin \varphi_{\mathbf{k},+} \cos \varphi_{\mathbf{k},+}), \quad a_{\mathbf{k}} := \sum_{\eta} a_{\eta\mathbf{k},\eta}, \quad b_{\mathbf{k}} := \sum_{\eta} b_{\eta\mathbf{k},\eta}, \quad (\text{F67})$$

where we used the constraint $\varphi_{\mathbf{k},\eta} = \varphi_{-\mathbf{k},-\eta}$ in the second line. Here and in all of the following computations of this appendix, we focus on terms that do not vanish for $w_0 = 0$. This is motivated by the fact that the final energy gain will already be of order D_0^2 and so further corrections of order w_0 of its prefactor will be subleading. We find

$$a_{\mathbf{k}} = 2 \sum_{p=\pm} p (\epsilon_{(g,p),+}(\mathbf{k}) - \epsilon_{(b,p),+}(\mathbf{k})) + \frac{4}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{1,+}^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,+}^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{1,+}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2,+}^{gg}(\mathbf{k}, \mathbf{q}))^2 \right], \quad (\text{F68a})$$

$$- \frac{4}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{1,+}^{bg}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{1,+}^{gb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,+}^{bg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2,+}^{gb}(\mathbf{k}, \mathbf{q}))^2 \right] D_0^2$$

$$b_{\mathbf{k}} = -\frac{8}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{j=1,2} \left[\bar{F}_{j,+}^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_{j,+}^{bg}(\mathbf{k}, \mathbf{q}) - \bar{F}_{j,+}^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_{j,+}^{gb}(\mathbf{k}, \mathbf{q}) \right] D_0. \quad (\text{F68b})$$

We see that $b_{\mathbf{k},\eta} = 0$ if $D_0 = 0$ while $a_{\mathbf{k},\eta} > 0$. So $\sin \varphi_{\mathbf{k},\eta} = 0$ minimizes the energy and, hence, there is no mixing between the bands for $D_0 = 0$, as expected by symmetry. However, any arbitrarily small non-zero D_0 leads to $b_{\mathbf{k},\eta} \neq 0$ and the energy is lowered when the bands hybridize: Eq. (F66) is minimized by $\varphi_{\mathbf{k},\eta}^0$ obeying $\tan 2\varphi_{\mathbf{k},\eta}^0 = -b_{\mathbf{k},\eta}/a_{\mathbf{k},\eta}$ and with energy change

$$\min_{\varphi_{\mathbf{k},+}} \Delta E[\varphi_{\mathbf{k},+}] = \frac{1}{2} \sum_{\mathbf{k} \in \mathcal{A}_+} \left(a_{\mathbf{k}} - \sqrt{a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2} \right) \leq 0. \quad (\text{F69})$$

Since $b_{\mathbf{k}} \propto D_0$, the energy correction scales as $-g_1 V_{\mathcal{A}} D_0^2$, as indicated schematically in Table III.

Valley polarized. This state also preserves the $C_{2z}\Theta$ and valley $U(1)_v$ symmetries, such that $U_{\mathbf{k}}$ is again of the form of Eq. (F65); this time, however, without the constraint $\varphi_{\mathbf{k},\eta} = \varphi_{-\mathbf{k},-\eta}$ as C_{2z} and Θ are both broken individually. The energy change is of the form of Eq. (F66) with coefficients given by (again, only keeping terms to zeroth order in w_0)

$$a_{\mathbf{k},\eta} = 2 \left(\epsilon_{(g,+),\eta}(\mathbf{k}) - \epsilon_{(b,+),\eta}(\mathbf{k}) \right) + \frac{2}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{1,\eta}^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,\eta}^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{1,\eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2,\eta}^{gg}(\mathbf{k}, \mathbf{q}))^2 \right], \quad (\text{F70a})$$

$$- \frac{2}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{1,\eta}^{bg}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{1,\eta}^{gb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,\eta}^{bg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2,\eta}^{gb}(\mathbf{k}, \mathbf{q}))^2 \right] D_0^2$$

$$b_{\mathbf{k},\eta} = -\frac{4}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{j=1,2} \left[\bar{F}_{j,\eta}^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_{j,\eta}^{bg}(\mathbf{k}, \mathbf{q}) - \bar{F}_{j,\eta}^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_{j,\eta}^{gb}(\mathbf{k}, \mathbf{q}) \right] D_0. \quad (\text{F70b})$$

Minimization yields

$$\min_{\varphi_{\mathbf{k},\eta}} \Delta E[\varphi_{\mathbf{k},\eta}] = \frac{1}{2} \sum_{\eta} \sum_{\mathbf{k} \in \mathcal{A}_{\eta}} \left(a_{\mathbf{k},\eta}^2 - \sqrt{a_{\mathbf{k},\eta}^2 + b_{\mathbf{k},\eta}^2} \right) \leq 0. \quad (\text{F71})$$

It is easy to see that $a_{\mathbf{k},\eta} = a_{-\mathbf{k},-\eta}$ and $b_{\mathbf{k},\eta} = b_{-\mathbf{k},-\eta}$, following from C_{2z} symmetry of the form factors and dispersion, and that $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ in Eq. (F68) are given by $a_{\mathbf{k}} = 2a_{\mathbf{k},+}$ and $b_{\mathbf{k}} = 2b_{\mathbf{k},+}$. As such the energy gain is the same for the spin and for the valley polarized states at $w_0 = 0$.

