Strong coupling phases of partially filled twisted bilayer graphene narrow bands

Jian Kang$^{1}$ and Oskar Vafek$^{1,2}$

$^1$National High Magnetic Field Laboratory, Tallahassee, Florida, 32304, USA
$^2$Department of Physics, Florida State University, Tallahassee, Florida 32306, USA

We identify states favored by Coulomb interactions projected onto the Wannier basis of the four narrow bands of the “magic angle” twisted bilayer graphene. At the filling of two electrons/holes per moire unit cell, such interactions favor an insulating SU(4) ferromagnet. The kinetic terms select the ground state in which the two valleys with opposite spins are equally mixed, with vanishing magnetic moment per particle. We also find extended excited states, the gap to which decreases in magnetic field. An insulating stripe ferromagnetic phase is favored at one electron/hole per unit cell.

In addition to superconductivity, recent experiments on magic angle twisted bilayer graphene revealed insulating phases at carrier concentrations corresponding to partial occupation of the four narrow bands composite near the neutrality point [1,3]. Such correlated insulator phases seem to occur only when the bandwidth of the composite is reduced either by fine-tuning of the twist angle to the vicinity of the “magic” value $\sim 1.1^\circ$ or by tuning the applied pressure at $\sim 1.3^\circ$ [1,3]. Importantly, the insulating states occur at commensurate (rational) fillings corresponding to 2 electrons/holes per moire unit cell, with additional resistance peaks observed at fillings of 1 hole/electron per unit cell and 3 holes/electrons per unit cell [1,2]. This observation is hard to reconcile with the notion that the insulation is due to Fermi surface nesting, or the van Hove singularities, reconstructed by electron-electron interactions, because such band structure features generically occur at incommensurate fillings. Instead, the above observations suggest that the effective Coulomb interaction dominates the effective kinetic energy [1,3]. The former is given by the projection of the Coulomb interaction onto the Hilbert space spanned by the narrow bands and is $\sim \epsilon^2/\ell_m \sim 15meV$, where the moire period $\ell_m \sim 13nm$ and $\epsilon \approx 6$ is the dielectric constant of the encapsulating BN. The kinetic energy scale is given by the bandwidth. Although there is no direct measurement of the bandwidth, theoretical calculations routinely find it to be $\lesssim 10meV$ [1,4,9].

Such considerations hint that, even if the physical system is ultimately in an intermediate coupling regime, a strong coupling approach may be more successful in capturing the nature of the correlated phases. In this approach the interaction-only Hamiltonian is minimized first, and the kinetic energy term is then treated as a perturbation [7,10,19].

Here we present the analysis and the solution to the strong coupling limit by projecting the Coulomb interaction onto the microscopically constructed exponentially localized Wannier states (WSs) for the four narrow bands [2,9]. In doing so we find that there is a qualitative difference between the effect of the interactions in twisted bilayer graphene narrow bands and the much studied narrow band whose width is small due to the exponentially vanishing overlap of the well separated localized orbitals i.e. a solid in an atomic limit. In contrast, the small bandwidth in twisted bilayer graphene is a result of fine tuning (twist angle or pressure) and subtle interference of the WSs, and, unlike in the atomic limit, it is not necessarily a result of large spatial separation of the exponentially localized WSs. Indeed, as shown before, each WS of the twisted bilayer graphene narrow bands has three main peaks on neighboring sites of the triangular moire superlattice [5,7]. Therefore, for nearest neighbor WSs on say, sites $i$ and $j$, two peaks overlap significantly (see Fig.1). Even though the integral under both has to vanish by orthogonality, the integral under each separately does not. This leads to a dramatically new form of the interaction Hamiltonian projected onto the narrow band basis – containing terms beyond the “cluster Hubbard” term [7,11] – which in turn leads to different strong coupling phases as in the atomic limit. Specifically, the usual anti-ferromagnetic super-exchange mechanism fails and turns ferromagnetic. Due to approximate spin-valley SU(4) symmetry, the fully spin-valley polarized ferromagnet is found to be degenerate with a spin-valley entangled state whose average total magnetic moment per particle vanishes. We also find exact excited states, which are spatially extended, and whose gap is suppressed by Zeeman coupling to an external magnetic field, making it (or at least its order parameter) a candidate for the experimentally observed correlated insulator at 2 electrons/holes per moire unit cell. At 1 particle per moire unit cell we find that the projected interactions favor an insulating stripe SU(4) ferromagnet. This state may be a candidate for the insulator observed at the 1/8 filling [3] if the SU(4) degeneracy is lifted in favor of the physical spin ferromagnet.

We start by writing the full Hamiltonian as

$$H = K + U,$$

where the kinetic energy $K$ is described by the tight-binding model [9] based on the WSs and where the Coulomb interaction is

$$U = \frac{1}{2} \sum_{\mathbf{r}, \mathbf{r}' \sigma, \sigma'} c^\dagger_{\sigma}(\mathbf{r}) c_{\sigma}(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') c^\dagger_{\sigma'}(\mathbf{r}') c_{\sigma'}(\mathbf{r}').$$
Figure 1. The centers of the hexagons correspond to the triangular moire lattice spanned by primitive vectors \( \mathbf{L}_{1,2} \). The Wannier state (WS) wavefunction centered on the moire honeycomb site \( j \) has three peaks at the neighboring triangular moire sites (grey circles with vertical stripes). The WS on the neighboring site \( i \) overlaps with it on the two hexagons (red horizontal stripes). An example of a four fermion interaction term, which is beyond the extended Hubbard model, appearing in the strong coupling Hamiltonian Eqs. (6-9-11), is also shown schematically.

Projected onto the four narrow bands is equivalent to expanding \( c_{\sigma}(\mathbf{r}) \) solely in terms of the narrow bands WSs

\[
c_{\sigma}(\mathbf{r}) = \frac{1}{3} \sum_{\mathbf{R}} \sum_{p=1}^{6} \sum_{j=\pm 1} w_{\mathbf{R} + \delta_p, j}(\mathbf{r})d_{j,\sigma}(\mathbf{R} + \delta_p),
\]

where integers \( m, n \) define the triangular moire lattice vectors \( \mathbf{R} = m\mathbf{L}_1 + n\mathbf{L}_2 \), the eigenvalue of the AA site centered 3-fold rotation \( \exp(j2\pi i/3) \) is labeled by \( j = \pm 1 \) and \( \delta_{1,\ldots,6} \) are basis vectors connecting the honeycomb sites to the triangular sites (see Fig. 1). To an excellent approximation, WSs with \( j = \pm 1 \) correspond to different valleys with very little valley mixing. The factor of 1/3 is due to each honeycomb site position \( \mathbf{R} + \delta_p \) being counted three times.

The Coulomb interaction \( V(\mathbf{r}) \) is screened due to the presence of the metallic gates \([1-3]\). The separation between the gates sets the length-scale beyond which the image charges exponentially diminish the repulsion \([20]\). Interestingly, the gate separation is comparable to the moire unit cell. This, as well as the form of \( w_{\mathbf{R} + \delta_p, j}(\mathbf{r}) \) justifies keeping only \( \mathbf{R} = \mathbf{R}' \) in the sum below:

\[
U = \frac{1}{2} \sum_{\mathbf{R}, \mathbf{R}' \neq \mathbf{R}} \sum_{\mathbf{r}, \mathbf{r}' \in \mathcal{O}_{\sigma,\sigma'}=\uparrow,\downarrow} n_{\sigma}(\mathbf{R} + \mathbf{r})V(\mathbf{R} + \mathbf{r} - \mathbf{R}' - \mathbf{r}')n_{\sigma'}(\mathbf{R}' + \mathbf{r}')
\]

\[
\approx \frac{1}{2} \sum_{\mathbf{R}} \sum_{\mathbf{r}, \mathbf{r}' \in \mathcal{O}_{\sigma,\sigma'}=\uparrow,\downarrow} n_{\sigma}(\mathbf{R} + \mathbf{r})V(\mathbf{r} - \mathbf{r}')n_{\sigma'}(\mathbf{R}' + \mathbf{r}'),
\]

where \( n_{\sigma}(\mathbf{r}) = c_{\sigma}^\dagger(\mathbf{r})c_{\sigma}(\mathbf{r}) \) and the sums over \( \mathbf{r}, \mathbf{r}' \) are restricted to be within the moire hexagon centered at the origin (see shaded \( \mathcal{O} \) in Fig. 1).

