Layer pseudospin skyrmions and topological moiré bands in twisted homobilayers

Fengcheng Wu,1 Timothy Lovorn,2 Emanuel Tutuc,3 Ivar Martin,1 and A. H. MacDonald2

1Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA
2Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA
3Department of Electrical and Computer Engineering, Microelectronics Research Center, The University of Texas at Austin, Austin, Texas 78758, USA

We show that moiré bands of homobilayer van der Waals materials can be topologically nontrivial, and illustrate the tendency by studying valence band states in the ±K valleys of twisted semiconducting transition metal dichalcogenides (TMDs). Because of the large spin-orbit splitting at their monolayer valence band maxima, the low energy bands of TMD homobilayers can be described using a two-band model with a layer-pseudospin magnetic field $\Delta(r)$ that has the periodicity of the moiré pattern. We show that for balanced bilayers $\Delta(r)$ has a topologically non-trivial skyrmion lattice texture in real space, and that the topmost moiré valence bands provide an accurate realization of the Kane-Mele quantum spin-Hall model. When the bilayer is polarized by a vertical displacement potential, the band Chern numbers are driven to zero before $\Delta(r)$ becomes topologically trivial. Because the bands narrow at small twist angles, a rich set of broken symmetry insulating states can occur when the number of electrons per moiré unit cell is an integer.

Introduction.— Moiré superlattices form in van der Waals bilayers with small differences between the lattice constants or orientations of the individual layers, and often dramatically alter electronic properties [11-15]. In the presence of long-period moiré patterns, electronic states can be described by continuum model Hamiltonian with the moiré periodicity and spinors whose dimension is equal to the total number of bands, summed over layers, in the energy range of interest. Application of Bloch’s theorem then gives rise to moiré bands [6]. Because the moiré pattern often generates spatial confinement, moiré bands can be narrow, enhancing the importance of electronic correlations. The flat bands of magic-angle twisted bilayer graphene, in which correlated insulating and superconducting states have been discovered [11-13], provide a prominent example. The study of moiré flat bands has recently become an active area of experimental and theoretical research centered on efforts to identify promising bilayer structures, and on topological characterization and many-body interaction physics [14-16].

When the two layers are formed from the same material (homobilayers), both must be treated on an equal footing. The ±K-valley valence bands of semiconductor TMD monolayers are particularly simple because strong spin-orbit coupling and broken inversion symmetry lifts spin-degeneracy [17], and the corresponding homobilayer can be described by a two-band model with layer pseudospins. The moiré pattern’s periodic modulation can then be accounted for by a scalar potential and a pseudo magnetic field $\Delta(r)$ whose components are the coefficients of the layer Pauli matrix expansion of the two-band Hamiltonian, i.e. $\Delta_x$ and $\Delta_y$ are the real and imaginary parts of the interlayer tunneling amplitude and $\Delta_z$ is the potential difference between layers. The field $\Delta(r)$ inherits the moiré pattern periodicity and plays a key role in the discussion below. Focusing for definiteness on the valence bands of WSe$_2$ and on AA stacking [Fig. 1], we find that for balanced bilayers $\Delta(r)$ has a skyrmion lattice texture in real space, and that the the moiré bands carry valley-contrasting Chern numbers. When the bilayer is polarized by a vertical displacement potential, the band Chern numbers are driven to zero before $\Delta(r)$ becomes topologically trivial in real space. In partially filled flat bands, interactions become dominant and can, for example, break time-reversal symmetry to form quantum anomalous Hall states. We anticipate that homobilayer moiré systems will emerge as an attractive platform to study the interplay between topology and correlation.