Next, we discuss the *quantum Hall state* SLP_- . While this state also preserves valley $U(1)_v$ symmetry and spin rotations, it breaks the \mathbf{k} -space-local $C_{2z}\Theta$ symmetry. As such, the transformation has to be of the more general form in Eq. (F64), involving three generators, $\zeta_{x,y,y}$. Due to the preserved C_{2z} symmetry, we have $\varphi_{\mathbf{k},+}^{(j)} = \varphi_{-\mathbf{k},-}^{(j)}$. Exactly as before, general energetic reasoning leads to $\varphi_{\mathbf{k},\eta}^{(j)} = 0$ if $\mathbf{k} \notin \mathcal{A}_{\eta}$. Instead of optimizing all three associated angles $\varphi_{\mathbf{k},+}^{(j)}$ simultaneously, let us first investigate them one at a time: first of all, $j = 3$ leads to $\Delta E^{\text{HF}} = 0$ (for $w_0 = 0$) in Eq. (F20). The generator ζ_x does lead to non-zero ΔE^{HF} , which is found to be of the form of Eq. (F66); this time with $b_{\mathbf{k},\eta} = 0$ and $a_{\mathbf{k},\eta}$ as given in Eq. (F70a). Since $a_{\mathbf{k},\eta} > 0$, at least for sufficiently small D_0 , the system does not gain any energy as a consequence of hybridization described by $(U_{\mathbf{k}})_{\eta,\eta'} = \delta_{\eta,\eta'} e^{i\varphi_{\mathbf{k},\eta} \zeta_x}$. The only generator left is $j = 2$, i.e., $(U_{\mathbf{k}})_{\eta,\eta'} = \delta_{\eta,\eta'} e^{i\varphi_{\mathbf{k},\eta} \zeta_y}$; in fact, this one allows for a reduction of energy and we obtain, to order w_0^0 , exactly the same of $a_{\mathbf{k},\eta}$ and $b_{\mathbf{k},\eta}$ as defined in Eq. (F70). In order to take into account the C_{2z} constraint, $\varphi_{\mathbf{k},\eta} = \varphi_{-\mathbf{k},-\eta}$, we again simplify from Eq. (F66) to Eq. (F67). Upon noting, as above, that $a_{\mathbf{k},\eta} = a_{-\mathbf{k},-\eta}$ and $b_{\mathbf{k},\eta} = b_{-\mathbf{k},-\eta}$, we have $a_{\mathbf{k}} = 2a_{\mathbf{k},+}$ and $b_{\mathbf{k}} = 2b_{\mathbf{k},+}$ and the energy reduction due to hybridization of the SLP_- is the same as that of the VP state (at $w_0 = 0$).

Upon noting that $\mathcal{M}_{\mathbf{k}} = i\sigma_0 s_0 (\eta_0 \pm \eta_z)$, $\mathbf{k} \in \mathcal{A}_{\pm}$, in Eq. (F40) for our $U_{\mathbf{k}}$ studied here, it is easy to compute the mixing matrix:

$$Q^b \mathcal{M}_{\mathbf{k}} + \mathcal{M}_{\mathbf{k}} \sigma_z = i(\sigma_y + \sigma_z)(\eta_0 \pm \eta_z), \quad \mathbf{k} \in \mathcal{A}_{\pm}, \quad (\text{F72})$$

i.e., both graphene-like bands in the valley that is at low energies mixes with both TBG-like bands in the same valley. This agrees well with the correlator near the K point of the SLP_- state obtained in HF, see Fig. 17.

To continue with the *valley Hall state* SLP_+ , we note that $C_{2z}\Theta$ is also broken in this state, leading again to the general form of $U_{\mathbf{k}}$ given in Eq. (F64) with three generators. This time Θ (rather than C_{2z}) is preserved and we have $\varphi_{\mathbf{k},+}^{(j)} = (-1)^j \varphi_{-\mathbf{k},-}^{(j)}$. As above, we study first each generator, $j = 1, 2, 3$, separately and note that $j = 1$ and $j = 3$ do not allow to lower the energy (again, for $w_0 = 0$ and not too large values of D_0). However, $j = 2$ does allow to lower the energy via hybridization for any non-zero D_0 . Writing $\varphi_{\mathbf{k},+} = \varphi_{\mathbf{k},+}^{(2)}$, the energy gain is of the form (F67) with

prefactors

$$\begin{aligned}
a_{\mathbf{k}} &= 2 \sum_{p=\pm} p \left(\epsilon_{(g,p),+}(\mathbf{k}) - \epsilon_{(b,p),+}(\mathbf{k}) \right) \\
&+ \frac{4}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{1,+}^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,+}^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{1,+}^{gg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2,+}^{gg}(\mathbf{k}, \mathbf{q}))^2 \right], \\
&- \frac{4}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \left[(\bar{F}_{1,+}^{bg}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{1,+}^{gb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_{2,+}^{bg}(\mathbf{k}, \mathbf{q}))^2 - (\bar{F}_{2,+}^{gb}(\mathbf{k}, \mathbf{q}))^2 \right] D_0^2
\end{aligned} \tag{F73a}$$