Substituting the Eq. (3) into the above form, with numerically calculated \( w_{\mathbf{R} + \delta_p, j}(\mathbf{r}) \) from the microscopic model \([6]\) we find that to an excellent approximation we can replace \( V(\mathbf{r} - \mathbf{r}') \) by its average over a region of size set by the extent of \( w_{\delta_p, j}(\mathbf{r}) \) within the moire hexagon \( V_0 \), and because \( V(\mathbf{r}) \) is dominated by the small wavevectors, we can ignore the valley mixing terms \([5]\). Thus,

\[
U \approx \frac{V_0}{2} \sum_{\mathbf{R}} \left( \sum_{j=\pm 1} \sum_{\sigma,\sigma'=\uparrow,\downarrow} O_{j,\sigma}(\mathbf{R}) \right)^2
\]

where \( O_{j,\sigma}(\mathbf{R}) = \sum_{\mathbf{r} \in \mathcal{O}} n_{j,\sigma}(\mathbf{R} + \mathbf{r}) \) and

\[
n_{j,\sigma}(\mathbf{R} + \mathbf{r}) = \frac{1}{9} \sum_{\mathbf{R}, \mathbf{R}' \neq \mathbf{R}} \sum_{p,p'=1}^6 w_{\mathbf{R} + \delta_p, j}(\mathbf{r})w_{\mathbf{R}' + \delta_p, j}(\mathbf{r})d_{j,\sigma}^\dagger(\mathbf{R} + \delta_p)d_{j,\sigma}(\mathbf{R}' + \delta_{p'})
\]

\[
\approx \frac{6}{p,p'=1} w_{\delta_p, j}(\mathbf{r})w_{\delta_p, j}(\mathbf{r})d_{j,\sigma}^\dagger(\mathbf{R} + \delta_p)d_{j,\sigma}(\mathbf{R} + \delta_{p'}) .
\]

It is clear that \( O_{j,\sigma}(\mathbf{R}) \) is a superposition of not only density-like operators \( d_{j,\sigma}^\dagger(\mathbf{R} + \delta_p)d_{j,\sigma}(\mathbf{R} + \delta_{p'}) \) with \( p = p' \), but also hopping-like terms with \( p \neq p' \) which may be of the same order of magnitude. For example,
such as those with $p = 1$ and $p’ = 2$ to be negative of the terms with $p = 5$ and $p’ = 4$, etc. In what follows, we assume for clarity that the 3 peaks of each WS reside entirely within the 3 neighboring hexagons with no support elsewhere. We relax this assumption in the Supplementary material without any change to our conclusions [21].

To summarize,

$$O_{j,\sigma}(R) = \frac{1}{3} Q_{j,\sigma}(R) + \alpha_1 T_{j,\sigma}(R),$$

where

$$Q_{j,\sigma}(R) = \sum_{p=1}^{6} d_{j,\sigma}^{\dag}(R + \delta_p) d_{j,\sigma}(R + \delta_p),$$

$$T_{j,\sigma}(R) = \sum_{p=1}^{6} \left( e^{i\eta_p} d_{j,\sigma}^{\dag}(R + \delta_{p+1}) d_{j,\sigma}(R + \delta_p) + h.c. \right),$$

where $e^{i\eta_p} = (-)^{p-1} e^{i(-)^{p-1} \theta_1}$, $\delta_7 = \delta_1$, and

$$\alpha_1 e^{i\theta_1} = \sum_{r \in \mathbb{C}} w_{R+\delta_{j,\sigma}}(r) w_{R+\delta_{j,\sigma}}(r).$$

(12)

$\alpha_1 e^{i\theta_1}$ is generally a complex number and $\theta_{+1} = -\theta_{-1}$. This phase factor can be absorbed by applying a global $U(1)$ transformation on WSs. In the rest of the paper, we will therefore assume $\theta_{+1} = -\theta_{-1} = 0$. For our WSs constructed from the projection method [21], $\alpha_1 \approx 0.23$. Although not all the above interaction terms have been included in the model of Ref. [5], and although the Coulomb interaction is not assumed screened in Ref. [5], similar value for $\alpha_1$ can be estimated from their ratio of the nearest-neighbor exchange and nearest neighbor density repulsion as $\alpha_1^{(K)} \approx \frac{1}{4} \sqrt{J_1/V_1} \approx 0.16$ (see Table I of Ref. [5]). The nature of the ground state in the strong coupling limit is insensitive to such differences.

We emphasize that it is not necessary to include the kinetic energy terms $K$ in Eqn. (1) to induce correlation among various sites; such sizable value of $\alpha_1$ makes the projected interaction term (non-local even in the strong coupling limit, and as we will see it dictates the nature of the ground state. It is therefore worth understanding why $\alpha_1$ is sizable. In the atomic limit, this overlap is exponentially small. As a consequence, the interactions usually include only the on-site terms, giving rise to the Hubbard model; $\alpha_1$ would then be set by the ratio of the bandwidth and the on-site repulsion. In our case, as mentioned, the two of the three peaks of the neighboring WSs spatially overlap and $\alpha_1 \sim \mathcal{O}(1)$. This stems from the fact that the emergent two-fold symmetry $C_2^\prime$ (see Fig 1) is not locally implemented for our valley filtered WSs [22]. Otherwise, when combined with (locally implemented) $C_2''$ (see Fig 1) and the emergent valley $U(1)$ symmetry, all the WSs would have to have the same parity under $C_2''$ [22], leading to $\alpha_1 = 0$. However, $C_2''$ cannot be locally implemented simultaneously with the valley $U(1)$, $C_2’$, and the time reversal symmetry [7] [22]. $\alpha_1 \sim \mathcal{O}(1)$ is thus rooted in the non-trivial topological properties of the narrow bands [7] [22] [26].

As the first step, we therefore need to find the spectrum of the interaction $U$ in Eqn. (6). This is non-trivial because the commutator $[O_{\sigma,j}(R), O_{\sigma,j}(R')]$ does not vanish for nearest neighbors $R$ and $R'$ due to $\alpha_1 \neq 0$. However, the ground state of (5) can be exactly solved for special fillings, including 2 particles/holes per unit cell. To see this, note that $\sum_{R} \sum_{j} \sum_{\sigma} O_{j,\sigma}(R) = N$, where $N$ is the total particle number operator. Therefore, we can write (6) exactly as

$$\frac{V_0}{2} \sum_{R} \left( n_0 - \sum_{j,\sigma} O_{j,\sigma}(R) \right)^2 + \frac{V_0}{2} n_0^2 N$$

(13)

where $N_R$ is the total number of moire unit cells. Because $N$ is fixed in the quantum number sector of interest, the last two terms are fixed. The ground state thus minimizes the first term. But the first term is a sum of squares of Hermitian operators, and if we can find a state in which each term vanishes, we find the ground state. Let $n_0 = 2$. Then the state

$$|\Phi_0\rangle = \prod_{R} d^{\dag}_{j=1,\uparrow}(R + \delta_1) d_{j=1,\uparrow}(R + \delta_2) \prod_{R} 0$$

(14)

makes the first term vanish for every $R$, and is therefore a ground state. This state corresponds to a fully spin/valley polarized ferromagnet with two electrons per moire unit cell. Although it is a ground state, it is not the only one. Due to the $SU(4)$ symmetry of Eq. (6), the ground state is $(2N_R + 3)(2N_R + 2)(2N_R + 1)/6$ fold degenerate. This $SU(4)$ ground state manifold includes states as (see Fig 2(a))

$$|\Phi_1\rangle = \prod_{R} d_{j=1,\downarrow}(R + \delta_1) d^{\dag}_{j=1,\downarrow}(R + \delta_2) \prod_{R} 0$$

(15)

Note that the expectation value of the square of the total (magnetic) spin operator $|\Phi_1\rangle S_{tot}^{\mu} |\Phi_1\rangle = \mathcal{O}(N_R)$, which means that the magnetic moment per particle vanishes in the thermodynamic limit. $|\Phi_1\rangle$ is therefore not a ferromagnet.