Aligned bilayers— To derive a moiré continuum Hamiltonian, we start by analyzing the electronic structure of an aligned bilayer [18]. Because the ±K valleys are related by time-reversal symmetry $\mathcal{T}$, we can focus on the +K valley. In an AA stacked TMD homobilayer [Fig. 1], the valence states at the +K valley valence band maximum are mainly of $d_{x^2−y^2}$ + id$_{xy}$ orbital character and spin up (↑) along z axis [19], and are separated from spin-down (↓) states by strong spin-orbit splitting. Retaining only the spin up valence-band states at the +K valley

![FIG. 1. (a) Top view of AA stacked TMD MX$_2$ homobilayer with a small twist angle $\theta$ and an in-plane displacement $d_0$. $M$ (X) stands for metal (chalcogen) atoms. The inset provides a schematic side view. (b) ±K valence bands in an AA stacked homobilayer when interlayer coupling is neglected.](image-url)
yields the two-band $k \cdot p$ Hamiltonian [20],
\[
\mathcal{H}_t(\theta = 0, d_0) = \left( \begin{array}{c}
-\frac{\hbar^2 k^2}{2m^*} + \Delta_b(d_0) \\
\Delta_T(d_0)
\end{array} \right) \left( \begin{array}{c}
\Delta_T(d_0) \\
-\frac{\hbar^2 k^2}{2m^*} + \Delta_t(d_0)
\end{array} \right),
\]
(1)
with parameters that depend on the displacement $d_0$ between the aligned layers. In [1], $b$ and $t$ refer to bottom (b) and top (t) layers, $k$ is momentum measured from +$K$ point, $m^*$ is the valence band effective mass that is approximately independent of $d_0$ [20], $\Delta_b, t$ are layer-dependent energies, and $\Delta_T$ is an inter-layer tunneling amplitude. The dependence of $\Delta_\alpha (\alpha = b, t, T)$ on $d_0$ is constrained by the symmetry properties of the bilayer. The two-dimensional lattice periodicity of the aligned bilayers implies that the $\Delta_\alpha$ are periodic functions of $d_0$. A $z \leftrightarrow -z$ mirror operation interchanges $b$ and $t$ and maps displacement $d_0$ to $-d_0$, implying that $\Delta_t(d_0) = \Delta_b(-d_0)$. Threefold rotation around the $z$ axis requires that $\Delta_b$ and $\Delta_t$ be invariant when $d_0$ is rotated by $2\pi/3$. These symmetry constraints suggest the following two-parameter lowest-harmonic parametrization:
\[
\Delta_\ell(d_0) = 2V \sum_{j=1,3,5} \cos(G_j \cdot d_0 + \ell \psi),
\]
(2)
where $\ell = 1$ for the $b$ layer and $\ell = -1$ for the $t$ layer, $G_j$ is the reciprocal lattice vector obtained by counterclockwise rotation of $G_1 = (4\pi)/(\sqrt{3}a_0)\hat{y}$ by angle $(j-1)\pi/3$, $a_0$ is the monolayer TMD lattice constant, and $V$ and $\psi$ respectively characterize the amplitude and shape of the potentials. Note that we have chosen the spatial averages of $\Delta_b, t$, which must be identical, as the zero of energy.

The $d_0$ dependence of $\Delta_T$ is most conveniently understood by assuming a two-center approximation [3] for tunneling between the metal $d_{x^2-y^2} + id_{xy}$ orbitals, and using a lowest-harmonic approximation. This leads to,
\[
\Delta_T(d_0) = w(1 + e^{-G_2 \cdot d_0} + e^{-iG_3 \cdot d_0}),
\]
(3)
where $w$ is a tunneling strength parameter. It is informative to highlight three high-symmetry displacement values: $d_{0,n} = n(a_1 + a_2)/3$ for $n = 0, \pm 1$, where $a_{1,2}$ are the primitive translation vectors of the aligned bilayer: $a_1 = a_0(1,0)$ and $a_2 = a_0(1/2, \sqrt{3}/2)$. For $n = 0$ the metal atoms of the two layers are aligned, $\Delta_t = \Delta_b = 6V \cos(\psi)$ and $\Delta_T = 3w$; the valence band maximum states are then symmetric and antisymmetric combinations of the isolated layer states. For $n = \pm 1$ the metal atoms in one layer are aligned with the chalcogen atoms in the other layer, and $\Delta_T$ vanishes as a result of the threefold rotational symmetry $C_{3z}$. We determine the model parameters by fitting the eigenvalues of $\mathcal{H}_t(k = 0)$ at the three displacements to corresponding values from fully relativistic band structure calculations performed using the local-density approximation as implemented in Quantum Espresso [21]. We find that
\[
\begin{align*}
\Delta(r) & = (\Delta_x, \Delta_y, \Delta_z) \equiv (\text{Re}\Delta_T, \text{Im}\Delta_T, \frac{\Delta_b - \Delta_t}{2}).
\end{align*}
\]