$$\begin{aligned}
&+ \frac{32}{N} \sum_{\mathbf{G}} V(\mathbf{G}) \left(\sum_{\mathbf{k}'} \bar{F}_4^{bb}(\mathbf{k}', \mathbf{G}) \right) (\bar{F}_{2,+}^{bb}(\mathbf{k}, \mathbf{G}) - \bar{F}_{2,+}^{gg}(\mathbf{k}, \mathbf{G})) D_0^2 \\
b_{\mathbf{k}} &= -\frac{8}{N} \sum_{\mathbf{q}} V(\mathbf{q}) \sum_{j=1,2} \left[\bar{F}_{j,+}^{bb}(\mathbf{k}, \mathbf{q}) \bar{F}_{j,+}^{bg}(\mathbf{k}, \mathbf{q}) - \bar{F}_{j,+}^{gg}(\mathbf{k}, \mathbf{q}) \bar{F}_{j,+}^{gb}(\mathbf{k}, \mathbf{q}) \right] D_0.
\end{aligned} \tag{F73b}$$

Comparison with the expressions for the VP and SLP₋ shows that the energy reduction due to hybridization is almost the same for the SLP₊. The only difference is the additional Hartree term in $a_{\mathbf{k}}$, which is expected to be only a very small correction due to Eq. (A38) and since it is higher order in D_0 . As such, we have $g_3 \simeq g_1$ in Table III.

The corresponding $\mathcal{M}_{\mathbf{k}} = i\sigma_0 s_0 (\eta_0 \pm \eta_z)$, $\mathbf{k} \in \mathcal{A}_{\pm}$, in Eq. (F40) leads to the mixing contribution in $Q_{\mathbf{k}}$

$$Q^b \mathcal{M}_{\mathbf{k}} + \mathcal{M}_{\mathbf{k}} \sigma_z = i(\sigma_y \eta_z + \sigma_z)(\eta_0 \pm \eta_z), \quad \mathbf{k} \in \mathcal{A}_{\pm}. \tag{F74}$$

This means that both the upper and the lower graphene-like bands in the valley that is at low energies mixes with both TBG-like bands in the same valley, in perfect agreement with Fig. 17.

Finally, the situation is slightly more complex for the SSLP₋ state, as it also breaks SU(2) spin-rotation symmetry down to residual rotation along one spin axis (say s_z). Consequently, we will have to allow for six generators,

$$(U_{\mathbf{k}})_{\eta, \eta'} = \sigma_0 \delta_{\eta, \eta'} e^{i \sum_{j=1}^3 (\varphi_{\mathbf{k}, \eta}^{(j,0)} \zeta_j s_0 + \varphi_{\mathbf{k}, \eta}^{(j,3)} \zeta_j s_3)}, \quad \varphi_{\mathbf{k}, \eta}^{(j,a)} = \varphi_{-\mathbf{k}, -\eta}^{(j,a)}, \tag{F75}$$

where the constraint comes from C_{2z} symmetry. Let us investigate every generator separately. First, one finds no change in energy for $\zeta_3 s_0$ and $\zeta_3 s_3$. In all other cases, the energy change is again of the form (F66). For $\zeta_{1,2} s_3$ and $\zeta_1 s_0$ we get $b_{\mathbf{k}, \eta} = 0$ and $a_{\mathbf{k}, \eta} > 0$ (for sufficiently small D_0) and, hence, no associated energetic gain. Hence, we are, again, only left with a single generator and $(U_{\mathbf{k}})_{\eta, \eta'} = \delta_{\eta, \eta'} e^{i \varphi_{\mathbf{k}, \eta} \zeta_2 s_0}$ with $\varphi_{\mathbf{k}, \eta} = \varphi_{-\mathbf{k}, -\eta}$. The prefactors are given by Eq. (F70) and we obtain the same energetic gain as for the SLP₋ (or VP and SP) state. As anticipated in Appendix F3, this mixing induces a gap for the SSLP₋ state as well, as also seen in our HF numerics, despite not explicitly ordering in the graphene-like bands (for small D_0).

5. Hund's Coupling

In the real system, $SU(2)_+ \times SU(2)_-$ is broken down to $SU(2)_s$. Keeping $U(1)_v$ as an exact symmetry, any spin-rotation invariant non-interacting Hamiltonian will automatically be invariant under $SU(2)_+ \times SU(2)_-$ and, hence, the explicit breaking of this enhanced spin symmetry can only be captured by interaction terms. We refer to all such interaction terms collectively as ‘‘intervalley Hund’s interactions’’.

