Moiré flat Chern bands and correlated quantum anomalous Hall states generated by spin-orbit couplings in twisted homobilayer MoS$_2$

Benjamin T. Zhou$^1$, Shannon Egan$^1$, and Marcel Franz$^{1,1}$

$^1$Department of Physics and Astronomy & Stewart Blusson Quantum Matter Institute, University of British Columbia, Vancouver BC, Canada V6T 1Z4

We predict that in a twisted homobilayer of transition-metal dichalcogenide MoS$_2$, spin-orbit coupling in the conduction band states from $\pm K$ valleys can give rise to moiré flat bands with nonzero Chern numbers in each valley. The nontrivial band topology originates from a unique combination of angular twist and local mirror symmetry breaking in each individual layer, which results in unusual skyrmionic spin textures in momentum space with skyrmion number $S = \pm 2$. Our Hartree-Fock analysis further suggests that density-density interactions generically drive the system at 1/2-filling into a valley-polarized state, which realizes a correlated quantum anomalous Hall state with Chern number $\mathcal{C} = \pm 2$. Effects of displacement fields are discussed with comparison to nontrivial topology from layer-pseudospin magnetic fields.

Introduction.— The discovery of possible correlated insulating states [1] and superconductivity [2] in magic-angle twisted bilayer graphene has paved a new avenue toward engineering electronic structures where interactions play a decisive role [3–19]. Lately, nontrivial topological properties brought about by the moiré patterns were also unveiled [20, 21]. Under appropriate symmetry breaking conditions, flat bands in twisted graphene acquire nonzero Chern numbers [22–25], which manifest experimentally as quantum anomalous Hall (QAH) states when spin or valley degeneracies are lifted spontaneously through electronic correlations [26–28]. The interplay between correlation and topology uncovered in these systems indicates possibilities for novel topological phases beyond conventional non-interacting theories.

Motivated by advances in twisted graphene systems, explorations into moiré superlattices formed by other two-dimensional materials, such as transition-metal dichalcogenides (TMDs) [29–35] and copper oxides [36, 37], have also seen rapid progress recently. In particular, moiré flat bands in a twisted homobilayer TMD formed by the valence band (VB) states in each K-valley were shown to carry nonzero Chern numbers [29]. The nontrivial band topology arises from combined effects of moiré potential and interlayer coupling, which act together as a layer-pseudospin magnetic field and create skyrmionic pseudospin textures in superlattice cells. The strong spin-valley locking due to giant Ising-type spin-orbit coupling (SOC) of order $\sim 100$ meVs [38, 39] further entails a possible quantum spin Hall state.

While flat bands from VB states have enjoyed wide interest, moiré physics arising from conduction band (CB) states in twisted TMDs remains largely unexplored. One possible complication lies in the relatively weak SOC in the conduction bands on the scale of a few to tens of meVs [39, 40], which may cause crossings between spin-up and spin-down moiré bands. Recent works on untwisted TMDs, on the other hand, revealed that the small spin-orbit splitting in CB states, when combined with Rashba SOC introduced by mirror symmetry breaking, can lead to nontrivial Berry phase effects [41–43]. Notably, upon assembling two identical TMD layers into a twisted bilayer, the mirror symmetry is broken locally in each layer (Fig.1(a)-(b)) and Rashba SOC is expected to arise. How this SOC would affect the band topology of twisted TMDs is an outstanding question.

In this Letter we establish a novel topological phase generated by SOC in CB moiré bands in twisted TMDs, which is essentially different from those in twisted graphene [20–25] where SOC is negligible, and those in the VB moiré bands of twisted TMDs where SOC plays no essential role in creating nonzero Chern numbers [29]. Focusing on the specific case of a twisted homobilayer MoS$_2$ we predict that an interplay among angular twist, local symmetry breaking and spin-orbit coupling in the CB states creates moiré flat bands with nonzero Chern numbers $\mathcal{C} = \pm 2$, and density-density interactions generically drive the system at 1/2-filling into a valley-polarized correlated QAH state. We further show that the Chern bands generated by SOC stay robust against displacement fields which would otherwise destroy the nontrivial topology from layer-pseudospins [29].

Continuum model for twisted homobilayer MoS$_2$.— Stacking two monolayers of MoS$_2$ with a small relative

FIG. 1: Schematic crystal structure of twisted bilayer MoS$_2$ with (a) top view and (b) side view. Electrons in different layers experience opposite effective electric fields due to local lack of mirror symmetry. (c) Moiré Brillouin zone (MBZ) formed by twisting hexagonal Brillouin zones of each layer. The in-plane spinor field exhibits opposite vorticities at $K_{m,+}$ and $K_{m,-}$ of MBZ due to layer-dependent Rashba effects.
twist angle $\theta_0$ results in a moiré superlattice with lattice constant $a_M = a/\theta_0$, where $a$ is the lattice constant of each monolayer. The system has a symmetry group of $D_3 \otimes \{U_\nu(1), \mathcal{T}\}$, where $U_\nu(1)$ stands for valley conservation, $\mathcal{T}$ for time-reversal symmetry, and the point group $D_3$ is generated by the three-fold rotation about the $z$-axis ($C_{3z}$) and a two-fold rotation about the $y$-axis ($C_{2y}$) (Fig.1(a)). For an aligned bilayer at $\theta_0 = 0$, the mirror symmetry $\mathcal{M}_z$ about the horizontal 2D mirror plane, which is respected in a monolayer, is broken locally on each layer. This generates uniform electric fields of opposite signs in different layers, which remain effective upon a small angular twist with $\theta_0 \sim 1^\circ$ (Fig.1(b)). Based on models of monolayer MoS$_2$ with broken $\mathcal{M}_z$ [41–43], the effective Hamiltonian for CB minima at $\pm k$ valleys of two twisted isolated layers can be written as:

$$\mathcal{H}_{\xi,0}^E = \sum_{k,\alpha,\beta} c_{k,\alpha,\beta}^\dagger c_{k,\alpha,\beta},$$  

$$h_{\xi,0}^E(k) = \frac{\hbar^2(k - \xi K)^2}{2m^*} - \mu + \xi \alpha_0 \sigma_z + (-1)^l \alpha_0 [k - \xi K] \times \sigma \cdot \hat{z},$$

where $\xi = \pm$ denotes the valley index, $l = 1(2)$ labels the top (bottom) layer, $\alpha, \beta$ label the spin indices with the $\sigma$-matrices representing the usual Pauli matrices for spins. $K_{1} \equiv K_{m,+}$ and $K_{2} \equiv K_{m,-}$ are the shifted $+K$-points on layer 1 and layer 2 (Fig.1(c)) due to the angular twist of $\theta_0$ equivalent to rotating the top (bottom) layer by an angle $\theta_0/2$ ($-\theta_0/2$) about the $z$-axis. In Eq.1, $m^* \approx 0.5m_e$ is the CB effective mass at $K$, $\mu$ is the Fermi energy measured from band minima. The $\alpha_0$-term and $\beta_{\alpha_0}$-term account for the Rashba SOC and the Ising SOC, respectively. The Ising SOC has a valley-dependent sign due to time-reversal symmetry and causes a splitting of $\Delta_{\alpha_0} = 2|\beta_{\alpha_0}|$ in the spin-up and spin-down CB states. The Rashba SOC has a layer-dependent sign imposed by $C_{2y}$ which swaps the two layers. Spatial variations of $\alpha_0$ are discussed in Supplemental Material (SM).

Spatial modulations due to the formation of moiré pattern can be captured by introducing coupling terms between states at $k$ and $k + G_{M,j}$ [23, 29], where $G_{M,j} = \frac{2\pi}{3\sqrt{3}a_M} \left(\cos \frac{(j-1)\pi}{3}, \sin \frac{(j-1)\pi}{3}\right)$ are the reciprocal vectors of the moiré superlattice. The momentum-space moiré potential is given by:

$$\mathcal{H}_M = \sum_{k} \sum_{j=1}^{6} \sum_{\alpha=\uparrow,\downarrow} \sum_{\alpha'=\uparrow,\downarrow} c_{k,\alpha}^\dagger V_{ij} G_{M,j} c_{k+G_{M,j},\alpha'},$$

where $V_{ij} G_{M,j} = V_M (\cos \psi - i(-1)^j \sin \psi)$ are complex coupling parameters with amplitude $V_M$ and phase $\psi$.

The effective inter-layer tunneling for states at $K$-valleys is modeled following the general recipe of two-center approximation [4, 29], which can be written as:

$$\mathcal{H}_{\xi,0}^T = -u_0 \sum_{k} c_{k,1}^\dagger (c_{k,2} + c_{k+\xi G_{M,2},2} + c_{k+\xi G_{M,3},2}) + \text{h.c.},$$

where $u_0$ denotes the effective inter-layer tunneling amplitude near $K$, which is extrapolated to be $u_0 \approx 10$ meV using a combined approach of empirical models and Slater-Koster methods (see the Supplemental Material [44] for details). The total non-interacting continuum model thus reads: $\mathcal{H}_0 = \sum_{\xi=\pm} \sum_{l=1,2} \mathcal{H}_{\xi,0}^E + \mathcal{H}_T + \mathcal{H}_M$ with parameters tabulated in Table I.