FIG. 2. (a) Brillouin zones of the bottom (blue hexagon) and top (red hexagon) layers in a twisted bilayer, and the moiré Brillouin zone (black hexagon). (b) The layer pseudospin skyrminn lattice in the moiré pattern. The color map illustrates the periodic spatial pattern of the $\hat{z}$ component of the layer pseudospin magnetic field $\Delta_z$, and the arrows illustrate the in-plane components $\Delta_x, \Delta_y$. The white lines outline a single moiré unit cell. The dots highlight the high symmetry positions $R^M_M, R^X_M$, and $R^Y_M$.

\[\begin{align*}
(V, \psi, w) & \approx (8.9 \text{ meV}, 91°, 9.7 \text{ meV}). \text{ Although these parameter values are used below, the qualitative effects that we discuss are largely generic.}
\end{align*}\]
As illustrated in Fig. 2(b), $\Delta_+(r)$ vanishes along the links that connect nearest-neighbor $R^M_M$ sites and has maximum and minimum values at $R^M_M$ and $R^M_M$. The in-plane pseudospin field, which accounts for interlayer tunneling, has vortex and antivortex structures centered on $R^M_M$ and $R^M_M$. Here $R^M_\alpha$ denote high-symmetry sites at which $\alpha$ atoms of the bottom layer are locally aligned with $\beta$ atoms of the top layer. It follows that $\Delta(r)$ forms a skyrmion lattice, i.e., that the direction of the $\Delta(r)$ covers the unit sphere once in each moiré unit cell (MUC). We have explicitly confirmed this property by numerically evaluating the winding number [24]:

$$N_w \equiv \frac{1}{4\pi} \int_{\text{MUC}} d^3 r \frac{\Delta \cdot (\partial_\ell \Delta \times \partial_\alpha \Delta)}{|\Delta|^3} = 1. \quad (6)$$

Skyrmion lattice pseudospin textures in position space suggest [25] the possibility of topological electronic bands in momentum space, although we will find that the connection is not one-to-one.

**Topological bands.** The moiré band Hamiltonian is readily diagonalized using a plane wave representation. The result for $\theta = 1^\circ$ is illustrated in Fig. 3(a). The $C_{2\gamma}T$ symmetry of the Hamiltonian maps $\kappa_+ \rightarrow \kappa_-$ and therefore enforces degeneracy between these points. For the two topmost moiré bands of the $+K$ valley, wave functions in the $b$ (t) layer are concentrated near the $R^M_M$ ($R^M_M$) sites, which are $\Delta_b$ ($\Delta_+$) maxima. Because of the layer-dependent momentum shifts $\kappa_\pm$ in the kinetic energies, the moiré band wave functions vary rapidly over the MBZ. In particular, the wave function of the topmost moiré band at $\kappa_+$ and $\kappa_-$ are respectively localized in layers $b$ and $t$. By integrating the Berry curvature $F$ over the MBZ [20], we confirm that the Chern numbers $C$ of the two topmost $+K$ valley moiré bands in Fig. 3 are non-trivial ($C = \pm 1$) at $\theta = 1^\circ$. The corresponding bands at the $-K$ valley must have the opposite Chern numbers due to the $T$ symmetry. Spin-valley locking implies that when the chemical potential is in the gap between the two topmost bands, the twisted homobilayer is not only a valley Hall insulator but also a quantum spin Hall insulator [27, 28].