In order to associate a sign of the Hund’s coupling with a member of a given Hund’s pair of states, we use the simple explicit form of the Hund’s interaction given in Eq. (3.16) of the main text. By evaluation their respective Hartree and Fock energies, we can express the contribution of this interaction to each state in Table II:

$$\begin{aligned}
\Delta E_{\text{SP}} &= -\Delta E_{\text{SVP}} = \frac{2J_H}{N} \sum_{\mathbf{G}} \left(\left[\sum_{\mathbf{k}} \bar{F}_1^{bb}(\mathbf{k}, \mathbf{G}) \right]^2 - 4 \left[\sum_{\mathbf{k}} \bar{F}_3^{bb}(\mathbf{k}, \mathbf{G}) \right]^2 D_0^4 \right) \\
\Delta E_{\text{IVC}_-} &= -3\Delta E_{\text{SIVC}_-} = -\frac{3J_H}{2N} \sum_{\mathbf{k}, \mathbf{q}} [(\bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_2^{bb}(\mathbf{k}, \mathbf{q}))^2 - 4(\bar{F}_3^{bb}(\mathbf{k}, \mathbf{q}))^2 D_0^4 - 4(\bar{F}_4^{bb}(\mathbf{k}, \mathbf{q}))^2 D_0^4 \\
&\quad + 4(\bar{F}_5^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 + 4(\bar{F}_6^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 - (\bar{F}_7^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 D_0^4 - (\bar{F}_8^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 D_0^4] \\
\Delta E_{\text{IVC}_+} &= -3\Delta E_{\text{SIVC}_+} = -\frac{3J_H}{2N} \sum_{\mathbf{k}, \mathbf{q}} [(\bar{F}_1^{bb}(\mathbf{k}, \mathbf{q}))^2 + (\bar{F}_2^{bb}(\mathbf{k}, \mathbf{q}))^2 - 4(\bar{F}_3^{bb}(\mathbf{k}, \mathbf{q}))^2 D_0^4 - 4(\bar{F}_4^{bb}(\mathbf{k}, \mathbf{q}))^2 D_0^4 \\
&\quad - 4(\bar{F}_5^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 - 4(\bar{F}_6^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 + (\bar{F}_7^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 D_0^4 + (\bar{F}_8^{bb}(\mathbf{k}, \mathbf{q}))^2 w_0^2 D_0^4] \quad (\text{F76}) \\
\Delta E_{\text{SSLP}_{\pm}} &= \mp \frac{2J_H}{N} \sum_{\mathbf{G}} \left(\left[\sum_{\mathbf{k}} \bar{F}_2^{bb}(\mathbf{k}, \mathbf{G}) \right]^2 - 4 \left[\sum_{\mathbf{k}} \bar{F}_4^{bb}(\mathbf{k}, \mathbf{G}) \right]^2 D_0^4 \right) \\
\Delta E_{\text{SLP}_{\pm}} &= 0 \\
\Delta E_{\text{VP}} &= 0
\end{aligned}$$

The form of the energies in the chiral-decoupled limit, $w_0 = D_0 = 0$, is indicated in the last two columns in Table III. Their signs can be found in the last column of Table II.

By construction of the order parameters in Table II, we see that a state and its Hund's partner have an energy contributions from J_H of opposite sign. Furthermore, all states that are their own Hund's partners, are not affected energetically to linear order in J_H .

Appendix G: Superconducting instabilities

1. Fluctuation-induced superconductivity

Finally, in this appendix we provide more details on how we determine the leading superconducting instabilities in Sec. VII B that result from fluctuations of the candidate particle-hole orders of MSTG identified in this work. Since the applied formalism is closely connected to Ref. 54, we will be brief and focus on the novel aspects relevant here—the spinless valley time-reversal symmetry $\tilde{\Theta}$ of some of the states in Table II and what happens in the absence of electron-phonon coupling—and illustrate the formalism explicitly with a minimal model for the IVC₋.

Integrating out the bosonic fields ϕ_q^j in the action \mathcal{S} in Eq. (7.1) leads to an effective electron-electron interaction. Projecting this interaction to the electronic degrees of freedom, $f_{k;\eta,s}$, with band energies $\epsilon_{\eta}(\mathbf{k})$, that are closest to the chemical potential for each momentum \mathbf{k} , we perform a decoupling in the Cooper channel (focusing on intervalley pairing which is expected to be dominant), with Hubbard-Stratonovich fields $\Delta_{s,s'}^{\eta}$, and obtain the saddle-point equations [54]

$$\Delta_{s,s'}^{\eta}(k) = - \sum_{\eta'} \int_{k'} \frac{\mathcal{V}_{\bar{s},\bar{s}'}^{s,s'}(k, \eta; k', \eta')}{\omega_{\eta'}^2 + \epsilon_{\eta'}^2(k')} \Delta_{\bar{s},\bar{s}'}^{\eta'}(k'). \quad (\text{G1})$$

Here the interaction matrix elements have the form $\mathcal{V}_{s_3, s_4}^{s_1, s_2}(k, \eta; k', \eta') = \mathcal{W}_{s_3, s_4}^{s_1, s_2}(k, \eta; k', \eta') + \Delta \mathcal{V}_{s_3, s_4}^{s_1, s_2}(k, \eta; k', \eta')$, where the first part is the $\text{SU}(2)_+ \times \text{SU}(2)_-$ symmetric contribution of the phonons from S_{phonon} and the second is a result of the particle-hole fluctuations,

$$\Delta \mathcal{V}_{s_3, s_4}^{s_1, s_2}(k, \eta; k', \eta') = -\frac{1}{2} \Lambda_{-\eta, s_2; -\eta', s_3}^j(-\mathbf{k}, -\mathbf{k}') [\chi(k - k')]_{j, j'} \Lambda_{\eta, s_1; \eta', s_4}^{j'}(\mathbf{k}, \mathbf{k}'), \quad (\text{G2})$$

where $\Lambda_{\eta, s; \eta', s'}^j(\mathbf{k}, \mathbf{k}') := \langle u_{\eta, s}(\mathbf{k}) | \lambda^j(\mathbf{k}, \mathbf{k}') | u_{\eta', s'}(\mathbf{k}') \rangle$ are the matrix elements of the coupling vertices λ^j in Eq. (7.3) with the respect to the wavefunctions $u_{\eta, s}(\mathbf{k})$ of the low-energy fermions $f_{k;\eta,s}$.