**Moiré flat Chern bands and skyrmionic spin textures.**— From general considerations, level spacings among the lowest moiré bands correspond roughly to the quantization energy $\Delta E \sim \hbar^2 \pi^2/(2m^* a_M^2)$ of an electron confined in a superlattice cell with size $a_M^2 = a^2/\theta_0^2$. For $\theta_0 \sim 1^\circ - 2^\circ$, $\Delta E \sim 1 - 10$ meV in twisted MoS$_2$, which is comparable to the spin-orbit splitting $\Delta_{\alpha_0} = 2|\beta_{\alpha_0}| = 3$ meV caused by the Ising SOC [39]. If spin is conserved (e.g., in the absence of Rashba SOC), spin-up and spin-down moiré bands are expected to cross in general.

By setting $\alpha_0 = 0$ in $\mathcal{H}_0$, we find that for $\xi = +$, the 1st-spin-down and the 2nd-spin-up moiré bands cross each other for $\theta_0 \in (1.25^\circ, 1.5^\circ)$. As a specific example, spin-resolved moiré bands at $\theta_0 = 1.4^\circ$ with $\alpha_0 = 0$ are shown in Fig.2(a). With spin conservation in this case, the spin Chern number is well-defined, and the spin-up and spin-down bands involved in the level crossing have Chern numbers ($C_{1,\uparrow} = +1$, $C_{2,\uparrow} = -1$), which originate from the layer-pseudospin magnetic fields as in moiré bands from VB states [29]. Upon turning on $\alpha_0$, crossing points are gapped out by the layer-dependent Rashba SOC.
which mix the spin-up and spin-down species (Fig. 2(b)). As a result of this mixing, the original $C_{1\uparrow}$ and $C_{2\uparrow}$ from layer pseudospins annihilate each other; meanwhile, a new pair of flat bands (2nd and 3rd bands in Fig. 2(b)) with Chern numbers $C = \pm 2$ emerge (see SM [44] for details on Chern number calculations). The nontrivial gap induced by Rashba SOCs is further signified by giant Berry curvatures of order $10^4 \text{Å}^2$ in the MBZ (Fig. 2(c)).

The unusual Chern numbers $C = \pm 2$ arise from a unique combination of angular twist and local mirror-symmetry breaking: as moiré bands at opposite $K_{m,+}$ and $K_{m,-}$ points originate from states in different layers (Fig. 1(c)), the layer-dependent Rashba SOCs (Eq. 1) create an in-plane spinor field with opposite vorticities (Fig. 1(c)), the layer-dependent Rashba SOCs (Eq. 1) create an in-plane spinor field with opposite vorticities (Fig. 1(c)).

Including density-density interactions, the low-energy effective Hamiltonian reads:

\[
\mathcal{H}_{\text{eff}} = \mathcal{H}_{0,\text{eff}} + \mathcal{H}_{I,\text{eff}},
\]

where $\mathcal{H}_{0,\text{eff}} = \sum_{\mathbf{k},\mathbf{q},\xi\xi'} E_{\xi}(\mathbf{k}) c_{\xi}^{\dagger}(\mathbf{k}) c_{\xi'}(\mathbf{k})$ is the non-interacting part with band energies $E_{\xi}(\mathbf{k})$, and the interacting Hamiltonian compatible with $U(1)$ and $T$ reads

\[
\mathcal{H}_{I,\text{eff}} = \sum_{\mathbf{k},\mathbf{q},\xi\xi'} \frac{V(\mathbf{q})}{A} \Lambda^{\xi}(\mathbf{k} + \mathbf{q}, \mathbf{k}) \Lambda^{\xi}(\mathbf{k}' - \mathbf{q}, \mathbf{k}') \tag{4}
\]

where $V(\mathbf{q})$ is the Fourier transform of the two-body interaction, $A$ is the total area of the system, and $\Lambda^{\xi}(\mathbf{k} + \mathbf{q}, \mathbf{k}) \equiv \langle u_{\xi,k+q} | u_{\xi,k} \rangle$ (periodic part of the Bloch states) are the form factors arising from projecting the density-density interaction to active bands. The general trial Hartree-Fock ground state at 1/2-filling has the form: $|\Phi\rangle = \prod_{\mathbf{k}\in\text{MBZ}} [w_{\xi}(\mathbf{k}) c_{\xi}^{\dagger}(\mathbf{k}) + \bar{w}_{\xi}(\mathbf{k}) c_{\xi}^{\dagger}(\mathbf{k})] |\Phi_0\rangle$, where $|\Phi_0\rangle$ is the vacuum state with all lower-lying bands filled, and $w_{\xi}(\mathbf{k})$ are the variational parameters.

Minimizing $\mathcal{E} = \langle \Phi | \mathcal{H}_{\text{eff}} | \Phi \rangle$ with mean-field approximations for Fock exchange terms (see SM [44]) leads to a set of self-consistent equations:

\[
\begin{align*}
\Delta_+ + \Delta_- &= J_0, \\
\Delta_+ - \Delta_- &= \frac{J_1}{2} \sum_{\mathbf{k}} \frac{(\Delta_+ - \Delta_-) - \delta(\mathbf{k})}{D(\mathbf{k})}, \\
\Delta_+ &= \frac{J_1}{2} \sum_{\mathbf{k}} \frac{\Delta_+ - \Delta_-}{D(\mathbf{k})}.
\end{align*}
\]

Here, $J_0$ and $J_1$ denote the mean-field intra-valley and inter-valley exchange coupling constants, $\Delta_+ = J_0 n_\xi (\bar{n}_\xi)$: mean occupancy for valley $\xi$ and $\Delta_{\pm}$ are the intra-valley and inter-valley order parameters, which characterize the average energy gains from intra-valley and inter-valley exchange interactions, respectively. $\delta(\mathbf{k}) \equiv E_+(\mathbf{k}) - E_-(\mathbf{k})$ denotes energy difference between bands from two valleys $\xi = \pm$, $D(\mathbf{k}) = \sqrt{\Delta_+^2 + |\delta(\mathbf{k}) - \Delta_M|^2}/4$, where we introduce the valley magnetic order $\Delta_M \equiv \Delta_+ - \Delta_- = J_0 (\overline{n}_+ - \overline{n}_-)$.

Solutions of Eq. S27 are divided into two classes: (i) the valley-polarized (VP) state [6, 23], in which one out of the two nearly degenerate moiré bands from two valleys is fully filled, with spontaneous $T$-symmetry breaking: $\Delta_M = \xi J_0, \Delta_+ = 0, w_{\xi}(\mathbf{k}) = 1$; (ii) the inter-valley coherent (IVC) state, in which the Slater determinant is formed by coherent superpositions of states from the two valleys, with spontaneous U(1)-symmetry breaking [6, 23, 24]: $\Delta_M = 0, \Delta_+ \approx J_1/2, w_{\pm}(\mathbf{k}) \neq 0$. By comparing the energies of the VP and IVC states, we obtain the mean-field $J_0 - J_1$ phase diagram for the range $J_0 \sim J_1$.}

### Table I: Parameters used in $\mathcal{H}_0$ for twisted homobilayer MoS$_2$ with monolayer lattice constant $a = 3.16 \text{Å}$.

| Parameter | Value |
|-----------|-------|
| $\mu$     | 0 meV |
| $\alpha_{so}$ | 80 meV $\cdot$ Å $^{-1}$ |
| $\beta_{so}$ | 1.5 meV |
| $V_M$      | 10 meV |
| $\psi$     | -89.6$^\circ$ |
| $u_0$      | 10 meV |

### FIG. 3: (a) $J_0 - J_1$ phase diagram of the system with the 2nd moiré band at 1/2-filling. Under repulsive density-density interactions, only the $J_1 \leq J_0$ regime can be accessed [44], where the valley-polarized (VP) state is generically favored and the system becomes a correlated QAH insulator. (b) Evolution of the order parameters $\Delta_+$ and $\Delta_-$ with the ratio $J_0/J_1$ is tuned along the line-cut (green dashed line in (a)) across the phase boundary. A first-order transition happens at the critical point $J_0 \approx J_1$. |
The IVC state is favored throughout the entire $J_1 \geq J_0$ regime, while the VP phase takes up most of the $J_1 < J_0$ regime. By tuning $J_0/J_1$ across the phase boundary (green dashed line in Fig.3(a)), a first-order transition occurs at $J_0 \simeq J_1$ (Fig.3(b)).

It is worth noting that given a general repulsive interaction $V(q) > 0$, the inequality $2|\Delta^+(k', k)||\Delta^-(k, k')| \leq |\Delta^+(k', k)|^2 + |\Delta^-(k, k')|^2$ together with $\mathcal{T}$-symmetry leads to $V_0 \geq J_1$, regardless of details of the form factors and the microscopic interaction [44]. Thus, electron correlations most likely drive the system at 1/2-filling into a $\mathcal{T}$-broken VP state, which realizes a correlated QAH state with $\mathcal{C} = \pm 2$. The case of general filling away from 1/2 is discussed in SM [44].