To gain deeper insight into the topological bands, we construct a tight binding model. The real space distribution of the wave functions suggests a two-orbital model for the first two moiré bands:

$$H_{\text{TB}} = \sum_{\ell,s} \sum_{R,R'} t_0 c_R^\dagger c_{R'}^s, \quad (7)$$

\[ + \sum_{\ell,s} \sum_{R,a_M} t_1 \epsilon^{\pm \kappa_{\ell R} a_M} c_R^\dagger c_{R+a_M}^s \delta_{RR'}, \]

where $s = \pm$ denotes spin (or equivalently valley $\pm K$), and $\ell = \pm$ labels orbitals localized in the bottom ($+1$) and top ($-1$) layers and centered around the $R^M_M$ and $R^M_M$ sites. The second line of (7) captures inter-layer hopping between nearest neighbors on the honeycomb lattice. Its form is constrained by the requirements that the energy spectra have threefold rotational symmetry and be identical at $\kappa_+$ and $\kappa_-$ points. The first line of (7) captures intralayer hopping between next nearest neighbors on the honeycomb lattice; the bond and spin-dependent phase factors $\exp(is\kappa_{\ell R} a_M)$, which take values of $\exp(\pm 2\pi s/3)$, are analogous to the Peierls substitution and account for the momentum shift $\kappa_\ell$ in (6). The Hamiltonian (7) is equivalent to the Kane-Mele model [27, 28], and to two time-reversed-partner copies of the Haldane model [29]. It correctly captures both the topological character and the energy dispersion of the first two bands in Fig. 3(a). The Kane-Mele model applies only at small twist angles. Beyond a critical angle $\theta^*_b \approx 1.35^\circ$ the gap between the second and the third bands closes at the $\gamma$ point, as illustrated in Fig. 3. When $\theta$ crosses $\theta^*_b$ from below, the Chern number of the first $+K$ band remains as $+1$, while the Chern numbers for the second and third...
bands change from $(-1, 0)$ to $(+1, -2)$. Although, the two-orbital model (7) is not fully applicable for $\theta > \theta^*_1$ it still captures the main character of the first two bands in regions of momentum space away from the $\gamma$ point. The system remains as a quantum spin Hall insulator when the chemical potential is in the gap between the first and the second bands until $\theta$ reaches $\theta^*_2 \approx 2.44^\circ$, beyond which there is no global gap between the first two bands.

**Field induced topological transition.**—Because the two sublattices in (7) are associated with different layers, a vertical electric field generates a staggered sublattice potential, which can induce a topological phase transition [27–29]. To study this transition, we add a layer dependent potential $tV_z/2$ to the moiré Hamiltonian (1) so that $\Delta_z \rightarrow (\Delta_b - \Delta_t + V_z)/2$. (We neglect the small spatial modulation of $V_z$ due to variation in the vertical distance between layers in the moiré pattern [20].) The magnitude of $V_z$ has a critical value $|V_z|_c$, at which the gap between the first and the second moiré bands closes at $\kappa_{\pm}$ points. When $|V_z| > |V_z|_c$, wave functions in the first moiré band are primarily localized in one single layer and the band becomes topologically trivial. The tight-binding model (7) predicts that $|V_z|_c$ is equal to splitting between the first and the second bands at $\kappa_{\pm}$ when $V_z = 0$, because the interlayer hopping term in (7) vanishes at these momenta. In Fig. (4c) we compare values of $|V_z|_c$ calculated from the tight-binding and the full moiré band Hamiltonian, showing that they match well, particularly for small twist angles (long moiré period). We note that there is no one-to-one correspondence between the Chern numbers $C$ of the electronic bands and the winding number $N_w$ of the pseudospin field, which remains non-trivial until $V_z$ equals $|\Delta_b - \Delta_t|$ evaluated at the $R^X_M$ or $R^M_X$ sites.