The ground state degeneracy is lifted by the kinetic terms, $K$ in Eqn. (1), which in general break the $SU(4)$ symmetry. The valley $U(1)$ symmetric hopping terms $t(R + \delta, R' + \delta') d^{\dag}_{R+\delta,j,\sigma} d_{R'+\delta,j,\sigma}$ favor the state with two valleys equally mixed, because then the second order process is least blocked. For the same reason the hopping terms that mix the valleys favor the state in which the two valleys carry opposite spins. The ground states is then given by Eqn. (15) up to a global spin $SU(2)$ rotation. The non-magnetic ground state depicted in Fig 2(a) is thus favored by the kinetic terms.

We can also find some of the excited eigenstates of
Eq. (6) exactly. In particular, the “cluster Hubbard” terms, with ground state at this filling analytically, even in the strong coupling limit because $\alpha_1 \neq 0$. However, the ground state can be found if $\alpha_1$ is small. The leading term in $U$ is given by the “cluster Hubbard” terms, with ground states for which each hexagon contains three fermions, and $\sum_{j,\sigma} Q_{j,\sigma}(R) = 3$. Such ground states are highly degenerate even without counting the valley and spin degrees of freedom. The linear order and the second order of the cross term $\sum_{R} \left( \sum_{j,\sigma} Q_{j,\sigma}(R) \right) \left( \sum_{j',\sigma} T_{j',\sigma}(R) \right)$ vanish for the same reason as discussed above. Therefore, to the order $O(\alpha_1^2)$, only the term $\sum_{R} \left( \sum_{j,\sigma} T_{j,\sigma}(R) \right)^2$ contributes. This contribution is minimized if (1) each hexagon contains exactly three occupied sites; (2) each occupied site is in the same state; (3) the number of bonds connecting an occupied site and an unoccupied site is minimized. These constraints favor the stripe $SU(4)$ ferromagnetic phase as the ground state, see Fig. 2(b) with the energy correction $\delta E = \alpha_1^2 N R V_0/2$. This phase is also an insulator due to the existence of the charge gap.

To summarize, we analysed the Coulomb interactions (screened by the gates) projected to the exponentially localized Wannier states for the four narrow bands in the “magic angle” twisted bilayer graphene. The projected interaction is highly non-local and is beyond extended Hubbard models. Such novel interactions result from the non-trivial topological properties of the narrow bands, giving rise to the $SU(4)$ ferromagnetic ground states at 1/4 and 1/8 fillings. At 1/4 filling, the kinetic terms break the $SU(4)$ symmetry and select the state in which two valleys with opposite spins are equally mixed (Fig. 2(a)). This state, although still $SU(4)$ ferromagnetic, is (physical) spin non-magnetic in the thermodynamic limit, with a charge gap suppressed by the magnetic field. We also argue that the stripe $SU(4)$ ferromagnetic insulator phase is the ground state at 1/8 filling (Fig. 2(b)). If the $SU(4)$ degeneracy is lifted in favor of the physical spin ferromagnet, such state could be a candidate for the experimentally observed insulator at the 1/8 filling. The mechanism of such symmetry breaking remains an open problem.

JK was supported by the National High Magnetic Field Laboratory through NSF Grant No. DMR-1157490 and the State of Florida. O. V. was supported by NSF DMR-1506756.

---

Figure 2. Schematic of the ground states at (a) 1/4 filling (2 electrons/holes per moire unit cell) and at (b) 1/8 filling (1 electron/hole per moire unit cell).
[5] M. Koshino, N.F.Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Maximally Localized Wannier Orbitals and the Extended Hubbard Model for Twisted Bilayer Graphene, Phys. Rev. X 8, 031087 (2018).

[6] J. Kang and O. Vafek, Symmetry, Maximally Localized Wannier States, and a Low-Energy Model for Twisted Bilayer Graphene Narrow Bands, Phys. Rev. X 8, 031088 (2018).

[7] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Origin of Mott Insulating Behavior and Superconductivity in Twisted Bilayer Graphene, Phys. Rev. X 8, 031089 (2018).

[8] X. Lin and D. Tomanek, Minimum model for the electronic structure of twisted bilayer graphene and related structures, Phys. Rev. B 98, 081410(R) (2018).

[9] M. Angeli, D. Mandelli, A. Valli, A. Amaricci, M. Capone, E. Tosatti, M. Fabrizio, Emergent D6 symmetry in fully-relaxed magic-angle twisted bilayer graphene, arXiv:1809.11140.

[10] C. Xu and L. Balents, Topological Superconductivity in Twisted Multilayer Graphene, Phys. Rev. Lett. 121, 087001 (2018).

[11] X. Y. Xu, K. T. Law, and P. A. Lee, Kekulé valence bond order in an extended Hubbard model on the honeycomb lattice, with possible applications to twisted bilayer graphene, Phys. Rev. B 98, 121406 (2018).

[12] B. Padhi, C. Setty, and P. W. Phillips, Doped Twisted Bilayer Graphene near Magic Angles: Proximity to Wigner Crystallization not Mott Insulation, Nano Lett. 18, 6175 (2018).

[13] J. F. Dodaro, S. A. Kivelson, Y. Schattner, X.-Q. Sun, and C. Wang, Phases of a phenomenological model of twisted bilayer graphene, Phys. Rev. B 98, 075154 (2018).

[14] A. Thomson, S. Chatterjee, S. Sachdev, M. S. Scheurer, Triangular antiferromagnetism on the honeycomb lattice of twisted bilayer graphene, Phys. Rev. B 98, 075109 (2018).

[15] D. M. Kennes, J. Lischner, C. Karrasch, Strong Correlations and d+id Superconductivity in Twisted Bilayer Graphene, arXiv:1805.06310.

[16] C.-M. Jian, C. Xu, More Insulators viewed as the Surface of three dimensional Symmetry Protected Topological Phases, arXiv:1810.03610.

[17] B. Padhi, P. Phillips, Pressure-Induced Metal-Insulator Transition in Twisted Bi-layer Graphene, arXiv:1810.00884.

[18] M. Ochi, M. Koshino, K. Kuroki, Possible correlated insulating states in magic-angle twisted bilayer graphene under strongly competing interactions, Phys. Rev. B 98, 081102 (2018).

[19] J. W. F. Venderbos, R. M. Fernandes, Correlations and electronic order in a two-orbital honeycomb lattice model for twisted bilayer graphene, arXiv:1808.10410.

[20] R. E. Throckmorton and O. Vafek, Fermions on bilayer graphene: Symmetry breaking for B = 0 and ν = 0, Phys. Rev. B 86, 115447 (2015).

[21] See Supplementary Material for more detailed calculations of the interaction constants and the derivation of the ground states as well as the extended excited states.

[22] L. Zou, H. C. Po, A. Vishwanath, T. Senthil, Band Structure of Twisted Bilayer Graphene: Emergent Symmetries, Commensurate Approximants and Wannier Obstructions, Phys. Rev. B 98, 085435 (2018).

[23] Zhida Song, Zhijun Wang, Wujun Shi, Gang Li, Chen Fang, B. Andrei Bernevig, All “Magic Angles” Are “Stable” Topological, arXiv:1807.10676.

[24] H. C. Po, L. Zou, T. Senthil, A. Vishwanath, Faithful Tight-binding Models and Fragile Topology of Magic-angle Bilayer Graphene, arXiv:1808.02482.

[25] J. Ahn, S. Park, B.-J. Yang, Failure of Nielsen-Ninomiya theorem and fragile topology in two-dimensional systems with space-time inversion symmetry: application to twisted bilayer graphene at magic angle, arXiv:1808.05375.

[26] J. Liu, J. Liu, X. Dai, A complete picture for the band topology in twisted bilayer graphene, arXiv:1810.03103.
Supplementary Material for “Strong coupling phases of partially filled twisted bilayer graphene narrow bands”

I. CONSTRUCTION OF THE WANNIER STATES

In our study, we construct the Wannier states (WSs) of the twisted bilayer graphene with the twist angle of \( \sim 1.3^\circ \) and \( m = 25 \), \( n = 26 \) (for details and notation, see Ref. [S2]). We will follow the projection method [S1] and choose the initial ansatz to have the same symmetry as the final WSs. Although very similar to what has been done in [S2], the initial ansatz here is chosen slightly differently in order to improve the localization of WSs and to maintain the nearly perfect valley polarization:

- \( h_1 \): As shown in Fig. 1(a), our \( h_1 \) is defined as \( \Psi_{\Gamma,E_+,\epsilon} \) only on sublattice A inside the triangle \( 0 - L_1 - L_2 \), and 0 otherwise. This choice guarantees that \( h_1 \) transform in the same way as \( \Psi_{\Gamma,E_+,\epsilon} \) (with \( C_3 \) the eigenvalue of \( \epsilon = \exp(i2\pi/3) \)) under the three-fold rotation around the center of the triangle.