Effects of displacement fields and comparison with Chern bands from layer pseudospins.— As shown in Fig.2(a), with $\alpha_{\text{so}} = 0$, nonzero Chern numbers also arise in CB moiré bands due to layer pseudospin magnetic fields, with the same origin as those in VB moiré bands [29]. Thus, two different mechanisms for moiré Chern bands, one from layer pseudospin and the other from SOCs, coexist in the CB moiré bands. In contrast, the Rashba SOC is unimportant in the CB moiré bands where spin-up and spin-down bands are separated by 100 – 400 meV [39], and the CB moiré band topology is governed by layer pseudospin alone.

The nontrivial topology from layer pseudospin is shown to be prone to displacement fields which can destroy the skyrmionic pseudospin textures by polarizing the layer-pseudospins [29]. On the other hand, the SOC mechanism has an independent origin in the avoided level crossings between spin-up and spin-down bands and does not require nontrivial layer pseudospin textures. Thus one would expect the nontrivial band topology in the CB moiré bands to be more robust against displacement fields than its VB counterpart.

To demonstrate the effects of displacement fields, we follow Refs. [29] and introduce a layer-dependent potential $V_{D,i} = (-1)^i V_z/2, (i = 1, 2)$ in $\mathcal{H}_0$. The moiré bands at $\theta_0 = 1.4^\circ$ under finite $V_z$ are shown in Fig.4(a), where we set $V_z = 4$ meV strong enough to destroy the nontrivial topology generated by layer pseudospins (Fig.4(b)). Clearly, the avoided level crossings still occur and a pair of Chern bands with Chern numbers $(+1, -1)$ remain. The Chern number is reduced from $\mathcal{C} = \pm 2$ to $\mathcal{C} = \pm 1$ under $V_z$ because the displacement field drives a band inversion near $K_{m+}$, and mediates a Chern number exchange $\Delta \mathcal{C} = \pm 1$ between the 1st and 2nd moiré bands.

The topological phase diagram for CB moiré bands as a function of $\theta_0$ and $V_z$ is shown in Fig.4(b). Due to the extra mechanism from SOCs, the critical displacement field for the nontrivial topological regime is enhanced to be $V_z^{c1}$, which is 2-4 times of the critical $V_z^{c1}$ (yellow dashed line) needed to destroy the layer pseudospin mechanism for $\theta_0 \sim 1^\circ$ ($V_z^{c1}$ obtained by setting $\alpha_{\text{so}} = 0$ in $\mathcal{H}_0$). Importantly, there is a wide parameter regime in the phase diagram (depicted in orange) where the layer pseudospin mechanism proposed in Ref. [29] fails while the CB moiré bands remain topological, which confirms that the CB states are more robust than VB states against displacement fields.

Conclusion and Discussions.— While on-going activities in twisted TMDs have focused mainly on VB moiré bands [29–35], our proposal of moiré Chern bands generated by SOC opens a new pathway into the largely unknown territory of CB moiré physics. When these Chern bands are half-filled, according to our predictions, electron interactions can lead to a spontaneously $\mathcal{T}$-broken VP phase and realize a correlated QAH state, which will manifest itself through quantized Hall conductance in transport measurements.

To realize the topological moiré bands generated by SOC, any finite Rashba SOC which is allowed by the $D_3$-symmetry of the system would be sufficient, in principle, to induce a nontrivial gap. Notably, as MoS$_2$ is intrinsically semiconducting (with Fermi level $\sim 0.8$ eV below the conduction band edge) [38, 39, 46], it is necessary to gate the chemical potential such that the CB moiré bands are filled in the first place. With the dual-gate setup used widely in experiments [1, 34], local mirror symmetry breaking can be further enhanced by interfacial contact with gating electrodes, which enhances the SOC gap in the non-interacting bands. Due to the correlated nature of the QAH state at 1/2-filling, the actual topological gap to be manifested experimentally is not solely determined by the SOC gap in the non-interacting bands; instead it should be largely

FIG. 4: (a) CB Moiré bands for valley $\xi = +$ of twisted bilayer MoS$_2$ at $\theta_0 = 1.4^\circ$ and $V_z = 4$ meV. Under finite $V_z$, spin-up and spin-down bands still cross, and the avoided level crossing due to SOC result in a pair of bands with $\mathcal{C} = \pm 1$. (b) Topological phase diagram of CB moiré bands as a function of $\theta_0$ and $V_z$, with an upper boundary $V_z^{c2}$ of the entire topological regime. Yellow dashed line: critical $V_z^{c1}$ above which nontrivial topology generated by layer-pseudospin is destroyed. Values of $V_z^{c1}$ are similar to those found in VB moiré bands [29]. Area enclosed by the red dashed lines: regime with level crossings between spin-up and spin-down CB moiré bands when $\alpha_{\text{so}} = 0$. Area in orange: topological regime where layer pseudospin mechanism in Ref.[29] fails.
Acknowledgement. — B.T.Z. thanks Zefei Wu and Ziliang Ye for inspiring discussions on experimental aspects of twisted MoS$_2$. This work was supported by NSERC and the Canada First Research Excellence Fund, Quantum Materials and Future Technologies Program. B.T.Z. further acknowledges the support of the Croucher Foundation.

* Corresponding author: benjamin.zhou@ubc.ca
† Corresponding author: franz@phas.ubc.ca

[1] Yuan Cao, Valla Fatemi, Ahmet Demir, Shiang Fang, Spencer L. Tomarken, Jason Y. Luo, Javier D. Sanchez-Yamagishi, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, Ray C. Ashoori & Pablo Jarillo-Herrero, Correlated insulator behaviour at half-filling in magic-angle graphene superlattices, *Nature* **556**, 80-84 (2018).

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

[3] Stephen Carr, Daniel Massart, Shiang Fang, Paul Cazeaux, Mitchell Luskin, and Efthimios Kaxiras, Twistrronics: Manipulating the electronic properties of two-dimensional layered structures through their twist angle, *Phys. Rev. B* **95**, 075420 (2017).

[4] Rafi Bistritzer and Allan H. MacDonald, Moiré bands in twisted double-layer graphene, *PNAS* **108** (30) 12233-12237 (2011).

[5] Eva Y. Andrei & Allan H. MacDonald, Graphene bilayers with a twist, *Nature Materials* **19**, 1265–1275 (2020).

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

[7] Noah F. Q. Yuan and Liang Fu, Model for the metal-insulator transition in graphene superlattices and beyond, *Phys. Rev. B* **98**, 045103 (2018).

[8] Mikito Koshino et al., Maximally Localized Wannier Orbitals and the Extended Hubbard Model for Twisted Bilayer Graphene, *Phys. Rev. X* **8**, 031087 (2018).

[9] Hiroki Iseobe, Noah F. Q. Yuan and Liang Fu, Unconventional Superconductivity and Density Waves in Twisted Bilayer Graphene, *Phys. Rev. X* **8**, 041041 (2018).

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

[11] Fengcheng Wu, A. H. MacDonald, and Ivar Martin, Theory of Phonon-Mediated Superconductivity in Twisted Bilayer Graphene, *Phys. Rev. Lett.* **121**, 257001 (2018).

[12] Cheng-Cheng Liu, Li-Da Zhang, Wei-Qiang Chen, and Fan Yang, Chiral Spin Density Wave and $d+i d$ Superconductivity in the Magic-Angle-Twisted Bilayer Graphene, *Phys. Rev. Lett.* **121**, 217001 (2018).

[13] Biao Lian, Zhijun Wang, and B. Andrei Bernevig, Twisted Bilayer Graphene: A Phonon-Driven Superconductor, *Phys. Rev. Lett.* **122**, 257002 (2019).

[14] J. González and T. Stauber, Kohn-Luttinger Superconductivity in Twisted Bilayer Graphene, *Phys. Rev. Lett.* **122**, 026801 (2019).

[15] Ming Xie and A. H. MacDonald, Nature of the Correlated Insulator States in Twisted Bilayer Graphene, *Phys. Rev. Lett.* **124**, 097601 (2020).

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

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

[18] Youngjoon Choi, Jeannette Kemmer, Yang Peng, Alex Thomson, Harpreet Arora, Robert Polski, Yiran Zhang, Hechen Ren, Jason Aicea, Gil Refael, Felix von Oppen, Kenji Watanabe, Takashi Taniguchi & Stevan Nadj-Perge, Electronic correlations in twisted bilayer graphene near the magic angle, *Nature Physics* **15**, 1174–1180 (2019).

[19] Yonglong Xie, Biao Lian, Berthold Jäck, Xiaomeng Liu, Cheng-Li Chiu, Kenji Watanabe, Takashi Taniguchi, B. Andrei Bernevig & Ali Yazdani, Spectroscopic signatures of many-body correlations in magic-angle twisted bilayer graphene, *Nature* **572**, 101–105 (2019).

[20] Jianpeng Liu, Junwei Liu, and Xi Dai, Pseudo Landau level representation of twisted bilayer graphene: Band topology and implications on the correlated insulating phase, *Phys. Rev. B* **99**, 155415 (2019).

[21] Zhida Song, Zhijun Wang, Wujun Shi, Gang Li, Chen Fang, and B. Andrei Bernevig, All Magic Angles in Twisted Bilayer Graphene are Topological, *Phys. Rev. Lett.* **123**, 036401 (2019).