Since the system has $\text{SU}(2)_s$ symmetry, we can always decompose pairing into singlet and triplet, which correspond to $\Delta_{s,s'}(k) = (i s_y)_{s,s'} \gamma_k^{\eta}$, $\gamma_k^{\eta} = \gamma_{-k}^{-\eta}$, and $\Delta_{s,s'}(k) = (i s_y \mathbf{d} \cdot \mathbf{s})_{s,s'} \Gamma_k^{\eta}$, $\Gamma_k^{\eta} = -\Gamma_{-k}^{-\eta}$, respectively. Their respective

saddle-point equations immediately follow from Eq. (G1) and read as

$$\gamma_k^+ = \int_{k'} \frac{M_{k,k'}^s}{\omega_{n'}^2 + \epsilon_+^2(\mathbf{k}')} \gamma_{k'}^+, \quad M_{k,k'}^s = -\frac{1}{4} \sum_{s,s',\eta,\eta'} ss' \mathcal{V}_{-s',s'}^{s,-s}(\eta k, \eta; \eta' k', \eta'), \quad (\text{G3a})$$

and

$$\Gamma_k^+ = \int_{k'} \frac{M_{k,k'}^t}{\omega_{n'}^2 + \epsilon_+^2(\mathbf{k}')} \Gamma_{k'}^+, \quad M_{k,k'}^t = -\frac{1}{4} \sum_{s,s',\eta,\eta'} \eta \eta' \mathcal{V}_{-s',s'}^{s,-s}(\eta k, \eta; \eta' k', \eta'), \quad (\text{G3b})$$

respectively.

To make progress, we need to take into account time-reversal symmetry. To begin with $\tilde{\Theta}$, we have $\tilde{\Theta} |u_{\eta,s}(\mathbf{k})\rangle = \eta |u_{-\eta,-s}(-\mathbf{k})\rangle$ (technically, we are here using $\tilde{\Theta} = s_x \eta_y \mathcal{K}$ but this can equally well be used as a definition of $\tilde{\Theta}$ in Table II). For any set of bosonic modes, ϕ_q^j that are all even ($\tilde{t}_\phi = +$) or odd ($\tilde{t}_\phi = -$) under $\tilde{\Theta}$, we further have $\tilde{\Theta} \lambda^j(\mathbf{k}, \mathbf{k}') \tilde{\Theta}^\dagger = \tilde{t}_\phi \lambda^j(-\mathbf{k}, -\mathbf{k}')$ and, hence,

$$\Lambda_{\eta,s;\eta',s'}^j(\mathbf{k}, \mathbf{k}') = \eta \eta' \tilde{t}_\phi \left[\Lambda_{-\eta,-s;-\eta',-s'}^j(-\mathbf{k}, -\mathbf{k}') \right]^*. \quad (\text{G4})$$

Similarly, we have $\Lambda_{\eta,s;\eta',s'}^j(\mathbf{k}, \mathbf{k}') = \sigma \sigma' t_\phi \left[\Lambda_{-\eta,-s;-\eta',-s'}^j(-\mathbf{k}, -\mathbf{k}') \right]^*$ for bosonic modes even ($t_\phi = +$) or odd ($t_\phi = -$) under Θ_s [54, 66]. This allows us to rewrite the relevant $\text{SU}(2)_+ \times \text{SU}(2)_-$ -symmetry breaking part, $\Delta M^{s,t}$, in $M^{s,t} = M_0^{s,t} + \Delta M^{s,t}$ (with $M_0^{s,t}$ related to \mathcal{W} and phonons) as

$$(\Delta M^s)_{k,k'} = \frac{1}{8} \tilde{t}_\phi \sum_{s,s',\eta,\eta'} ss' \eta \eta' \mathcal{F}_{k,\eta,s;k',\eta',s'}, \quad (\Delta M^t)_{k,k'} = \frac{1}{8} \tilde{t}_\phi \sum_{s,s',\eta,\eta'} \mathcal{F}_{k,\eta,s;k',\eta',s'}, \quad (\text{G5})$$

when applying $\tilde{\Theta}$. We here introduced

$$\mathcal{F}_{k,\eta,s;k',\eta',s'} = \left[\Lambda_{\eta,s;\eta',s'}^j(\eta \mathbf{k}, \eta' \mathbf{k}') \right]^* [\chi(\eta k - \eta' k')]_{j,j'} \Lambda_{\eta,s;\eta',s'}^{j'}(\eta \mathbf{k}, \eta' \mathbf{k}') \geq 0, \quad (\text{G6})$$

which is non-negative, since χ has to be positive definite due to stability [54, 66]. For conventional time-reversal symmetry, Θ_s , singlet and triplet switch their role [54]

$$(\Delta M^s)_{k,k'} = \frac{1}{8} t_\phi \sum_{s,s',\eta,\eta'} \mathcal{F}_{k,\eta,s;k',\eta',s'}, \quad (\Delta M^t)_{k,k'} = \frac{1}{8} t_\phi \sum_{s,s',\eta,\eta'} ss' \eta \eta' \mathcal{F}_{k,\eta,s;k',\eta',s'}. \quad (\text{G7})$$