- \( h_2 \): Apply complex conjugation to \( h_1 \). Therefore, \( h_1 \) and \( h_2 \) transform to each other under time reversal, and \( h_2 \) has the eigenvalue of \( \epsilon^* \) under the three-fold rotation around the center of the triangle.

- \( h_3 \): Apply \( C'_2 \) to \( h_1 \). Note that \( h_3 \) is nonzero only inside the triangle \( 0 - (L_2 - L_1) - L_2 \). In addition, \( h_3 \) has the eigenvalue of \( \epsilon^* \) under the three-fold rotation around the center of the triangle.

- \( h_4 \): Apply \( C'_2 \) to \( h_2 \). It is obvious that \( h_4 \) and \( h_3 \) transform to each other under time reversal. In addition, \( h_4 \) has the eigenvalue of \( \epsilon \) under the three-fold rotation around the center of the triangle \( 0 - (L_2 - L_1) - L_2 \).

This ansatz is chosen to improve the localization of the WSs obtained from the projection method.

II. SCREENED COULOMB POTENTIAL

In this section, we will present our numerical result to confirm that the projected Coulomb potential can be written in the form of Eqn. 6 with the constraint that \( \sum_{\mathbf{R},j,\sigma} \hat{O}_{j,\sigma}^\dagger(\mathbf{R}) = \hat{N} \).
As explained in the text, the Coulomb potential is screened due to the image charges induced by two gates above and below the twisted bilayer graphene. As derived in Ref. [S6], the screened Coulomb interaction is given by

\[
V(r) = U_\xi \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{\sqrt{(r/\xi)^2 + n^2}},
\]

where \( \xi \approx 10\text{nm} \) is half of the distance between two gates, and \( \epsilon \approx 6 \) is the dielectric constant of BN. This leads to \( U_\xi = e^2/(4\pi\epsilon\xi) = 24\text{meV} \). Eqn. [S1] works as long as \( r \neq 0 \). In our formula, the on-site repulsion (\( r = 0 \)) is set to be \( 2 \times 9.3 = 18.6\text{eV} \). Since the Coulomb potential decays exponentially for \( |r| > \xi \) and projected on the same hexagon, we consider only the interactions that \( r \) and \( r' \) are located in the same hexagon. Therefore, the Coulomb interaction is well approximated by

\[
U = \frac{1}{2} \sum_{R} \sum_{\sigma\sigma'} \sum_{R,\delta R,\delta'} V(r - r') n_\sigma(R) n_{\sigma'}(R + r') .
\]

We next project the fermion number operator \( n(r) \) to the WSs for the narrow bands. As each site contains four different states with \( j = \pm 1 \) and \( \sigma = \uparrow, \downarrow \),

\[
n_\sigma(r \in \mathcal{O}) \rightarrow \sum_{jj'} \sum_{R,\delta R,\delta'} w_{R+\delta,j}(r) w_{R',\delta',j}(r) d_{j,\sigma}^\dagger (R + \delta) d_{j',\sigma}(R' + \delta'),
\]

where \( R \) (\( R' \)) specifies the unit cell and \( \delta \) (\( \delta' \)) is \( \delta_1 \) or \( \delta_2 \) refers to the two honeycomb lattice sites in the unit cell (illustrated in Fig. 1). If each WS has only three peaks well localized around the neighboring triangular lattice sites, the projected \( n(r) \) is dominated by the WS at the sites of the hexagon. In this supplementary material, however, we will not make this assumption and argue that our conclusions in the main text still holds.

**A. Cluster Hubbard terms**

As explained in the main text, the cluster Hubbard terms can be written as

\[
U_0 = \frac{1}{2} \sum_{\sigma\sigma'} \sum_{j,j'} \sum_{R, p, p'=1} V_{p,j,j',p'}(R + \delta_p) \rho_{j,\sigma}(R + \delta_p) \rho_{j',\sigma'}(R + \delta_{p'}) \quad \text{where}
\]

\[
V_{p,j,j',p'}^{(0)} = \sum_{r,r' \in \mathcal{O}} V(r - r')|w_{\delta_p,j}(r)|^2 |w_{\delta_{p'},j'}(r')|^2.
\]

\( \rho_{j,\sigma}(R + \delta_p) = d_{j,\sigma}^\dagger (R + \delta_p) d_{j,\sigma}(R + \delta_p) \) is the projected on-site fermion number operator. Since \( w_{\delta_p,j}(r) = (w_{\delta_p,j,1}(r))^\dagger \), the interaction constant \( V_{p,j,j',p'}^{(0)} \) is independent of the valley indices \( j \) and \( j' \). In the following, we simplify the notation \( V_{p,j,j',p'}^{(0)} = V_{pp'}^{(0)} \). The on-site charging interaction constants are given in Tab. S1

| \( V_{11}^{(0)} \) | \( V_{12}^{(0)} \) | \( V_{13}^{(0)} \) | \( V_{14}^{(0)} \) | \( V_{23}^{(0)} \) |
|---|---|---|---|---|
| 9.03 | 8.92 | 8.79 | 8.75 | 8.90 |

Table S1. The on-site charging interactions. All other on-site interaction constants can be obtained by the symmetry transformation \( C_3 \) and \( C_2' \). All numbers here are in the unit of meV.

All these numbers are almost identical. Thus, we set \( V_{pp'}^{(0)} \approx V^{(0)} = 8.87\text{meV} \).

**B. Assisted Nearest Neighbor Hopping**

As explained in the main text, the interaction term also includes the assisted nearest neighbor hopping terms. The additional interaction terms can be grouped into two parts: the cross terms between the nearest neighbor hopping
and the on-site density, and the square of the nearest hopping:

\[ U_1 = \frac{1}{2} \sum_R \sum_{p, p'} \sum_{j, j', \sigma} \sum_{\sigma'} \left\{ \left( V^{(1)}_{p, j; p', j'} \rho_{j, \sigma} ( R + \delta_p ) d^\dagger_{j', \sigma'} ( R + \delta_{p'} + 1 ) + h.c. \right) + \left( V^{(2)}_{p, j; p', j'} d_{j, \sigma} ( R + \delta_p ) d^\dagger_{j', \sigma'} ( R + \delta_{p'} + 1 ) + h.c. \right) + \left( V^{(3)}_{p, j; p', j'} d^\dagger_{j, \sigma} ( R + \delta_p ) d_{j', \sigma'} ( R + \delta_{p'} + 1 ) + h.c. \right) \right\} \tag{S5} \]

with the definition of \( \delta_7 \equiv \delta_1 \). Same convention will be used in our paper for notation convenience. These interaction constants are calculated as

\[
V^{(1)}_{p, j; p', j'} = \sum_{r, r' \in \mathbb{O}} V ( r - r' ) \left| w_{\delta_{p, j}(r)} \right|^2 w_{\delta_{p', j'}(r')} \tag{S6}
\]

\[
V^{(2)}_{p, j; p', j'} = \sum_{r, r' \in \mathbb{O}} V ( r - r' ) w_{\delta_{p, j}}(r) w_{\delta_{p, j'}}(r') w_{\delta_{p', j}}(r') w_{\delta_{p', j'}}(r') \tag{S7}
\]

\[
V^{(3)}_{p, j; p', j'} = \sum_{r, r' \in \mathbb{O}} V ( r - r' ) w_{\delta_{p, j}}(r) w_{\delta_{p, j'}}(r') w_{\delta_{p', j}}(r') w_{\delta_{p', j'}}(r') \tag{S8}
\]

It is obvious that the first interaction constants \( V^{(1)} \) is independent of the indices \( j \) and \( V^{(1)}_{p, j; p', j'} = ( V^{(1)}_{p, j; p', j'} )^* \). So we will drop the index \( j \) in \( V^{(1)} \). These interaction constants are listed in Table. S2.

| \( V^{(1)}_{1, 1c} \) | \( V^{(1)}_{1, 2c} \) | \( V^{(1)}_{1, 3c} \) | \( V^{(1)}_{1, 4c} \) | \( V^{(1)}_{1, 5c} \) | \( V^{(1)}_{1, 6c} \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 7.49e\text{0.744}\text{e} | 7.54e\text{0.249}\text{e} | 7.29e\text{0.744}\text{e} | 7.46e\text{0.249}\text{e} | 7.36e\text{0.744}\text{e} | 7.63e\text{0.247}\text{e} |

Table S2. The crossing term between the on-site density and the nearest neighbor hopping. Other interaction constants of the crossing term can be obtained by the symmetry transformation \( C_3 \), \( C_2 \), and time reversal. All numbers here are in the unit of meV.