[22] Nick Bultinck, Shubhayu Chatterjee, and Michael P. Zaletel, Mechanism for Anomalous Hall Ferromagnetism in Twisted Bilayer Graphene, *Phys. Rev. Lett.* **124**, 166601 (2020).

[23] Ya-Hui Zhang, Dan Mao, Yuan Cao, Pablo Jarillo-Herrero, and T. Senthil, Nearly flat Chern bands in moiré superlattices, *Phys. Rev. B* **99**, 075127 (2019).

[24] Ya-Hui Zhang, Dan Mao, and T. Senthil, Twisted bilayer graphene aligned with hexagonal boron nitride: Anomalous Hall effect and a lattice model, *Phys. Rev. Research* **1**, 033126 (2019).

[25] Wen-Yu He, David Goldhaber-Gordon & K. T. Law, Giant orbital magnetoelectric effect and current-induced magnetization switching in twisted bilayer graphene, *Nature Communications* **11**, 1650 (2020).

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

[27] M. Serlin, C. L. Tschirhart, H. Polshyn, Y. Zhang, J. Zhu, K. Watanabe, T. Taniguchi, L. Balents, A. F. Young, Intrinsic quantized anomalous Hall effect in a moiré heterostructure, *Science* **367**, 900-903 (2020).

[28] Kevin P. Nuckolls, Myungchul Oh, Dillon Wong, Biao K. Kosmider, J. W. Gonzalez, and J. Fernandez-Rossier, Di Xiao, Gui-Bin Liu, Wanxiang Feng, Xiaodong Xu, and Pavel A. Volkov, Justin H. Wilson, J. H. Pixley, Magic correlated ground at 1/2-filling and discussions on the case of general filling; (iv) estimate of spatial variation in Rashba SOC strength, which includes Refs.[47–50].

[43] Kyung-Han Kim and Hyun-Woo Lee, Berry curvature in monolayer MoS2 with broken mirror symmetry, *Phys. Rev. B* **97**, 235423 (2018).

[44] See the Supplemental Material for details on: (i) calculation of effective inter-layer tunneling parameter $t_0$; (ii) method used for computing Chern numbers in the moiré bands; (iii) self-consistent Hartree-Fock calculations for correlated ground at 1/2-filling and discussions on the case of general filling; (iv) estimate of spatial variation in Rashba SOC strength, which includes Refs.[47–50].

[45] B. A. Bernevig, Chapter 8.3, *Topological Insulators and Topological Superconductors*, Princeton University Press (2013).

[46] Kin Fai Mak, Changgu Lee, James Hone, Jie Shan, and Tony F. Heinz, Atomically Thin MoS2: A New Direct-Gap Semiconductor, *Phys. Rev. Lett.* **105**, 136805 (2010).

[47] Kowsalya Devi Rasamani, Farbod Alimohammadi, Yang Sun, Interlayer-expanded MoS2, *Materials Today* **20**, 83-91, (2017).

[48] Ferdows Zahid, Lei Liu, Yu Zhu, Jian Wang, and Hong Guo, A generic tight-binding model for monolayer, bilayer and bulk MoS2, *AIP Advances* **3**, 052111 (2013).

[49] Takahiro Fukui, Yasuhiro Hatsugai, and Hiroshi Suzuki, Chern Numbers in Discretized Brillouin Zone: Efficient Method of Computing (Spin) Hall Conductances, *J. Phys. Soc. Jpn.* **74**, 1674 (2005).

[50] Qin-Fang Yao, Jia Cai, Wen-Yi Tong, Shi-Jing Gong, Ji-Qing Wang, Xiangang Wan, Chun-Gang Duan, and J. H. Chu, Manipulation of the large Rashba spin splitting in polar two-dimensional transition-metal dichalcogenides, *Phys. Rev. B* **95**, 165401 (2017).
Supplemental Material for
“Moiré Chern bands and correlated quantum anomalous Hall states
generated by spin-orbit couplings in twisted homobilayer MoS2”

Benjamin T. Zhou,1 Shannon Egan,1 Marcel Franz1

1Department of Physics and Astronomy & Stewart Blusson Quantum Matter Institute, University of British
Columbia, Vancouver BC, Canada V6T 1Z4

I. Calculation of effective inter-layer tunneling parameter \( w_0 \)

In the main text, we mention that an effective inter-layer tunneling parameter \( w_0 \) is extrapolated using a combined
approach of empirical model and Slater-Koster method. Here, we present details on the calculation of \( w_0 \).

The inter-layer tunneling between two MoS\(_2\) monolayers with a small twist angle can be conveniently modeled by
the two-center approximation following the general recipe introduced in previous studies on twisted bilayer graphene [S1]
and twisted homobilayer MoTe\(_2\) [S2]. Particularly, electronic states near the conduction band minimum of monolayer
transition-metal dichalcogenides (TMDs) are predominantly from the 4d\(_{z^2}\)-orbitals of the transition metal atoms [S3]
and the inter-layer tunneling is expected to be mediated by the \( \sigma\)-bonding between 4d\(_{z^2}\)-orbitals in each layer. To start with, we consider fermionic operators that create Bloch states of the form:

\[
c_{k,1,\alpha}^\dagger = \frac{1}{\sqrt{N}} \sum_{R} e^{i k \cdot R} c_{1,\alpha}^\dagger (R),
\]

\[
c_{k',2,\alpha}^\dagger = \frac{1}{\sqrt{N}} \sum_{R'} e^{i k' \cdot (R' + d_0)} c_{2,\alpha}^\dagger (R' + d_0).
\]

Here, \( k, R \) and \( k', R' \) denote the momentum and lattice vectors of layer 1 and 2, respectively. \( c_{l,\alpha}^\dagger (R) \) creates a localized Wannier orbital of \( d_{z^2} \) character with spin \( \alpha \) at site \( R \) in layer \( l \). In general, the inter-layer tunneling Hamiltonian is given by:

\[
\mathcal{H}_T = \sum_{k,k',\alpha,\beta} c_{k,1,\alpha}^\dagger T_{\alpha\beta}(k,k') c_{k',2,\beta}^\dagger + h.c.
\]

(S2)

where the tunneling matrix \( T \) reads:

\[
T_{\alpha\beta}(k,k') = \langle k, 1, \alpha | \mathcal{H}_T | k', 2, \beta \rangle
\]

\[
= \frac{1}{N} \sum_{R,R'} e^{-i k \cdot R} e^{i k' \cdot (R' + d_0)} \langle R, 1, \alpha | \mathcal{H}_T | R' + d_0, 2, \beta \rangle
\]

\[
= -\frac{1}{N} \sum_{R,R',q} e^{-i k \cdot R} e^{i k' \cdot (R' + d_0)} t_{\alpha\beta}(q) e^{i q \cdot (R' - d_0)}
\]

\[
= -\frac{1}{N} \sum_{R,R',q} \frac{t_{\alpha\beta}(q)}{A} e^{i (k' - q) \cdot d_0} \left( \sum_{R} e^{i (q - k) \cdot R} \right) \left( \sum_{R'} e^{i (k' - q) \cdot R'} \right)
\]

\[
= -\frac{1}{N} \sum_{G,G'} t_{\alpha\beta} \left( k + G \right) \Omega e^{-i G' \cdot d_0} \delta_{k + G,k' + G'}
\]

(S3)

Here, \( t_{\alpha\beta}(q) = \int d^2 \tau t_{\alpha\beta}(\tau)e^{-i q \cdot \tau} \) is the Fourier transform of \( t_{\alpha\beta}(\tau) \), \( A \) is the total area of the system, \( \Omega = A/N \) is the
area of the primitive unit cell. In the last step of the derivation above, we made use of the identity
\( \sum_{R} e^{i (q - k) \cdot R} = N \sum_{G} \delta_{q,k + G} \).

To simplify the expression of \( T_{\alpha\beta}(k,k') \) above, we note that for the relevant electronic states near the \( K \)-points with
\( k \sim \xi K \), the dominant \( t_{\alpha\beta}(q) \)-terms that enter the continuum model for valley \( \xi \) involve only 3 different reciprocal
lattice vectors \( G = 0, \xi G_2, \xi G_3 \) for the top layer (\( G' = 0, \xi G_2', \xi G_3' \) accordingly for the bottom layer). Thus, the
Effective inter-layer tunneling Hamiltonian can be written as:

\[ H_{T, eff}^\xi = -\frac{1}{\Omega} \sum_k \sum_{\alpha,\beta}^4 c_{k,1,\alpha}^\dagger t_{\alpha\beta}(\xi K) c_{k,2,\beta} + c_{k,1,\alpha}^\dagger t_{\alpha\beta}(\xi K + \xi G_{2}) e^{-i\xi G_{2}^\dagger d_0} c_{k+\xi G_{M,2},\beta} \\
+ c_{k,1,\alpha}^\dagger t_{\alpha\beta}(\xi K + \xi G_{3}) e^{-i\xi G_{3}^\dagger d_0} c_{k+\xi G_{M,3},\beta} + h.c. \\
+ \sum_{k,\alpha,\beta} w_{\alpha\beta} e^{-i\xi G_{2}^\dagger d_0} c_{k+\xi G_{M,3},\beta} + h.c. \]

where \( w_{\alpha\beta} = t_{\alpha\beta}(K)/\Omega \), \( G_{M,2} = G_{j} - G_{j}' \). Explicitly, the tunneling matrices connecting states at \( k \) on layer 1 to those at \( k, k + \xi G_{M,2}, k + \xi G_{M,3} \) on layer 2 are listed below:

\[ T_{\alpha\beta}(k, k) = -w_{\alpha\beta}, T_{\alpha\beta}(k, k + \xi G_{M,2}) = -e^{-i\xi G_{2}^\dagger d_0} w_{\alpha\beta}, T_{\alpha\beta}(k, k + \xi G_{M,3}) = -e^{-i\xi G_{3}^\dagger d_0} w_{\alpha\beta}. \]

For simplicity, we consider the case with \( d_0 = 0 \) where metal sites from the two layers are aligned at zero twist angle, and assume \( w_{\alpha\beta} = \omega \delta_{\alpha\beta} \) given that the inter-layer tunneling is usually dominated by spin-preserving processes. This leads to the form of the inter-layer tunneling Hamiltonian in Eq.3 of the main text.