Upon writing $\Delta_s(k) = \gamma_k^+ / \sqrt{\omega_n^2 + \epsilon_+^2(\mathbf{k})}$ and $\Delta_t(k) = \Gamma_k^+ / \sqrt{\omega_n^2 + \epsilon_+^2(\mathbf{k})}$, we can recast solving the saddle-point equations in Eq. (G3) as solving for the largest eigenvalue $\lambda_{t,s}(T)$ of the symmetric matrix $(\tilde{M}^{s,t})_{k,k'} := M_{k,k'}^{t,s} / \sqrt{(\omega_n^2 + \epsilon_+^2(\mathbf{k}))(\omega_{n'}^2 + \epsilon_+^2(\mathbf{k}'))}$ as a function of temperature T . The critical temperature, $T_c^{s,t}$, in the two channels corresponds to $\lambda_{t,s}(T_c^{s,t}) = 1$. Defining $(\Delta \tilde{M}^{s,t})_{k,k'} := \Delta M_{k,k'}^{t,s} / \sqrt{(\omega_n^2 + \epsilon_+^2(\mathbf{k}))(\omega_{n'}^2 + \epsilon_+^2(\mathbf{k}'))}$, we immediately see from Eqs. (G5) and (G7), that $(\Delta \tilde{M}^s)_{k,k'} \geq (\Delta \tilde{M}^t)_{k,k'}$ if $t_\phi = +$ or $\tilde{t}_\phi = -$ and $(\Delta \tilde{M}^t)_{k,k'} \geq (\Delta \tilde{M}^s)_{k,k'}$ if $t_\phi = -$ or $\tilde{t}_\phi = +$.

Let us first discuss the situation without electron-phonon coupling and, hence, $\tilde{M}^{s,t} = \Delta \tilde{M}^{s,t}$. As follows from the Perron-Frobenius theorem, given two square matrices A and B of equal size with $A_{ij} \geq |B_{ij}|$ and A being irreducible [73], the largest eigenvalue of A (the ‘‘Perron root’’) is larger than the magnitude of the largest eigenvalue of B . We then have $\lambda_s(T) \geq \lambda_t(T)$ and, hence, $T_c^s \geq T_c^t$ if $t_\phi = +$; this means that singlet will generically dominate over triplet. Similarly, if $\tilde{t}_\phi = +$, we get $\lambda_t(T) \geq \lambda_s(T)$ leading to $T_c^t \geq T_c^s$, which means that triplet will dominate. Note that if the fluctuations have a well-defined behavior under both Θ_s and $\tilde{\Theta}$ and we further have $t_\phi = \tilde{t}_\phi$, Eqs. (G5) and (G7) imply that $\Delta M^s = \Delta M^t$, i.e., singlet and triplet remain degenerate.

In the scenario proposed in Ref. 54, where electron-phonon coupling yields an important contribution in stabilizing superconductivity but predominantly in the $\text{SU}(2)_+ \times \text{SU}(2)_-$ -symmetric channel, we treat the symmetry-breaking $\Delta \tilde{M}^{s,t}$ as a small perturbation to the $\text{SU}(2)_+ \times \text{SU}(2)_-$ -symmetric $\tilde{M}_0^{s,t} := \tilde{M}^{s,t} - \Delta \tilde{M}^{s,t}$ associated with the phonons. As follows from the analysis of Ref. 66, we must have $\Delta_{t,s}^0(k) > 0$ for the unperturbed

saddle-point solutions, i.e., leading eigenstates of $\widetilde{M}_0^{s,t}$. As such, it is clear that the leading (first-order) correction $\delta\lambda_{s,t} = \int_{\mathbf{k},\mathbf{k}'} \Delta_{t,s}^0(\mathbf{k})(\Delta\widetilde{M}^{s,t})_{\mathbf{k},\mathbf{k}'}\Delta_{t,s}^0(\mathbf{k}')$ is larger for singlet (triplet), which then dominates, if $t_\phi = +$ ($t_\phi = -$) or $\tilde{t}_\phi = -$ ($\tilde{t}_\phi = +$).

We illustrate these statements more explicitly for the IVC₋ state, which is odd under Θ_s and even under $\tilde{\Theta}$, i.e., characterized by $t_\phi = -$ and $\tilde{t}_\phi = +$. As already mentioned in Sec. VII B, a minimal description is given by $[\chi(i\Omega_n, \mathbf{q})]_{j,j'} = \chi_0\delta_{j,j'}/(\Omega_n^2 + c^2\mathbf{q}^2 + \xi^{-2})$, $j = 1, 2$, and $\lambda^j(\mathbf{k}, \mathbf{k}') = s_0\eta_j f_{\mathbf{k},\mathbf{k}'}$, where $f_{\mathbf{k},\mathbf{k}'}$ is a matrix in band space with $f_{\mathbf{k},\mathbf{k}'} = f_{\mathbf{k}',\mathbf{k}}^\dagger = -f_{\mathbf{k}',\mathbf{k}}^T$. This leads to

$$\mathcal{F}_{k,\eta,s;k',\eta',s'} = \frac{2\chi_0\delta_{s,s'}\delta_{\eta,-\eta'}}{\Omega_n^2 + c^2\mathbf{q}^2 + \xi^{-2}} |\langle u_{\eta,s}(\mathbf{k})|f_{\mathbf{k},\mathbf{k}'}|u_{\eta,s}(\mathbf{k}')\rangle|^2 \geq 0. \quad (\text{G8})$$