Therefore, these cross terms can be approximated as

\[
V^{(1)} \sum_R \left( \sum_{j', \sigma'} \sum_{p' = 1, 3, 5} e^{i \theta_j} d^\dagger_{j', \sigma'} ( R + \delta_p ) d_{j', \sigma'} ( R + \delta_{p'} + 1 ) \right) - e^{-i \theta_j} d^\dagger_{j, \sigma} ( R + \delta_{p' + 1} ) d_{j', \sigma'} ( R + \delta_{p'} + 2 ) + h.c. \tag{S9}
\]

with \( \theta \approx 0.75\pi \) and \( V^{(1)} \approx 7.46\text{meV} \).

| \( V^{(2)}_{1, 1c} \) | \( V^{(2)}_{1, 2c} \) | \( V^{(2)}_{1, 3c} \) | \( V^{(2)}_{1, 4c} \) | \( V^{(2)}_{1, 5c} \) | \( V^{(2)}_{1, 6c} \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 6.14e\text{-0.510}\text{e} | 6.26e\text{0.995}\text{e} | 6.10e\text{-0.505}\text{e} | 6.22e\text{-0.505}\text{e} | 6.40e\text{-0.495}\text{e} | 6.39e\text{-0.497}\text{e} |
| \( V^{(2)}_{1, 1c} \) | \( V^{(2)}_{1, 2c} \) | \( V^{(2)}_{1, 3c} \) | \( V^{(2)}_{1, 4c} \) | \( V^{(2)}_{1, 5c} \) | \( V^{(2)}_{1, 6c} \) |
| 6.41 | 6.39e\text{0.498}\text{e} | 6.08 | 6.18e\text{0.499}\text{e} | 6.48 | 6.37 |

| \( V^{(3)}_{1, 1c} \) | \( V^{(3)}_{1, 2c} \) | \( V^{(3)}_{1, 3c} \) | \( V^{(3)}_{1, 4c} \) | \( V^{(3)}_{1, 5c} \) | \( V^{(3)}_{1, 6c} \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 6.39 | 6.38e\text{0.498}\text{e} | 6.07 | 6.16e\text{0.499}\text{e} | 6.46 | 6.36 |
| \( V^{(3)}_{1, 1c} \) | \( V^{(3)}_{1, 2c} \) | \( V^{(3)}_{1, 3c} \) | \( V^{(3)}_{1, 4c} \) | \( V^{(3)}_{1, 5c} \) | \( V^{(3)}_{1, 6c} \) |
| 6.14e\text{-0.510}\text{e} | 6.27e\text{0.995}\text{e} | 6.10e\text{-0.505}\text{e} | 6.22e\text{0.995}\text{e} | 6.40e\text{-0.495}\text{e} | 6.39e\text{-0.497}\text{e} |

Table S3. The square of the nearest neighbor hopping. Other interaction constants can be obtained by the symmetry transformation \( C_3 \), \( C_2 \), and time reversal. All numbers here are in the unit of meV.

Again, the interaction constants in Tab. S3 suggest that the 2nd and 3rd terms in Eqn. S3 can be well approximated as

\[
\frac{V^{(2)}_{2}}{2} \sum_R \left( \sum_{j', \sigma'} \sum_{p' = 1, 3, 5} \left( e^{i \theta_j} d^\dagger_{j', \sigma'} ( R + \delta_p ) d_{j', \sigma'} ( R + \delta_{p'} + 1 ) - e^{-i \theta_j} d^\dagger_{j, \sigma} ( R + \delta_{p' + 1} ) d_{j', \sigma'} ( R + \delta_{p'} + 2 ) \right) + h.c. \right)^2
\]
with $V^{(2)} \approx 6.28\text{meV}$. It is interesting that $V_2 V_0/(V_1)^2 = 1.001 \approx 1$. This suggests that the whole interaction can be written in a simple form:

$$U = \frac{V_0}{2} \sum_{R} (\alpha_0 Q(R) + \alpha_1 T(R))^2$$

where

$$Q(R) = \sum_{p,j,\sigma} \rho_{p,j,\sigma}(R + \delta_p)$$

$$T_1(R) = \sum_{p=1,3,5} \sum_{j,\sigma} (e^{i\theta_j} d_{j,\sigma}^\dagger (R + \delta_p) d_{j,\sigma}(R + \delta_{p+1}) - e^{-i\theta_j} d_{j,\sigma}^\dagger (R + \delta_p) d_{j,\sigma}(R + \delta_{p+2}) + h.c.)$$

In addition, we note that the ratio between two coefficients $\alpha_1/\alpha_0 = V_1/V_0 = 0.84$, very close to the ratio of two WS overlaps inside the hexagon:

$$t_0 = \sum_{r \in \mathcal{O}} |w_1(r)|^2 = 0.286, \quad t_1 = \sum_{r \in \mathcal{O}} w_1^*(r)w_4(r) = 0.23e^{0.743\pi i}, \quad t_1/t_0 = 0.81e^{0.743\pi i}.$$  

We see that the phase $\theta \approx \text{Arg}[t_1]$ and $|t_1|/t_0 \approx \alpha_1/\alpha_0$, suggesting that the approximation introduced in the main text agrees with our numerical calculation very well. Therefore, it is natural to choose that

$$\alpha_0 = t_0, \alpha_1 = |t_1| \implies V_0 = V^{(0)}/\alpha_0^2 = 108\text{meV}, \quad U = \frac{V_0}{2} \sum_{R} (\alpha_0 Q(R) + \alpha_1 T_1(R))^2.$$  

Note that the phase $\theta \neq 0$, but a gauge transformation can be applied to absorb it:

$$\begin{cases}
    d_{j,\sigma}(R + \delta_p) \rightarrow e^{i\theta_j/2} d_{j,\sigma}(R + \delta_p) & \text{for } p = 1, 3, 5 \\
    d_{j,\sigma}(R + \delta_p) \rightarrow e^{-i\theta_j/2} d_{j,\sigma}(R + \delta_p) & \text{for } p = 2, 4, 6.
\end{cases}$$

After applying this gauge transformation, we see that

$$T_1(R) = \sum_{p=1,3,5} \sum_{j,\sigma} (d_{j,\sigma}^\dagger (R + \delta_p) d_{j,\sigma}(R + \delta_{p+1}) - d_{j,\sigma}^\dagger (R + \delta_p) d_{j,\sigma}(R + \delta_{p+2})).$$

C. Assisted Next-nearest Neighbor Hopping

As explained in the text, the assisted hopping between next-nearest neighbor is small because of the orthogonality between two overlapped peaks. However, our numerical calculation shows that this term is still significant because our constructed WSs also contains several secondary peaks, giving rise to the sizable overlap between the next nearest neighbor WSs within the hexagon:

$$\alpha_2 = \sum_{r \in \mathcal{O}} w_{\delta_p}^*(r)w_{\delta_{p+2}}(r) \approx -0.13.$$  

The interaction, when including this next-nearest neighbor assisted hopping, can still be written in a similar form:

$$U = \frac{V_0}{2} \sum_{R} (\alpha_0 Q(R) + \alpha_1 T_1(R) + \alpha_2 T_2(R))^2$$

where $T_2(R) = \sum_{p,\sigma} \sum_{j=\pm 1} d_{j,\sigma}^\dagger(R + \delta_p)d_{j,\sigma}(R + \delta_{p+2}) + h.c.$  