To evaluate the effective tunneling strength \( t(q = K) \) near the \( K \)-points, we use the following simple model for \( t(r) \) which incorporates the Slater-Koster \( \sigma \)-bonding between \( d_{2z} \)-orbitals in the two different layers with an empirical exponential decay:

\[ t(r) = |n^2(r) - (l^2(r) + m^2(r))/2|^2 V_{d2d}(r), \]

\[ V_{d2d}(r) = t_0 e^{-\lambda \sqrt{d^2 + |r|^2}}. \]

Here, \( d = 0.7 \text{ nm} \) is the inter-layer distance between two TMD monolayers [S4], \( n(r), l(r), m(r) \) are the directional cosines of the vector \( r \) connecting the two position vectors in the two layers, with

\[ n^2(r) = \frac{d^2}{d^2 + |r|^2}, \quad l^2(r) + m^2(r) = \frac{|r|^2}{d^2 + |r|^2}. \]

The value of \( \lambda \) is estimated based on the empirical fact that \( t' \approx 0.1t \), where \( t' \) and \( t \) are the next-nearest-neighbor and nearest-neighbor hopping amplitudes. Thus, \( t/t' = e^{\lambda a} \approx 10 \), and \( \lambda a = ln(10) \approx 2.3 \). Note that the form of \( t(r) \)
in Eq. S6 is a radial function \( t(r) \equiv t(r) \), and its Fourier transform has the general form of:

\[
\begin{align*}
    t(q) & = \int_0^\infty dr r J_0(qr)t(r) \\
    & = 2\pi \int_0^\infty dr r J_0(qr)\int_0^{2\pi} d\theta e^{-iqr\cos \theta} \\
    & = 2\pi \int_0^\infty dr r J_0(qr) t(r),
\end{align*}
\]

where \( J_0(x) \) is the Bessel function of order zero. With the form of \( t(r) \) in Eq. S6, \( t(q) \) can be obtained by numerical integration with the only fitting parameter \( t_0 \). To fix \( t_0 \), we note that states near the \( \Gamma \)-point of the valence band states in MoS\(_2\) also originate from 4\( d_{z^2}\)-orbitals [S3], and the inter-layer coupling causes an energy splitting of \( \Delta E \approx 2t(q = 0) \approx 800 \text{ meV} \) at the \( \Gamma \)-point in the valence bands in aligned bilayer MoS\(_2\) [S5]. This implies \( t(q = 0) \approx 400 \text{ meV} \) and allows us to fix \( t_0 \) and obtain the values of \( t(q) \) for all \( q \) as shown in Fig. S1. The effective interlayer coupling strength at the \( K \)-point in our continuum model is then extrapolated to be \( w_0 \approx 10 \text{ meV} \) at \( qa = Ka = 4\pi/3 \) (dashed vertical line in black in Fig. S1).

II. Chern numbers in moiré bands

A. Numerical method for computing Chern numbers

The Chern numbers of the CB moiré bands presented in Fig. 2 and Fig. 4 of the main text are calculated numerically using the standard formula introduced in Ref. [S6]. The procedures of the numerical method is outlined below. For the sake of simplicity, we consider a given CB moiré band and drop the band index in the following.

First, we construct a discretized moiré Brillouin zone (MBZ) with an \( N_x \times N_y \) grid. For each momentum point \( \mathbf{k} \) in the grid, we evaluate the following complex quantities:

\[
\begin{align*}
    U_x(\mathbf{k}) & = \langle u(\mathbf{k} + \Delta k_x \hat{x})|u(\mathbf{k}) \rangle, \\
    U_y(\mathbf{k}) & = \langle u(\mathbf{k} + \Delta k_y \hat{y})|u(\mathbf{k}) \rangle,
\end{align*}
\]

where \( |u(\mathbf{k})\rangle \) is the periodic part of the Bloch state at \( \mathbf{k} \). \( \hat{x}, \hat{y} \) are unit vectors in \( k_x \) and \( k_y \) directions, \( \Delta k_x \) and \( \Delta k_y \) measure the spacings between neighboring grid points in \( k_x \) and \( k_y \) directions. Note that the phase factors of \( U_x \) and \( U_y \) essentially account for the Berry connections on each link between two neighboring grid points. The gauge-invariant Berry flux threading through one plaquette around \( \mathbf{k} \) is given by \( \gamma_\mathbf{k} = \arg W(\mathbf{k}) \), where \( W(\mathbf{k}) \) is the Wilson loop:

\[
W(\mathbf{k}) = U_x(\mathbf{k})U_y(\mathbf{k} + \Delta k_x \hat{x})U_y^*(\mathbf{k} + \Delta k_y \hat{y})U_y^*(\mathbf{k}).
\]

The Chern number is then expressed as the sum over all Berry fluxes \( \gamma_\mathbf{k} \) through a total number of \( N_x \times N_y \) plaquettes:

\[
\mathcal{C} = \frac{1}{2\pi} \sum_\mathbf{k} \gamma_\mathbf{k}.
\]

We note that to obtain accurate results for Chern numbers in the continuum model, the total number of grid points must be large enough such that \( |\arg W(\mathbf{k})| < \pi \) for all \( \mathbf{k} \), and the total number \( M \) of the moiré Brillouin zones must also be large enough to ensure an accurate description of the low-energy physics at the exact \( M \to \infty \) limit. In our model, we find that \( N_x = N_y = 40 \) and \( M = 81 \) produce accurate numerical results for \( \mathcal{C} \) in the CB moiré bands.

B. Effects of Berry phase in monolayer transition-metal dichalcogenides

It is known that a Berry phase with value close to \( \pi \) can appear in a local region around \( K \)-points in the Brillouin zone of a monolayer transition-metal dichalcogenide (TMD) [S7]. This Berry phase originates from the massive Dirac Hamiltonian involving both the conduction and valence band edges at \( +K \)-point (see black dashed rectangles in Fig. S2), which is essentially different from both the Berry curvatures generated by the layer degrees of freedom in VB moiré bands [S2] (green dashed rectangles in Fig. S2)) and the giant Berry curvatures (Fig. 2(c) of the main text) generated by the spin and layer degrees of freedom in our continuum model (pink dashed rectangles in Fig. S2)).
While by summing over both the spin-up and spin-down valence bands in a monolayer TMD at one of the $K$-valleys, one acquires a total Berry phase close to $2\pi$ and appears to correspond to a “Chern number” of 1 for this valley, such a “Chern number” is, strictly speaking, not well-defined. This is because the Berry curvatures in a monolayer are summed over two bands and integrated only over a local region around $+K$ in the (large) monolayer Brillouin zone. The integral is not guaranteed to be quantized by any topological reasons. In contrast, the Chern numbers in both the CB moiré bands in the main text and the VB moiré bands in Ref. [S2] are well-defined quantities because the Berry curvatures of each single band are integrated over the entire (small) moiré Brillouin zone. The resultant Chern numbers of each band from valley $\xi = +$ or $\xi = -$ are guaranteed to be quantized by topology.

Importantly, we note that the Berry phase in monolayer TMDs do not affect the topology of CB moiré bands in a twisted homobilayer. This is because the Berry curvatures from the massive Dirac Hamiltonian is only on the order of $\sim 10^{3} \text{Å}^2$ [S7], three orders of magnitude smaller than the Berry curvatures $\sim 10^{4} \text{Å}^2$ in the moiré Chern bands (Fig.2(c) of the main text) generated by spin and layer degrees of freedom. Thus, the topology of CB moiré bands is dictated by the spin and layer degrees of freedom, which justifies why the contributions from remote valence bands ($\sim 1.6 \text{eV}$ away from conduction band states) can be neglected.