So only matrix elements with $ss'\eta\eta' = -1$ are non-zero such that $\Delta M^{s,t}$ are identical in Eqs. (G5) and (G7) for $t_\phi = -$ and $\tilde{t}_\phi = +$. We, thus, consistently see that IVC₋ fluctuations will favor the triplet state.

Finally, consider the SLP₋ state, which is odd under both Θ_s and $\tilde{\Theta}$ and, thus, has $t_\phi = \tilde{t}_\phi = -$. To be consistent with the analysis above, fluctuations of this state must affect singlet and triplet in exactly the same way. This can also be easily seen explicitly by noting that its coupling vertex λ in Eq. (7.3) cannot couple different valleys or spin species and, hence, $\mathcal{F}_{k,\eta,s;k',\eta',s'} \propto \delta_{s,s'}\delta_{\eta,\eta'}$. Consequently, we get $\Delta M^s = \Delta M^t$ in Eqs. (G5) and (G7).

2. Order parameter in the polarized phase

Finally, we also illustrate the spin-structure of the superconducting order parameter in the presence of flavor polarization, relevant for ν with $2 \lesssim |\nu| \lesssim 3$. For both SP-based and SVP-based band resetting, the superconducting order parameter in the TBG-like bands has the form

$$H_{\text{SC,p}}^b = \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \psi_{\mathbf{k};+}^\dagger \psi_{-\mathbf{k};-}^\dagger + \text{H.c.}, \quad (\text{G9})$$

where $\psi_{\mathbf{k};\pm}$ are the electronic operators of valley $\eta = \pm$ in the polarized TBG-like bands which are closest to the Fermi level for each \mathbf{k} . In principle, there are also finite band-off-diagonal matrix elements. However, since we are only interested in the predominant spin-structure of the superconductor, we do not need to take them into account. Note that, if the band resetting also involves intervalley coherent order, such as in Fig. 7(c,d), the labels \pm of $\psi_{\mathbf{k};\pm}$ have to be understood as combinations of the two valleys; the spin structure we discuss here will not be affected by it.

To transform back to the physical spin ($s = \uparrow, \downarrow$) basis, we write $\psi_{\mathbf{k};\eta} = \sum_s (u_{\mathbf{k};\eta}^*)_s f_{\mathbf{k};\eta,s}$, where the two-component spinor, $u_{\mathbf{k};\eta}$, encodes the nature of the flavor polarization. The singlet, $\Delta_{\mathbf{k}}^s$, and triplet, $\mathbf{d}_{\mathbf{k}}$, component of the order parameter, coupling as $H_{\text{SC,p}}^b = \sum_{\mathbf{k}} \Delta_{\mathbf{k}} f_{\mathbf{k};+,s}^\dagger ((s_0\Delta_{\mathbf{k}}^s + \mathbf{s} \cdot \mathbf{d}_{\mathbf{k}})i s_y)_{s,s'} f_{-\mathbf{k};-,s'}^\dagger + \text{H.c.}$, then follow from

$$((s_0\Delta_{\mathbf{k}}^s + \mathbf{s} \cdot \mathbf{d}_{\mathbf{k}})i s_y)_{s,s'} = (u_{\mathbf{k};+})_s (u_{-\mathbf{k};-})_{s'}. \quad (\text{G10})$$

For the case of SP, where the spin polarization is in the same direction in the two valleys, say along the s_z direction, we have

$$(u_{\mathbf{k};\pm})_s = \delta_{s,\uparrow}, \quad \Rightarrow \quad \Delta_{\mathbf{k}}^s = 0, \quad \mathbf{d}_{\mathbf{k}} = -\Delta_{\mathbf{k}}(1, i, 0)^T/2, \quad (\text{G11})$$

corresponding to non-unitary triplet pairing. A Zeemann field, \mathbf{B}_Z , will not affect the polarization nor the superconductor (except for pinning the direction of their magnetizations).

For the SVP-type of polarization, we have instead (choosing the polarization along the s_x axis)

$$u_{\mathbf{k};\pm} = \frac{1}{\sqrt{2}} \begin{pmatrix} \pm 1 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \Delta_{\mathbf{k}}^s = \Delta_{\mathbf{k}}/2, \quad \mathbf{d}_{\mathbf{k}} = \Delta_{\mathbf{k}}(1, 0, 0)^T/2, \quad (\text{G12})$$

leading to an equal admixture of singlet and triplet pairing [27]. Applying \mathbf{B}_Z , will pin the orientation of the spin to be orthogonal to it and induce a canting of angle θ that increases with $|\mathbf{B}_Z|$. For \mathbf{B}_Z along the s_z direction, we thus get

$$u_{\mathbf{k};\pm} = \frac{1}{\sqrt{2(1 + \sin\theta)}} \begin{pmatrix} \pm(1 + \sin\theta) \\ \cos\theta \end{pmatrix} \quad \Rightarrow \quad \Delta_{\mathbf{k}}^s = \cos\theta\Delta_{\mathbf{k}}/2, \quad \mathbf{d}_{\mathbf{k}} = \Delta_{\mathbf{k}}(1, i\sin\theta, 0)^T/2. \quad (\text{G13})$$

TABLE VI: We explicitly list the notation we use to denote the Pauli matrices which act on the sublattice, minivalley, valley, and spin degrees of freedom used throughout the text in the second column, and the notation used in [25] in the third column.