**Interaction effects.**—When the moiré bands are nearly flat, the density of states is strongly enhanced [Fig. 3(b)] and many-body interaction effects are magnified. Here we focus on interaction effects within the first two moiré bands at zero $V_z$ and small $\theta$. The on-site Coulomb repulsion $U_0$ scales as $e^2/(ca_W)$, where $c$ is an effective dielectric constant that depends on the three-dimensional dielectric environment, and $a_W$ is the spatial extent of the Wannier orbitals centered at $R^X_M$ or $R^M_X$ sites. For $\theta$ around $1^\circ$, we find that $U_0$ can be more than one order of magnitude larger than the hopping parameters $t_{0,1}$ [20]. In the strong correlation limit, we anticipate that the interplay between layer and spin/valley degrees of freedom will lead to unusual distinct insulating states at integer numbers of holes per MUC. For one hole per MUC, where the first moiré band is half filled, one candidate insulating state is ferromagnetic. Because the single-particle Hamiltonian has only $U(1)$ symmetry, perpendicular spin polarization is energetically preferred. The Ising spin anisotropy implies finite temperature phase transitions. When the first moiré band is completely spin-polarized, the system is a quantum anomalous Hall insulator. Similar physics could occur for three holes per MUC, where the second moiré band is half filled. For two holes per MUC (equivalently one hole per sublattice site of the honeycomb lattice in Kane-Mele model), there is a competition between the quantum spin Hall insulator and the antiferromagnetic Mott insulator [32], which occur for weak and strong interactions respectively. For some fractional numbers of holes per MUC, the flat bands may host fractional topological insulators [33].

**Discussion.**—Moiré bands with valley-contrasting Chern numbers can also be realized in some graphene-based moiré systems [15, 16, 34]. In this case however, quantum spin Hall states that might be induced by interactions cannot survive to accessible temperatures because electrons in graphene have accurate SU(2) spin symmetry which enhances fluctuation effects. In Ref. [35], quantum spin Hall nano-dots and nano-stripes have been proposed for TMD-based moiré systems in which the large gap between valence and conduction bands needs to be inverted by strong vertical electric field. In contrast, our proposal relies only on valence band states. Ref. [14] has suggested that the SU(2) Hubbard model can be simulated in TMD heterobilayers. In twisted TMD homobilayers, the two layers can be effectively decoupled by using a finite layer bias potential to drive the system into region (II) of the phase diagram in Fig. (4c). Thus, conventional one-orbital Hubbard model can also be studied in twisted homobilayers, with a greater scope for *in situ* manipulation of model parameters. We note that compared to heterobilayers, twisted TMD homobilayers may be experimentally realized with a more precise control of the twist angle by adopting the same techniques used for twisted graphene bilayers [17, 19].

![FIG. 4. (a) Moiré bands at twist angle $2\theta$. (b) Energy gaps between adjacent moiré bands as a function of $\theta$. The gap $\varepsilon_{ij}$ is the difference between the minimum energy of band $i$ and the maximum energy of band $j$. (c) Phase diagram as a function of angle $\theta$ and layer bias potential $V_z$. When the chemical potential is in the gap between the first and the second band, the system is a quantum spin Hall insulator in region (I) and a trivial insulator in region (II). In region (III), the two bands overlap in energy. The solid and dashed green lines show the critical bias potentials obtained using the full moiré Hamiltonian and the effective tight-binding model respectively. (b) and (c) have the same horizontal axis.](image)
Our study suggests an effective way to realize topological moiré bands in a large class of two dimensional materials \cite{23}, including those with intrinsic magnetism \cite{24}.

FW and IM were supported by the Department of Energy, Office of Science, Materials Science and Engineering Division. Work at Austin was supported by the Department of Energy, Office of Basic Energy Sciences under contract DE-FG02-02ER45958 and by the Welch foundation under grant TBF1473. We acknowledge HPC resources provided by the Texas Advanced Computing Center (TACC) at The University of Texas at Austin.