We found $\sum_R \alpha_0 Q(R) + \alpha_1 T_1(R) + \alpha_2 T_2(R) \approx \hat{N}$. This relation can be recovered by including the next nearest hopping beyond the hexagon. Here, we define

$$T_2'(R) = \sum_{j,\sigma} d_{j,\sigma}^\dagger(R + \delta_1)(d_{j,\sigma}(R + \delta_1 + L_1) + d_{j,\sigma}(R + \delta_1 + L_1 - L_2)) +$$

$$d_{j,\sigma}^\dagger(R + \delta_2)(d_{j,\sigma}(R + \delta_2 + L_2) + d_{j,\sigma}(R + \delta_2 + L_1)) +$$

$$d_{j,\sigma}^\dagger(R + \delta_3)(d_{j,\sigma}(R + \delta_3 + L_2) + d_{j,\sigma}(R + \delta_3 + L_2 - L_1)) +$$

$$d_{j,\sigma}^\dagger(R + \delta_4)(d_{j,\sigma}(R + \delta_4 - L_1) + d_{j,\sigma}(R + \delta_4 - L_1 + L_2)) +$$

$$d_{j,\sigma}^\dagger(R + \delta_5)(d_{j,\sigma}(R + \delta_5 - L_1) + d_{j,\sigma}(R + \delta_5 - L_2)) +$$

$$d_{j,\sigma}^\dagger(R + \delta_6)(d_{j,\sigma}(R + \delta_6 - L_2) + d_{j,\sigma}(R + \delta_6 - L_2 + L_1)) + h.c.$$  

(S16)
The interaction also includes the contribution from this additional next nearest neighbor hopping term:

\[ U = V_0 \sum_{\mathbf{R}} (\alpha_0 Q(\mathbf{R}) + \alpha_1 T_1(\mathbf{R}) + \alpha_2 T_2(\mathbf{R}) + \alpha'_2 T'_2(\mathbf{R}))^2 \] (S17)

\[ \alpha'_2 \approx \sum_{r \in \mathcal{O}} U_{\delta_1}(r) w_{L_1+\delta_1}(r) \approx 0.05 \] (S18)

Now, when summing over all the hexagons, we have

\[ \sum_{\mathbf{R}} \alpha_0 Q(\mathbf{R}) + \alpha_1 T_1(\mathbf{R}) + \alpha_2 T_2(\mathbf{R}) + \alpha'_2 T'_2(\mathbf{R}) \approx 3\alpha_0 \hat{N} \approx \hat{N} \]

This relation becomes almost exact when including more hopping terms beyond the hexagon. For simplicity, we stop here and assume that the interaction is

\[ U = \frac{V_0}{2} \sum_{\mathbf{R}} (\alpha_0 Q(\mathbf{R}) + \alpha_1 T_1(\mathbf{R}) + \alpha_2 T_2(\mathbf{R}) + \alpha'_2 T'_2(\mathbf{R}))^2 \]

with \( \alpha_0 = 1/3 \) and \( \alpha'_2 = -\alpha_2/2 \), so that

\[ \sum_{\mathbf{R}} \alpha_0 Q(\mathbf{R}) + \alpha_1 T_1(\mathbf{R}) + \alpha_2 T_2(\mathbf{R}) + \alpha'_2 T'_2(\mathbf{R}) = \hat{N} \].

**III. GROUND STATE AND EXCITED STATES AT 1/4 FILLING**

As explained in the main text, we will follow the strong coupling approach, i.e. to treat the hopping term as perturbation and find the ground states at 1/4 filling. It turns out that the product state

\[ |\Phi_{GS}\rangle = \prod_{\mathbf{R},\delta} d_{j,\sigma_n}^\dagger (\mathbf{R} + \delta)|0\rangle \]

is the ground state of the interaction \( U \). \( \delta \) refers to the two hexagon sites in each unit cell, and the creation operator \( d_{j,\sigma_n}^\dagger \) creates a fermion with the valley \( j \) and spin along the direction of \( \mathbf{n} \). Since the interaction \( U \) is SU(4) symmetric, any hoppings that conserve both spin and valley annihilate \(|\Phi_{GS}\rangle\) because the state on each honeycomb lattice site is identical. Therefore, for any hexagon,

\[ \sum_{j,\sigma} O_{j,\sigma}(\mathbf{R})|\Phi_{GS}\rangle = 6\alpha_0 \implies E_{GS} = 18N_R\alpha_0^2 V_0 \],

where \( N_R \) is the number of the unit cell. It seems that the energy of the ground state depends on the parameter \( \alpha_0 \), which depends on the constructed WSs. However, when including more cluster Hubbard terms from WSs located outside the hexagon,

\[ \alpha_0 + \alpha'_0 + \cdots = \frac{1}{3} \implies E_{GS} = 2N_R V_0 \] .

**A. SU(4) Symmetry Breaking**

In this subsection, we will discuss how the SU(4) degeneracy can be lifted by the kinetic terms in the Hamiltonian. The valley U(1) symmetric hopping terms can be generally written as [S2]

\[
K(\mathbf{R} + \delta, \mathbf{R}' + \delta') = \sum_{\sigma} \left( t(\mathbf{R} + \delta, \mathbf{R}' + \delta') d_{1,\sigma}^\dagger (\mathbf{R} + \delta) d_{1,\sigma} (\mathbf{R}' + \delta') + t^*(\mathbf{R} + \delta, \mathbf{R}' + \delta') d_{-1,\sigma}^\dagger (\mathbf{R} + \delta) d_{-1,\sigma} (\mathbf{R}' + \delta') \right) + h.c.
\]

\[
K = \sum_{\mathbf{R},\delta, \mathbf{R}',\delta'} K(\mathbf{R} + \delta, \mathbf{R}' + \delta')
\] (S19)
The most general form of the ground state is

$$|\Phi_{GS}\rangle = \prod_{R,\delta} \frac{1}{\sqrt{1 + |\beta|^2}} \left( d_{1,\sigma_1}^\dagger (R + \delta) + \beta \ d_{-1,\sigma_2}^\dagger (R + \delta) \right) |0\rangle,$$

(S20)

where $n_1$ ($n_2$) are the direction of the spin polarizations. Note that this is the most general form for the ground state of the interaction $U$. When $\beta = 0$ or $\infty$, the state is both valley and spin polarized. If $|\beta| = 1$, the state mixes both valleys with equal weights.

In our approach, the hopping term in Eqn. (S19) is treated as perturbation. It is clear that the first order perturbation vanishes as $\langle \Phi_{GS}|K|\Phi_{GS}\rangle = 0$ since any hopping changed the fermion number in two or more hexagons. The second order perturbation gives the correction as

$$\delta E_{GS}^{(2)} = -\sum_{\phi_{ex}} \frac{|\langle \phi_{ex}|K|\Phi_{GS}\rangle|^2}{|E_{GS} - E_{ex}|},$$

(S21)

where $\sum_{\phi_{ex}}$ sums over all the excited states. Since the spectrum of the excited states is almost impossible to solve, we instead maximize

$$\sum_{\phi_{ex}} |\langle \phi_{ex}|K|\Phi_{GS}\rangle|^2 = \left| \sum_{R+\delta, R'+\delta'} K(R+\delta, R'+\delta')|\Phi_{GS}\rangle \right|^2 = \sum_{R+\delta, R'+\delta'} \left| K(R+\delta, R'+\delta')|\Phi_{GS}\rangle \right|^2$$

$$\leq \sum_{R+\delta, R'+\delta'} (Im(t(R+\delta, R'+\delta')))^2$$

(S22)

where $\| \cdots \|$ is the norm of the state. Since the hopping constants $t$ are in general complex numbers, $|\beta| = 1$ to maximize the norm. Thus, the ground state is given by the mixture of two valleys with equal weights.