	Notation in this work	Notation in [25]
Sublattice	ρ	ρ
Minivalley	μ	μ
Valley	η	τ
Spin	s	σ

So we see that this state continuously transforms into the non-unitary triplet when applying a magnetic field. We note the behavior in magnetic field of these pairing states—the non-unitary triplet and singlet-triplet admixed phases—is the same in the absence of polarization, as shown in Ref. 27.

Appendix H: WZW terms in MSTG

In this appendix we will give more details on how we reached the conclusions in Sec. VII C. We will begin as in the main text with $|\nu| = 0$. In this case, as in the main text, the low-energy physics is captured by

$$H_0^D = \sum_{\mathbf{q}} \psi_{\mathbf{q}}^\dagger [\gamma_1 q_x + \gamma_2 q_y] \psi_{\mathbf{q}}, \quad \gamma_1 = \rho_x \mu_z s_0, \quad \gamma_2 = \rho_y \mu_0 s_0, \quad (\text{H1})$$

where we are using the notation for Pauli matrices defined in Table VI. We will assume a pairing Δ with 2 real components and an insulating order parameter m_j with three real components, incorporated as a Dirac masses,

$$H_1^D = \sum_{\mathbf{q}, j} \psi_{\mathbf{q}}^\dagger m_j \psi_{\mathbf{q}} + \sum_{\mathbf{q}} \left(\psi_{\mathbf{q}}^\dagger \Delta T \psi_{-\mathbf{q}}^\dagger + \text{H.c.} \right) \quad (\text{H2})$$

with $T = \mu_x s_y$. We want to consider all possible pairings Δ and insulating states m_j such that when the Dirac fermions are integrated out of the action corresponding to $H_0^D + H_1^D$, a WZW term will be produced. One way to do this would be to consider which of the options classified in [25] survive projection to a minivalley-valley locked subspace. Alternatively, we can consider which orders are compatible with the conditions derived in [25] for the projected form of the Dirac action. We will choose the second way in this appendix. We will first consider the compatible pairings Δ at $\nu = 0$. At $\nu = 0$, triplet pairing is excluded as we require $SU(2)_s$ spin symmetry to be intact before the bands acquire any ordering. A requirement for a pairing to be compatible with the existence of a WZW term is that it must satisfy the following conditions with the kinetic term's γ matrices

$$\gamma_i \Delta T = -\Delta T \gamma_i^T \neq 0 \quad i = 1, 2. \quad (\text{H3})$$

We will additionally assume pairings which do not break moiré translational symmetry and intervalley pairing terms which preserve $U(1)_\nu$. We have also chosen a momentum independent Δ . The only pairing compatible with all the above requirements is $\Delta = \mathbb{1}$, transforming under the A representation of the point group. As mentioned in the main text, this is the same pairing considered for TBG in [25] projected to the minivalley-valley locked subspace. For this pairing, we can then ask what are the compatible 3-component order parameters m_j , $j = 1, 2, 3$. The conditions the insulating states must satisfy with the pairing and γ matrices are:

$$\text{Tr}[\gamma_{i_1} \gamma_{i_2} m_{j_1} m_{j_2} m_{j_3}] \propto \epsilon_{i_1 i_2 i_3 j_1 j_2 j_3} \quad (\text{H4})$$

$$m_j \Delta T = \Delta T m_j^T \neq 0 \quad j = 1, 2, 3 \quad (\text{H5})$$

We find the only two options are a 3-component spin-Hall state $\rho_z \mu_z (s_x, s_y, s_z)$ and a 2-component moiré density wave state $\rho_x (\mu_x, \mu_y)$ combined with a third component sublattice polarized state ρ_z .

We can then consider the case at $|\nu| = 2$. As argued in the main text, the low-energy Dirac theory takes the same form as at $\nu = 0$, see Eq. (H1).

However, unlike the case at $\nu = 0$, we no longer assume $SU(2)_s$ spin symmetry. It is simple to see the options which were compatible at $\nu = 0$ for singlet pairing are still compatible at $\nu = 2$. On the other hand, the breaking of

spin rotation symmetry allows for additional options at $|\nu| = 2$ which were not previously possible at $|\nu| = 0$. It is easily verified that the triplet pairing $\Delta = \mu_z s_z$ satisfies Eq. (H3). We find two options which are compatible with this pairing. The first is the set of orders $\rho_z(s_x, s_y)$ and $\rho_z \mu_z s_z$, corresponding to our SSLP₊ and SSLP₋ respectively. The second is $\rho_x(\mu_x, \mu_y)$ and ρ_z , corresponding to a moiré density wave state and SLP₊ state respectively.