---

[1] B. Hunt, J. D. Sanchez-Yamagishi, A. F. Young, M. Yankowitz, B. J. LeRoy, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, and R. C. Ashoori, Science 340, 1427 (2013)
[2] C. R. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, M. Moon, M. Koshino, T. Taniguchi, K. Watanabe, K. L. Shepard, J. Hone, and P. Kim, Nature 497, 598 (2013)
[3] L. Wang, Y. Gao, B. Wen, Z. Han, T. Taniguchi, K. Watanabe, M. Koshino, J. Hone, and C. R. Dean, Science 350, 1231 (2015)
[4] L. M. Spann, A. A. Zibrov, H. Zhou, T. Taniguchi, K. Watanabe, M. P. Zaletel, and A. F. Young, Science 360, 62 (2018)
[5] G. Chen, L. Jiang, S. Wu, B. Lv, H. Li, K. Watanabe, T. Taniguchi, Z. Shi, Y. Zhang, and F. Wang, arXiv:1803.01985 (2018)
[6] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. U.S.A. 108, 12233 (2011)
[7] K. Kim, A. DaSilva, S. Huang, B. Fallahazad, S. Larentis, T. Taniguchi, K. Watanabe, B. J. LeRoy, A. H. MacDonald, and E. Tutuc, Proc. Natl. Acad. Sci. U.S.A. 114, 3364 (2017)
[8] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature 556, 80 (2018)
[9] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature 556, 43 (2018)
[10] D. K. Efimkin and A. H. MacDonald, Phys. Rev. B 98, 035404 (2018)
[11] N. F. Q. Yuan and L. Fu, Phys. Rev. B 98, 045103 (2018)
[12] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, arXiv:1803.09742 (2018)
[13] F. Wu, T. Lovorn, E. Tutuc, and A. H. MacDonald, arXiv:1804.03151, accepted by Phys. Rev. Lett. (2018)
[14] F. Wu, A. H. MacDonald, and I. Martin, arXiv:1805.08735 (2018)
[15] Y.-H. Zhang, D. Mao, Y. Cao, P. Jarillo-Herrero, and T. Senthil, arXiv:1805.08232 (2018)
[16] B. L. Chittari, G. Chen, Y. Zhang, F. Wang, and J. Jung, arXiv:1806.00462 (2018)
[17] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, Phys. Rev. Lett. 108, 196802 (2012)
[18] J. Jung, A. Raoux, Z. Qiao, and A. H. MacDonald, Phys. Rev. B 89, 205414 (2014)
[19] H. C. P. Movva, B. Fallahazad, K. Kim, S. Larentis, T. Taniguchi, K. Watanabe, S. K. Banerjee, and E. Tutuc, Phys. Rev. Lett. 118, 247701 (2017)
[20] See Supplemental Material at URL for results of a four-band model, additional discussion of interaction effects and description of moiré Hamiltonain in AB stacking configuration, which includes Ref. 38.
[21] P. Giannozzi et al., J. Phys.: Condens. Matter 21, 395502 (2009)
[22] F. Wu, T. Lovorn, and A. H. MacDonald, Phys. Rev. B 97, 035306 (2018)
[23] F. Wu, T. Lovorn, and A. H. MacDonald, Phys. Rev. Lett. 118, 147401 (2017)
[24] K. Moon, H. Mori, K. Yang, S. M. Girvin, A. H. MacDonald, L. Zheng, D. Yoshioka, and S.-C. Zhang, Phys. Rev. B 51, 5138 (1995)
[25] N. Nagaosa and T. Yokura, Nature Nanotechnology 8, 899 (2013)
[26] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. 82, 1959 (2010)
[27] C. L. Kane and E. J. Mele, Phys. Rev. Lett. 95, 226801 (2005)
[28] C. L. Kane and E. J. Mele, Phys. Rev. Lett. 95, 146802 (2005)
[29] F. D. M. Haldane, Phys. Rev. Lett. 61, 2015 (1988)
[30] C. Zhang, C.-F. Chau, X. Ren, M.-Y. Li, L.-J. Li, C. Jin, M.-Y. Chou, and C.-K. Shih, Sci. Adv. 3, e1601459 (2017)
[31] H. Yu, G.-B. Liu, J. Tang, X. Xu, and W. Yao, Sci. Adv. 3, e1701696 (2017)
[32] M. Hohenadler, T. C. Lang, and F. F. Assaad, Phys. Rev. Lett. 106, 100403 (2011)
[33] X.-L. Qi, Phys. Rev. Lett. 107, 126803 (2011)
[34] J. C. W. Song, P. Samutpraphoot, and L. S. Levitov, Proc. Natl. Acad. Sci. U.S.A. 112, 10879 (2015)
[35] Q. Tong, H. Yu, Q. Zhu, Y. Wang, X. Xu, and W. Yao, Nat. Phys. 13, 356 (2017)
[36] S. Lebegue, T. Bjorkman, M. Klintenberg, R. M. Nieminen, and O. Eriksson, Phys. Rev. X 3, 031002 (2013)
[37] X. Li, T. Cao, Q. Niu, J. Shi, and J. Feng, Proc. Natl. Acad. Sci. U.S.A. 110, 3738 (2013)
[38] Y. Wang, Z. Wang, W. Yao, G.-B. Liu, and H. Yu, Phys. Rev. B 95, 115429 (2017)
Supplemental Material