Furthermore, in the presence of the valley mixing hopping terms, the kinetic terms contain

$$K'_i = t_{1d}^i d_{1,\sigma}^\dagger (R_i + \delta_i) d_{-1,\sigma} (R_i + \delta_i) + (t'_{1d})^* d_{-1,\sigma}^\dagger (R_i + \delta_i) d_{1,\sigma} (R_i + \delta_i)$$

following the same approach, we maximize the norm of

$$\|K'|\Phi_{GS}\| = \sum_{il} \frac{|t'_{il}|^2}{1 + |\beta|^2} [1 + |\beta|^2 - 2|\beta|^2 |\langle \sigma_{n_1} | \sigma_{n_2} \rangle|^2]$$

This norm is maximized if and only if $\langle \sigma_{n_1} | \sigma_{n_2} \rangle = 0$, i.e. $n_1 = -n_2 = n$. Thus, the ground state is

$$\prod_{R,\delta} \frac{1}{\sqrt{2}} \left( d_{1,\sigma_1}^\dagger (R + \delta) + e^{i\theta} d_{-1,\sigma_2}^\dagger (R + \delta) \right) |0\rangle$$

Note that the ground state is SU(2) degenerate in spin space since the spin orientation $n$ is not fixed and the phase $\theta$ is arbitrary. For simplicity, we assume $n = \hat{z}$ and set $\theta = 0$.

$$|\Phi_{GS}\rangle = \prod_{R,\delta} \frac{1}{\sqrt{2}} \left( d_{1,1}^\dagger (R + \delta) + d_{-1,1}^\dagger (R + \delta) \right) |0\rangle$$

It is interesting to calculate the expectation value of $S^2$ where $S$ is the total spin:

$$S = \sum_r c_{\alpha}^\dagger (r) \sigma_{\alpha \beta} c_{\beta} (r) \rightarrow \sum_{R} \sum_{p=1}^2 \sum_j d_{j,\alpha}^\dagger (R + \delta_p) \sigma_{\alpha \beta} d_{j,\beta} (R + \delta_p)$$

(S23)

After projection to the WSs, the expectation value of $S^2$ for the ground state is:

$$\langle S^2_s \rangle = \sum_{R, \delta, \delta'} \sum_j \langle \Phi_{GS}|S_{\delta} (R_i + \delta_i) S_{\delta} (R_j + \delta_j) |\Phi_{GS}\rangle$$

(S24)
The formula above is nonzero only when \( R_i + \delta_i = R_j + \delta_j \):

\[
\langle S^2_z \rangle = \sum_{R, \delta} \langle \Phi_{GS} | S^2_z (R + \delta) | \Phi_{GS} \rangle = \frac{N_R}{4},
\]

where \( N \) is the number of honeycomb lattice sites. Similar results are obtained for \( \langle S^2_x \rangle \) and \( \langle S^2_y \rangle \). Therefore,

\[
\langle S^2 \rangle = \frac{3}{4} N_R \propto N_R \implies \sqrt{\langle S^2 \rangle} \propto N_R^{-\frac{1}{2}}
\]

The average magnetic moment per particle is proportional to \( 1/\sqrt{N_R} \), thus vanishes in the thermodynamic limit.

**B. Excited States**

Consider the state

\[
|\Phi_{N+1}\rangle = \frac{1}{\sqrt{N_R}} \sum_R d_{1,1}^{\dagger}(R + \delta_1) |\Phi_{GS}\rangle.
\]

(S25)

To show it is also the eigenstate of the interaction \( U \), note that

\[
O(R) = \alpha_0 Q(R) + \alpha_1 T_1(R) + \alpha_2 T_2(R) + \alpha_2' T_2'(R).
\]

This leads to

\[
[O(R), \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,1}^{\dagger}(R' + \delta_1)] = \frac{1}{\sqrt{N_R}} \sum_{p=1,3,5} \left( (\alpha_0 + 2\alpha_2 + 2\alpha_2') d_{1,1}^{\dagger}(R + \delta_p) + \alpha_2' \left( d_{1,1}^{\dagger}(R + \delta_1 + L_1) + d_{1,1}^{\dagger}(R + \delta_1 + L_1 - L_2) + d_{1,1}^{\dagger}(R + \delta_3 + L_2) + d_{1,1}^{\dagger}(R + \delta_3 + L_2 - L_1) + d_{1,1}^{\dagger}(R + \delta_5 - L_2) \right) \right)
\]

\[
= \sum_R \left[ O(R), \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,1}^{\dagger}(R' + \delta_1) \right] = \frac{3}{\sqrt{N_R}} \sum_{p=1,3,5} (\alpha_0 + 2\alpha_2 + 4\alpha_2') \sum_{R'} d_{1,1}^{\dagger}(R' + \delta_1)
\]

(S26)

\[
\implies \sum_R \left[ O(R), \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,1}^{\dagger}(R' + \delta_1) \right] = \frac{3}{\sqrt{N_R}} \sum_{p=1,3,5} \left( (\alpha_0 + 2\alpha_2 + 2\alpha_2') d_{1,1}^{\dagger}(R + \delta_p) + \alpha_2' \left( d_{1,1}^{\dagger}(R + \delta_1 + L_1) + d_{1,1}^{\dagger}(R + \delta_3 + L_2) + d_{1,1}^{\dagger}(R + \delta_3 + L_2 - L_1) + d_{1,1}^{\dagger}(R + \delta_5 - L_2) \right) \right)
\]

\[
= \sum_R \left[ O(R), \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,1}^{\dagger}(R' + \delta_1) \right] = \frac{3}{\sqrt{N_R}} \sum_{p=1,3,5} \left( (\alpha_0 + 2\alpha_2 + 2\alpha_2') d_{1,1}^{\dagger}(R + \delta_p) + \alpha_2' \left( d_{1,1}^{\dagger}(R + \delta_1 + L_1) + d_{1,1}^{\dagger}(R + \delta_3 + L_2) + d_{1,1}^{\dagger}(R + \delta_3 + L_2 - L_1) + d_{1,1}^{\dagger}(R + \delta_5 - L_2) \right) \right)
\]

(S27)
The energy of the excited state is given as

\[
\left( \sum_{R} O(R)^2 \right) \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,4}^\dagger (R' + \delta_1) |\Phi_{GS}\rangle
\]

\[
= \sum_{R} \left( 2 \left[ O(R), \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,4}^\dagger (R' + \delta_1) \right] O(R) + \left[ O(R), \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,4}^\dagger (R' + \delta_1) \right] \right) + \frac{1}{\sqrt{N_R}} \sum_{R'} d_{1,4}^\dagger (R' + \delta_1) O(R)^2 |\Phi_{GS}\rangle
\]

\[
= 6(\alpha_0 + 2\alpha_2 + 4\alpha_2')6\alpha_0 |\Phi_{N+1}\rangle + 3 \left( (\alpha_0 + 2\alpha_2 + 2\alpha_2')^2 + 2\alpha_2'^2 \right) |\Phi_{N+1}\rangle + 36\alpha_2^2 N_R |\Phi_{N+1}\rangle
\]

\[
= \left( 36\alpha_0(\alpha_0 + 2\alpha_2 + 4\alpha_2') + 3 (\alpha_0 + 2\alpha_2 + 2\alpha_2')^2 + 6\alpha_2'^2 + 36\alpha_2^2 N_R \right) |\Phi_{N+1}\rangle \quad (S28)
\]

Thus, we conclude that

\[
E_{N+1} = \frac{V_0}{2} \left( 36N_R\alpha_0^2 + 36\alpha_0(\alpha_0 + 2\alpha_2 + 4\alpha_2') + 3 (\alpha_0 + 2\alpha_2 + 2\alpha_2')^2 + 6\alpha_2'^2 \right)
\]

Note that the creation operator in Eqn. (S25) is applied only on one sublattice. We can apply the same operator on another sublattice and obtain the eigenstate with the same energy \(E_{N+1}\). In addition, the excited state could also be generated by the creation operator \(d_{1,1}^\dagger\) and \(\frac{1}{\sqrt{2}}(d_{1,1}^\dagger - d_{1,1}^\dagger)\). Thus, there exist six extended states with the same energy.