III. Self-consistent Hartree-Fock calculations

A. Hartree-Fock energy functional and Lagrange equations

When the Chern bands in a twisted bilayer MoS$_2$ are half-filled, the low-energy effective Hamiltonian is written as:

$$\mathcal{H}_{\text{eff}} = \mathcal{H}_{0,\text{eff}} + \mathcal{H}_{I,\text{eff}}$$  \hspace{1cm} (S12)

where the non-interacting Hamiltonian is given by

$$\mathcal{H}_{0,\text{eff}} = \sum_{\mathbf{k}} \sum_{\xi = \pm} E_{\xi}(\mathbf{k}) c_{\xi}^\dagger(\mathbf{k}) c_{\xi}(\mathbf{k})$$  \hspace{1cm} (S13)
Here, $\xi = \pm$ is the valley index, $E_\xi(k)$ is the energy of moiré band from valley $\xi$. The effective interacting Hamiltonian is given by:

$$\mathcal{H}_{t, eff} = \sum_{k, k', q} \frac{V(q)}{A} \sum_{\xi' \xi} \Lambda^\xi(k + q, k) \Lambda^{\xi'}(k' - q, k')$$

$$\times \xi^c(k + q) c^\dagger_{\xi'}(k' - q) c_{\xi'}(k) c_{\xi}(k),$$

where $V(q) = \int d^2r V(r) e^{iq \cdot r}$ is the Fourier transform of the two-body interaction $V(r_1 - r_2)$, $A$ is the total area of the system, $\Lambda^\xi(k + q, k) \equiv \langle u_{\xi + q, k} | u_{\xi, k} \rangle$ are the form factors arising from projecting the general density-density interaction to the active bands, which end up as the overlap between the cell-periodic parts of Bloch states within the same band from valley $\xi$. Note that due to the $U_\xi(1)$-symmetry, wave functions from different valleys do not overlap: $\Lambda^{+ -}(k', k) \equiv \langle u_{+, k'} | u_{-, k} \rangle = 0$.

As mentioned in the main text, the Coulomb energy scale $g_C \sim 10$ meV in twisted bilayer MoS$_2$ overwhelms the band width $W \sim 1$ meV in the Chern bands found in the continuum model. This motivates us to consider possible ground states at 1/2-filling as a general Hartree-Fock state of the form: $|\Phi\rangle = \prod_k |w_+(k)e^{i\theta_+(k)}c^+_k(k) + w_-(k)e^{i\theta_-(k)}c^-_k(k)\rangle |\Phi_0\rangle$, where $w_\xi(k)$ and $\theta_\xi(k)$ are real numbers and denote the amplitude and phase of the coefficients of valley $\xi$ at momentum $k$. $|\Phi_0\rangle$ is the vacuum state with all bands below the active bands being fully filled. The variational parameters can be determined by minimizing the ground state energy subject to the normalization condition $w^2_+(k) + w^2_-(k) = 1$. The energy of $|\Phi\rangle$ is given by:

$$E^\Phi \equiv \langle \Phi | \mathcal{H}_{t, eff} | \Phi \rangle = E^\Phi_0 + E_H + E^\Phi_F,$$

where $E^\Phi_0$, $E_H$ and $E^\Phi_F$ refer to the kinetic energy, Hartree energy, and the Fock exchange energy of $|\Phi_0\rangle$, respectively. It is straightforward to show that the Hartree term essentially accounts for the direct coupling: $E_H \propto n_e^2$, where $n_e$ is the electron density of the system, which is fixed at a given band filling. Thus, we neglect $E_H$ from now on and focus on $E^\Phi_0$ and $E^\Phi_F$:

$$E^\Phi_0 = \sum_k w^2_+(k)E_+(k) + w^2_-(k)E_-(k),$$

$$E^\Phi_F = -\sum_{k, q} \frac{V(q)}{A} \sum_{\xi \xi'} \Lambda^\xi(k + q, k) \Lambda^{\xi'}(k, k + q) w_\xi(k + q) w^{\xi'}(k) w^{\xi}(k)$$

$$\times e^{-i\theta_\xi(k + q)} e^{i\theta_{\xi'}(k + q)} e^{-i\theta_{\xi}(k)} e^{i\theta_{\xi'}(k)}.$$

Note that in order to minimize $E^\Phi_F$, the phase factors would adjust themselves such that $e^{-i\theta_{\xi'}(k + q)} e^{i\theta_{\xi}(k)} \Lambda^\xi(k + q, k) = [\Lambda^\xi(k, q, k)]$. This reduces $E^\Phi_F$ to the following form:

$$E^\Phi_F = -\sum_{k, q} \frac{V(q)}{A} \sum_{\xi \xi'} |\Lambda^\xi(k + q, k)|^2 |\Lambda^{\xi'}(k, k + q)| w_\xi(k + q) w^{\xi'}(k) w^{\xi}(k)$$

where we have used the identity $|\Lambda^\xi(k, q, k)| = |\Lambda^{\xi'}(k, q, k)|^* = |\Lambda^\xi(k + q, k)|$. To minimize $E^\Phi_0 + E^\Phi_F$ subject to the constraint $w^2_+(k) + w^2_-(k) = 1$, we introduce the Langrange multipliers $\lambda_k$ such that

$$\partial_{w_\xi(k)} \left[E^\Phi_0 + E^\Phi_F - \sum_k \lambda_k (w^2_+(k) + w^2_-(k))\right] = 0.$$

This leads to the following two equations:

$$\left\{ [E_+(k) - \sum_q g_+(q, k) w^2_+(k + q)] w_+(k) - [\sum_q g_{+-}(q, k) w_+(k + q)] w_-(k + q)] w_-(k) = \lambda_k w_+(k) \right\}$$

$$\left\{ [E_-(k) - \sum_q g_-(q, k) w^2_-(k + q)] w_-(k) - [\sum_q g_{-+}(q, k) w_-(k + q)] w_+(k + q)] w_+(k) = \lambda_k w_-(k) \right\}$$

where we introduced the shorthand notation:

$$\left\{ g_\xi(q, k) = V(q) |\Lambda^\xi(k + q, k)|^2 / A, (\xi = \pm) \right\}$$

$$\left\{ g_{+-}(q, k) = V(q) |\Lambda^+(k + q, k)||\Lambda^-(k + q, k)| / A. \right\}$$
As we seek for nontrivial solutions of Eq.S19, it is convenient to introduce the following two quantities:

\[
\Delta_{\pm}(k) = \sum_{q} g_{\pm}(q, k) w_{\pm}^{2}(k + q),
\]

(S21)

\[
\Delta_{\pm}(k) = \sum_{q} g_{\pm}(q, k) w_{\pm}(k + q) w_{\mp}(k + q),
\]

which allows us to reformulate Eq.S19 as an eigenvalue equation:

\[
\begin{pmatrix}
E_{+}(k) - \Delta_{+}(k) & -\Delta_{+-}(k) \\
-\Delta_{+-}(k) & E_{-}(k) - \Delta_{-}(k)
\end{pmatrix}
\begin{pmatrix}
w_{+}(k) \\
w_{-}(k)
\end{pmatrix} = \lambda_{k}
\begin{pmatrix}
w_{+}(k) \\
w_{-}(k)
\end{pmatrix}.
\]

(S22)

Note that \(\lambda_{k}\) has the physical interpretation as the energy contribution from \(k\) to the total energy of the Hartree-Fock state. This motivates us to focus on the lower branch of eigenvalues of the \(2 \times 2\) symmetric matrix:

\[
\lambda_{1}(k) = \frac{1}{2} [E_{+}(k) - E_{-}(k) - (\Delta_{+}(k) + \Delta_{-}(k))] - \sqrt{\Delta_{+-}^{2}(k) + \frac{1}{4} [E_{+}(k) - E_{-}(k) - (\Delta_{+}(k) - \Delta_{-}(k))]^{2}}
\]

with the corresponding eigenvector \([w_{+}(k), w_{-}(k)]^{T} = [\cos(\phi_{k}/2), \sin(\phi_{k}/2)]^{T}\), where

\[
\sin(\phi_{k}) = \frac{\Delta_{+-}(k)}{\sqrt{\Delta_{+-}^{2}(k) + \frac{1}{4} [E_{+}(k) - E_{-}(k) - (\Delta_{+}(k) - \Delta_{-}(k))]^{2}}} = 2 \cos(\phi_{k}/2) \sin(\phi_{k}/2),
\]

(S24)

\[
\cos(\phi_{k}) = \frac{1}{2} \sqrt{\Delta_{+-}^{2}(k) + \frac{1}{4} [E_{+}(k) - E_{-}(k) - (\Delta_{+}(k) - \Delta_{-}(k))]^{2}} = \cos^{2}(\phi_{k}/2) - \sin^{2}(\phi_{k}/2).
\]

The defining equations (Eq.S21) require that \(\Delta_{+}(k), \Delta_{-}(k), \Delta_{+-}(k)\) satisfy the following self-consistent equations:

\[
\Delta_{+}(k) = \frac{1}{N} \sum_{k'} J_{+}(k', k) \frac{1 + \cos(\phi_{k'})}{2},
\]

(S25)

\[
\Delta_{-}(k) = \frac{1}{N} \sum_{k'} J_{-}(k', k) \frac{1 - \cos(\phi_{k'})}{2},
\]

\[
\Delta_{+-}(k) = \frac{1}{N} \sum_{k'} J_{+-}(k', k) \frac{\sin(\phi_{k'})}{2}.
\]

Here, \(J_{\xi\xi'}(k, k')\) are the Fock exchange functions:

\[
J_{\xi\xi'}(k', k) = \Omega^{-1} \sum_{G_M} V(k' + G_M - k)|\Delta_{\xi}(k' + G_M, k)||\Delta_{\xi}(k, k' + G_M)|,
\]

(S26)

where \(\Omega\) is the area of the real-space moiré unit cell, \(G_M\) denotes the moiré reciprocal lattice vectors.

**B. Mean-field approximation**

To simplify Eq.S25, we employ the mean-field approximation which replaces the exchange coupling strengths \(J_{\xi\xi'}(k', k)\) by their averages: intravalley exchange coupling \(J_0 \equiv \langle J_{+}(k', k) \rangle = \langle J_{-}(k', k) \rangle\), and the intervalley exchange coupling \(J_1 \equiv \langle J_{+-}(k', k) \rangle\). This reduces Eq.S25 to a coupled set of self-consistent equations involving only constant order parameters:

\[
\Delta_{+} = \frac{J_0}{N} \sum_{k} \frac{1 + \cos(\phi_{k})}{2},
\]

(S27)

\[
\Delta_{-} = \frac{J_0}{N} \sum_{k} \frac{1 - \cos(\phi_{k})}{2},
\]

\[
\Delta_{+-} = \frac{J_1}{N} \sum_{k} \frac{\sin(\phi_{k})}{2},
\]

where \(J_0\) and \(J_1\) denote the intravalley and intervalley exchange couplings, respectively.
where

\[
\sin(\phi_k) = \frac{\Delta_+}{\sqrt{\Delta_+^2 + \frac{1}{4}[E_+(k) - E_-(k) - (\Delta_+ - \Delta_-)^2]}},
\]

\[
\cos(\phi_k) = \frac{1}{2} \frac{(\Delta_+ - \Delta_-) - (E_+(k) - E_-(k))}{\sqrt{\Delta_+^2 + \frac{1}{4}[E_+(k) - E_-(k) - (\Delta_+ - \Delta_-)^2]}}.
\]

By adding and subtracting the first two equations in Eq. S27, we end up with a compact form (Eq.5 of the main text):

\[
\Delta_+ + \Delta_- = J_0,
\]

\[
\Delta_+ - \Delta_- = \frac{J_0}{2N} \sum_k \frac{(\Delta_+ - \Delta_-) - \delta(k)}{D(k)},
\]

\[
\Delta_{12} = \frac{J_1}{2N} \sum_k \frac{\Delta_{12}}{D(k)}
\]

Here, \(\delta(k) = E_+(k) - E_-(k), D(k) = \sqrt{\Delta_+^2 + \frac{1}{4}[\delta(k) - (\Delta_+ - \Delta_-)]^2}\). Note that time-reversal enforces \(\delta(-k) = E_+(k) - E_-(k)\) to be an odd function of \(k\).

C. Hartree-Fock ground states under different regimes of mean-field exchange coupling \((J_0, J_1)\)

It is worth noting that there exist two possible solutions for the last equation in Eq. S29, which divides the whole set of solutions to Eq. S29 into two different classes: (i) \(\Delta_{12} = 0\), and (ii) \(\Delta_{12} \neq 0\). Particularly, in case (i) where \(\Delta_{12} = 0\), we have \(D(k) = \frac{1}{2}\delta(k) - (\Delta_+ - \Delta_-)\) for all \(k\) (or equivalently, either \(\Delta_+ = J_0, \Delta_- = 0\) or \(\Delta_+ = 0, \Delta_- = J_0\)). These two degenerate states are the valley-polarized (VP) solutions which spontaneously break time-reversal symmetry. The energy of the two degenerate VP states is given by

\[
E_{VP} = \sum_k E_+(k) - NJ_0.
\]

In case (ii) where \(\Delta_{12} \neq 0\), one can divide both sides of the 3rd equation in Eq. S29 by \(\Delta_{12}\), which simplifies the set of equations as below:

\[
\Delta_+ + \Delta_- = J_0,
\]

\[
\Delta_+ - \Delta_- = \frac{J_0}{2N} \sum_k \frac{\delta(k)}{D(k)},
\]

\[
1 = \frac{J_1}{2N} \sum_k \frac{1}{D(k)},
\]

where the three unknown mean-field order parameters \(\Delta_+, \Delta_-, \Delta_{12}\) that characterize this state need to be solved self-consistently. In particular, we note that \(\Delta_{12} \neq 0\) implies a nontrivial inter-valley order associated with this Hartree-Fock state, in which the microscopic valley is no longer conserved (i.e., spontaneous breaking of valley-\(U_v(1)\) symmetry). On the other hand, any solution with nonzero \(\Delta_M \equiv \Delta_+ - \Delta_-\) necessarily lifts the valley degeneracy and breaks time-reversal symmetry spontaneously. Thus, we call \(\Delta_M\) as the valley magnetic order. As it turns out, under time-reversal symmetry (e.g., in the absence of any external magnetic fields), the only \(U_v(1)\)-breaking solution must satisfy the condition \(\Delta_+ = \Delta_-\), i.e., there is no valley magnetic order \(\Delta_M = \Delta_+ - \Delta_-\) involved. This solution is known as the (generalized) inter-valley coherence (IVC) state [S8] (note: \(w_+(k) = w_-(k) \neq 1/\sqrt{2}\) in general except at time-reversal-invariant momenta). According to Eq. S23, the energy of the IVC state is:

\[
E_{IVC} = \sum_k \frac{1}{2} [E_+(k) + E_-(k) - (\Delta_+ + \Delta_-)] - D(k)
\]

\[
[E_-(k) = E_+(k), \Delta_+ + \Delta_- = J_0] = \sum_k [E_+(k) - D(k)] - \frac{N}{2} J_0.
\]
The ground state can then be determined by comparing the two energies $E_{IVC}$ and $E_{VP}$:

$$\Delta E = E_{IVC} - E_{VP} = \frac{N}{2} J_0 - \sum_k D(k),$$  \hspace{1cm} (S33)

where $\Delta_+, \Delta_-, \Delta_{+}$ involved in $D(k)$ are given by the self-consistent solutions of Eq. S31. The ground states obtained for each pair of $(J_0, J_1)$ map out a $J_0 - J_1$ mean-field phase diagram.

To develop some physical insights into the energetics of the system, we first consider some simple limits. First, we note that the flat band condition implies $J_0, J_1 \gg |\delta(k)|$. This combined with the self-consistency condition $1 = \frac{J_1}{2N} \sum_k \frac{1}{D(k)}$ suggests $D(k) \sim J_1/2$. Thus, when the intravalley exchange coupling dominates over its intervalley counterpart: $J_0 \gg J_1$, we have $\Delta E \sim N(J_0 - J_1)/2 \gg 0$. In this case, the VP state is favored. By a similar argument, when the intervalley exchange coupling dominates: $J_1 \gg J_0$, $\Delta E \sim N(J_0 - J_1)/2 \ll 0$ and the IVC state always wins. In particular, in the entire $J_1 \geq J_0$ regime, using the harmonic mean-arithmetic mean (HM-AM) inequality:

$$\frac{N}{\sum_{i=1}^N \frac{1}{x_i}} \leq \sum_{i=1}^N \frac{x_i}{N}, (x_i > 0, \forall i = 1, ..., N)$$  \hspace{1cm} (S34)

one can show that

$$\sum_k D(k) \geq \frac{N^2}{\sum_k \frac{1}{D(k)}} = \frac{N}{2} J_1,$$  \hspace{1cm} (S35)

where we used the last equation in Eq. S31. This shows that for $J_1 \geq J_0$, $\Delta E \leq N(J_0 - J_1)/2 \leq 0$ always holds, and the system favors the IVC state with nonzero $\Delta_+$ and $\Delta_M = \Delta_+ - \Delta_- = 0$.

Next, we consider the $J_0 > J_1$ regime. As we discussed above, at $J_0 = J_1$, the system always favors the IVC state, while in the limit $J_0 \gg J_1$ the system should favor the VP state. Thus, that upon increasing the ratio $J_0/J_1$, a phase transition must happen somewhere in the $J_0 > J_1$ regime. By solving the self-consistent equations (Eq. S29) numerically, we obtain the $J_0 - J_1$ phase diagram in Fig.3 of the main text, which reveals that the VP state is favored almost immediately when the $J_0 > J_1$ regime is accessed, and a first-order transition from IVC to VP phase occurs at $J_0 \simeq J_1$. This is due to the fact that the energetics of the system is essentially governed by $(J_0, J_1)$ under the condition $J_0, J_1 \gg |\delta(k)|$. As a result, $\Delta E \simeq N(J_0 - J_1)/2$ almost holds all the time, and the VP state is favored almost in the entire regime with $J_0 > J_1$.