For \(|\Phi_{N-1}\rangle\) state, we consider

\[
|\Phi_{N-1}\rangle = \frac{1}{\sqrt{N_R}} \sum_{R} \frac{1}{\sqrt{2}} \left( d_{1,\uparrow}(R + \delta_1) + d_{-1,\downarrow}(R + \delta_1) \right) |\Phi_{GS}\rangle.
\]

Following the same method, we conclude that this is also the eigenstate of the interaction \(U\) with the energy of

\[
E_{N-1} = \frac{V_0}{2} \left( 36N\alpha_0^2 - 36\alpha_0(\alpha_0 + 2\alpha_2 + 4\alpha_2') + 3 (\alpha_0 + 2\alpha_2 + 2\alpha_2')^2 + 6\alpha_2'^2 \right)
\]

Thus, the gap to the extended state is

\[
\Delta \leq E_{N+1} + E_{N-1} - 2E_N = 3V_0 \left( (\alpha_0 + 2\alpha_2 + 2\alpha_2')^2 + 2\alpha_2'^2 \right)
\]

For simplicity, we set \(\alpha_2' = -\alpha_2/2\) and \(\alpha_0 = 1/3\) to satisfy the constraint that \(\sum_R O(R) = \hat{N}\). This leads to an upper limit of the gap:

\[
\Delta \leq 3V_0 \left( \left( \frac{1}{3} + \alpha_2 \right)^2 + \frac{\alpha_2^2}{2} \right).
\]

Since \(-1/3 < \alpha_2 < 0\), this assisted next-nearest neighbor hopping term generally decrease the upper limit of the gap.

**IV. RISE OF SU(4) FERROMAGNETIC EXCHANGE INTERACTION**

In this section, we will discuss how the ferromagnetic interactions arise from the assistant hopping terms in \(U\). For simplicity, we include only the on-site particle number \(Q(R)\) and the nearest neighbor hopping \(T_1\) and set \(\alpha_0 = 1/3\). Thus

\[
U = \sum_{R} (O(R))^2 \quad \text{where} \quad O(R) = \frac{1}{3} Q(R) + \alpha_1 T_1(R).
\]
with $Q(R)$ and $T_{ij}(R)$ defined in Eqn. [S11] and [S14]. To understand how the ferromagnetic exchange interaction rises, consider the limit that $\alpha_1 \ll \alpha_0$ and treat the nearest neighbor assisted hopping terms as perturbation.

$$U = U_0 + U_1 + U_2$$  \hspace{1cm} (S29)

$$U_0 = \alpha_0^2 V_0 \sum_R (Q(R))^2 ,$$  \hspace{1cm} (S30)

$$U_1 = 2\alpha_0 V_0 \alpha_1 \sum_R Q(R) T(R),$$  \hspace{1cm} (S31)

$$U_2 = \alpha_1^2 V_0 \sum_R (T(R))^2 .$$  \hspace{1cm} (S32)

As explained in the main text, instead of considering the spectrum of $\sum_R (\alpha_0 Q(R) - 2)^2$ for $1/4$ filling. The lowest energy level is given by the state in which each hexagon contains 6 fermions $(Q(R)|\Phi_{GS} = 6|\Phi_{GS})$. Such states are highly degenerate even without including the valley and spin degrees of freedom. To lift this degeneracy, we consider the perturbative expansion of the small coefficient $\alpha_1$. Note the cross term $Q(R)T(R)$ can be written as

$$U_1 = 2\alpha_0 \alpha_1 V_0 \sum_{\langle ij \rangle} (Q_{R_1} - Q_{R_2}) T_{ij} ,$$

where $\langle ij \rangle$ is a nearest neighbor bond, and $R_1$ and $R_2$ refers to two hexagons that share this bond. $T_{ij}$ is the nearest hopping term on this bond. Therefore, the cross term $U_1$ annihilates all the states in the manifold of the ground states of $U_0$. As a consequence, the perturbative expansion of $U_1$ vanished up to the second order $O(\alpha_1^2)$. The contribution of $U_2$ contains the term

$$\langle \Phi_{GS}|d^\dagger_{j,\sigma}(R + \delta_p) d_{j,\sigma}(R + \delta_{p+q}) d^\dagger_{j',\sigma'}(R + \delta_{p'} d_{j',\sigma'}(R + \delta_{p'+q'}))|\Phi_{GS}\rangle$$

where $q, q' = \pm 1$ for nearest neighbor hopping. It can be shown that this term is nonzero only when $p = p' + q$ and $p' = p + q$. Therefore, we can focus only on the term

$$d^\dagger_{j,\sigma}(R + \delta_p) d_{j,\sigma}(R + \delta_{p+q}) d^\dagger_{j',\sigma'}(R + \delta_{p'}) d_{j',\sigma'}(R + \delta_{p'+q'}) .$$

This term gives the $SU(4)$ ferromagnetic exchange. To be more explicit, this term can be written as the ferromagnetic coupling of two spin operators when $j = j'$ as follows:

$$d^\dagger_{j,\sigma}(R + \delta_p) d_{j,\sigma}(R + \delta_{p+q}) d^\dagger_{j',\sigma'}(R + \delta_{p'}) d_{j',\sigma'}(R + \delta_{p'+q'})$$

$$= -2S_j(R + \delta_p) \cdot S_j(R + \delta_{p+q}) + n_j(R + \delta_p) - \frac{1}{2} n_j(R + \delta_p) n_j(R + \delta_{p+q})$$

\hspace{1cm} (S33)

V. GROUND STATE AT 1/8 FILLING

In this section, we study the ground state at 1/8 filling, ie. one particle/hole per unit cell. We cannot solve the ground state analytically with this filling, even in the strong coupling limit. Therefore, we follow the method in the previous section by treating $\alpha_1$ as a small expansion parameter.

With the “cluster Hubbard” terms only, the energy is minimized if each hexagon contains three fermions. As argued in the previous section, both the linear order and the second order of the cross terms vanish. Up to the second order $O(\alpha_1^2)$, the only contribution to the energy correction comes from

$$\sum_R \sum_p \sum_{q=\pm 1} \sum_{j,\sigma} \sum_{j',\sigma'} \langle \Phi_{GS}|d^\dagger_{j,\sigma}(R + \delta_p) d_{j,\sigma}(R + \delta_{p+q}) d^\dagger_{j',\sigma'}(R + \delta_{p'}) d_{j',\sigma'}(R + \delta_{p'+q'})|\Phi_{GS}\rangle$$

$$= \sum_{R,p \ q=\pm 1} \left\| \sum_{j,\sigma} d^\dagger_{j,\sigma}(R + \delta_p) d_{j,\sigma}(R + \delta_{p+q})|\Phi_{GS}\rangle \right\|^2 ,$$

\hspace{1cm} (S34)

where $|\Phi_{GS}\rangle$ is the ground state of the “cluster Hubbard” terms with the filling of 1/8.
Similar to the state with 1/4 filling, we place fermions with the same state (valley and spin) on the honeycomb site. However, half of the sites must be empty for the 1/8 filling. Therefore, we need to minimize the number of “dangling” bonds connecting an occupied site and an unoccupied site. In the stripe phase, each occupied site has exactly one dangling bond, as illustrated in Fig. 2(b). For any occupied site, if three neighboring sites are occupied, each of these sites must have two dangling bonds to satisfy the constraint that each hexagon has three fermions. As a consequence, the average of dangling bonds for this “star” configuration is 3/2. Therefore, the number of dangling bonds is minimized in the stripe phase, which is the ground state at least for small $\alpha_1$.

---

[S1] N. Marzari and D. Vanderbilt, Maximally Localized Generalized Wannier Functions for Composite Energy Bands, Phys. Rev. B 56, 12847 (1997); N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Maximally Localized Wannier Functions: Theory and Applications, Rev. Mod. Phys. 84, 1419 (2012); R. Sakuma, Symmetry-Adapted Wannier Functions in the Maximal Localization Procedure, Phys. Rev. B 87, 235109 (2013).

[S2] J. Kang and O. Vafek, Symmetry, Maximally Localized Wannier States, and a Low-Energy Model for Twisted Bilayer Graphene Narrow Bands, Phys. Rev. X 8, 031088 (2018).

[S3] 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, Unconventional superconductivity in magic-angle graphene superlattices, Nature 556, 43 (2018).

[S4] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Correlated insulator behaviour at half-filling in magic-angle graphene superlattices, Nature 556, 80 (2018).

[S5] M. Yankowitz, et al., Tuning superconductivity in twisted bilayer graphene, arXiv:1808.07865

[S6] R. E. Throckmorton and O. Vafek, Fermions on bilayer graphene: Symmetry breaking for $B = 0$ and $\nu = 0$, Phys. Rev. B 86, 115447 (2015).

[S7] T. O. Wehling, E. Şaşioğlu, C. Friedrich, A. I. Lichtenstein, M. I. Katsnelson, and S. Blügel, Strength of Effective Coulomb Interactions in Graphene and Graphite, Phys. Rev. Lett. 106, 236805 (2011).