D. Physically accessible regimes of $(J_0, J_1)$ under density-density repulsions

We note that the discussions above are based on pure algebraic analysis — it is clearly desirable to examine the accessible regimes of $(J_0, J_1)$ on physical grounds. Indeed, given a generic density-density repulsive interaction which respects $U_e(1)$ and $T$, the relation between $J_0$ and $J_1$ can be derived. In particular, the intravalley exchange coupling strength is

$$J_0 \equiv \langle J_{++}(k', k) \rangle \hspace{1cm} (S36)$$

$$= \frac{1}{N^2 \Omega} \sum_{k,k'} \sum_{G_M} V(k' + G_M - k) \Lambda^+(k' + G_M, k)^2$$

$$= \frac{1}{N^2 \Omega} \sum_{k,k'} \sum_{G_M} V(-k' - G_M + k) \Lambda^-(-k, -k' - G_M)^2$$

$$= \frac{1}{N^2 \Omega} \sum_{k,k'} \sum_{G_M} V(k' + G_M - k) \Lambda^-(k' + G_M, k)^2$$

$$= \langle J_{--}(k', k) \rangle.$$
Note that in the equations above we used the fact that $V(q) = V(-q)$ and the condition $\Lambda^+(k', k) = \Lambda^-(k', -k')$ imposed by time-reversal symmetry. For the inter-valley exchange coupling strength,\
\[
J_1 \equiv \langle J_+ - (k', k) \rangle = \frac{1}{N^2 \Omega} \sum_{k, k'} \sum_{G_M} V(k' + G_M - k) |\Lambda^+(k' + G_M, k)| |\Lambda^-(k, k' + G_M)| \\
\leq \frac{1}{N^2 \Omega} \sum_{k, k'} \sum_{G_M} V(k' + G_M - k) |\Lambda^+(k' + G_M, k)|^2 + |\Lambda^-(k' - G_M, -k)|^2
\]
\[
[V(q) = V(-q)] = \frac{1}{2N^2 \Omega} \left[ \left( \sum_{k, k'} \sum_{G_M} V(k' + G_M - k) |\Lambda^+(k' + G_M, k)|^2 \right) \right.
\]
\[
+ \left. \left( \sum_{k, k'} \sum_{G_M} V(-k' - G_M + k) |\Lambda^+(k' - G_M, -k)|^2 \right) \right]
\]
\[
= \frac{1}{2} 2 \langle J_+ (k', k) \rangle
\]
\[
= J_0.
\]

Therefore, we identify the physically accessible points $(J_0, J_1)$ are located in the $J_0 \geq J_1$ regime of the phase diagram. Notably, the relation $J_0 \geq J_1$ derived above holds regardless of details of the wave functions and the microscopic interaction $V(q)$ as long as the interaction is repulsive in nature. Thus, as shown in Fig.3 of the main text, the VP state takes up a large portion of the physical regime with $J_0 \geq J_1$ except a narrow strip with $J_0 \approx J_1$ where the IVC state is favored.

**E. Case of general filling away from 1/2**

Here, we briefly discuss the physics for general fillings away from 1/2. For general fillings above 1/2, one valley is fully filled, while the other is partially filled. On the other hand, for fillings below 1/2, one valley is empty while the other is partially filled. In both cases, the system is metallic and one expects anomalous Hall effects with non-quantized Hall conductance due to the large Berry curvatures hosted in the Chern bands (Fig.2(c) of the main text).

Moreover, at full band filling (i.e., both valleys being fully filled), we expect time-reversal symmetry to be restored and the opposite Chern numbers in opposite valleys give rise to a pair of counter-propagating edge states. This effect can be termed as the **quantum valley Hall effect**, analogous to the quantum spin Hall effect in which opposite Chern numbers from opposite spin species give rise to a pair of counter-propagating edge states.

**F. Correlated topological insulating gap at 1/2-filling**

As we mentioned in the Conclusion and Discussions section of the main text, under the flat band condition $W \ll g_C$ ($W$: band width, $g_C$: interaction strength) at small twist angles $\theta_0 \sim 1^\circ$, the actual topological insulating gap to be manifested experimentally at 1/2-filling is not solely determined by the gap induced by Rashba SOCs in the non-interacting bands. Instead, it is largely governed by the correlation-induced insulating gap for the valley-polarized (VP) state, which is approximately given by the intra-valley exchange coupling strength $J_0 \sim g_C \approx 10$ meVs (see estimation of $g_C$ in the section “Correlated QAH state at 1/2-filling” in the main text).

Physically, for a VP state, the intra-valley exchange interactions among electrons in the fully filled moiré band lowers the energy of each individual electron in this valley by an amount of $J_0$. On the other hand, the energy of the empty band in the other valley remains unaffected, as electron correlation cannot occur within an empty band. As a result, once the non-interacting bands acquire nonzero Chern numbers from any finite Rashba SOC, the correlated topological gap at 1/2-filling would be on the order of $J_0 = 10$ meVs. To explicitly demonstrate this, we use our Hartree-Fock formalism (see Eq.30) to calculate the correlated band structures for both valley $\xi = +$ and valley $\xi = -$ when the pair of 2nd moiré bands are half-filled (Fig.S3). As the spontaneous $T$-broken VP phase selects one out of the two valleys randomly, we assume without loss of generality that the 2nd moiré band from $\xi = +$ being fully filled, while its counterpart from $\xi = -$ is empty. Clearly, the topological insulating gap (indicated by double arrow in Fig.S3(c)-(d)) corresponds to the scale of $J_0$, instead of the gap induced by Rashba SOC in the non-interacting bands in Fig.S3(a)-(b).
FIG. S3: (a)-(b) Non-interacting bands of twisted homobilayer MoS2 for (a) valley $\xi = +$, and (b) valley $\xi = -$ with Rashba strength $\alpha_{so} = 30$ meV $\cdot \text{Å}$. Dashed lines in (a)-(b) indicate location of the Fermi energy $E_F$ when the 2nd moiré bands are 1/2-filled and interaction effects are neglected. In the absence of correlations, both bands from valley $\xi = +$ and valley $\xi = -$ are half-filled. (c)-(d) Correlated band structures when the 2nd moiré bands are half-filled, where the system enters the valley-polarized state. Electron correlations induce an insulating gap $\Delta_{CQAH} \sim J_0$, where we set $J_0 = 10$ meVs.

IV. Estimate of spatial variation in Rashba SOC strength

As we discussed in the Conclusion and Discussion section in the main text, the formation of moiré pattern in principle leads to spatial variations of the Rashba SOC strength on the moiré scale (the amplitude of variations denoted by $\delta \alpha_{so}$ in the following).

To estimate $\delta \alpha_{so}$, the key observation is that the spatial variation in Rashba strength is not independent of the moiré potential. In particular, the moiré potential causes a spatial variation in the inter-layer potential differences between the two constituent layers with its amplitude on the scale of $V_M \approx 10$ meVs (Table I of the main text). Given the inter-layer distance $d = 7 \text{Å}$ in bilayer MoS$_2$, this corresponds to a spatially modulating electric field with the amplitude of variations given by $\delta E_z = 1.4 \text{mV} \cdot \text{Å}^{-1}$. Assuming a linear scaling relation $\delta \alpha_{so} = \lambda \delta E_z$, we obtain an estimated value of $\delta \alpha_{so} = 0.08 \text{meV} \cdot \text{Å}$ by using a proportionality constant $\lambda = 0.0625 \text{e} \cdot \text{Å}^2$ for MoS$_2$ extrapolated from Fig.7 of Refs.[S9]. Thus, the amplitude of the spatial variation in Rashba SOC is negligibly small compared to previous reports on gate-induced Rashba SOC with $\alpha_{so} \sim 30 \text{meV} \cdot \text{Å}$ in each layer [S10]. This justifies the use of a uniform Rashba SOC in our continuum model as a good approximation.

* Corresponding author: benjamin.zhou@ubc.ca
† Corresponding author: franz@phas.ubc.ca
[S1] Rafi Bistritzer and Allan H. MacDonald, PNAS 108 (30) 12233-12237 (2011).
[S2] Fengcheng Wu, Timothy Lovorn, Emanuel Tutuc, Ivar Martin and A. H. MacDonald, Phys. Rev. Lett. 122, 086402 (2019).
[S3] Gui-Bin Liu, Wen-Yu Shan, Yugui Yao, Wang Yao, and Di Xiao, Phys. Rev. B 88, 085433 (2013).
[S4] Kowsalya Devi Rasamani, Farbod Alimohammadi, Yugang Sun, Materials Today 20, 83-91, (2017).
[S5] Ferdows Zahid, Lei Liu, Yu Zhu, Jian Wang, and Hong Guo, AIP Advances 3, 052111 (2013).
[S6] Takahiro Fukui, Yasuhiro Hatsugai, and Hiroshi Suzuki, J. Phys. Soc. Jpn. 74, 1674 (2005).
[S7] Di Xiao, Gui-Bin Liu, Wanxiang Feng, Xiaodong Xu, and Wang Yao, Phys. Rev. Lett. 108, 196802 (2012).
[S8] Po, H. C., Zou, L., Vishwanath, A. & Senthil, T. Phys. Rev. X 8, 031089 (2018).
[S9] Qun-Fang Yao, Jia Cai, Wen-Yi Tong, Shi-Jing Gong, Ji-Qing Wang, Xiangang Wan, Chun-Gang Duan, and J. H. Chu, Phys. Rev. B 95, 165401 (2017).
[S10] Benjamin T. Zhou, Katsuhisa Taguchi, Yuki Kawaguchi, Yukio Tanaka & K. T. Law, Spin-orbit coupling induced valley Hall effects in transition-metal dichalcogenides, Communications Physics 2, 26 (2019).