FOUR-BAND MODEL

In monolayer TMD, the $k.p$ Hamiltonian at $\pm K$ valleys is typically described by a massive Dirac model in the basis of conduction and valence band states [17], which is particularly important for the valley dependent optical selection rules. In the main text, we have discarded the conduction band states and used a parabolic band approximation for valence band states. There is a question about whether such an approximation is adequate to capture all important low-energy physics, in particular the valley band structure and the topological character. Below we answer this question in the affirmative by studying a moiré Hamiltonian that includes both the valence and conduction bands.

As in the main text, we study spin up states at $+K$ valley. For aligned bilayers, the $k.p$ Hamiltonian that includes both conduction and valence band states is given by:

$$H_\uparrow(\theta = 0, d_0) = \left( \begin{array}{c} T(d_0) \hbar k \cdot d_0 \end{array} \right),$$

(8)

where $h_\ell$ is the massive Dirac Hamiltonian in layer $\ell$, and $T(d_0)$ is the interlayer tunneling matrix. $h_\ell$ is given by:

$$h_\ell(k, d_0) = \left( \begin{array}{c} \Delta_\ell \Delta_{\ell,c}(d_0) \hbar v_F(k_x - i k_y) \end{array} \right).$$

(9)

Here the two components refer to the metal $d_{xz}$ and $d_{x^2-y^2} + id_{xy}$ orbital states, which carry different angular momenta and represent the main character respectively for conduction ($c$) and valence ($v$) band states. In (9), $\Delta_\ell$ is the average band gap, $\Delta_{\ell,c}(d_0)$ is the variation of band $\xi$ extrema energy as a function of $d_0$, and $v_F$ is the Fermi velocity. Here $\ell$ is the band index for $c$ and $v$. Similar to Eq.(2) in the main text, $\Delta_{\ell,\xi}$ can be parametrized as:

$$\Delta_{\ell,\xi}(d_0) = 2V_\xi \sum_{j=1,3,5} \cos(G_j \cdot d_0 + \ell \psi_\xi).$$

(10)

The tunneling matrix $T(d_0)$ is derived by using a two-center approximation [6] and also taking into account the threefold-rotational symmetry at high-symmetry displacements [6]. $T(d_0)$ is parametrized as:

$$T(d_0) = \left( \begin{array}{cc} w_c & w_{cv} \\ w_{cv} & w_v \end{array} \right)$$

$$+ \left( \begin{array}{cc} w_c e^{-i2\pi/3} & w_{cv} e^{i2\pi/3} \\ w_{cv} e^{-i2\pi/3} & w_v \end{array} \right) e^{-iG_2 \cdot d_0}$$

$$+ \left( \begin{array}{cc} w_c e^{-i2\pi/3} & w_{cv} e^{i2\pi/3} \\ w_{cv} e^{-i2\pi/3} & w_v \end{array} \right) e^{-iG_3 \cdot d_0}.$$  

(11)

The mirror operation ($z \leftrightarrow -z$) leads to the constraints that $w_c$ and $w_v$ are real numbers, and $w_{cv} = w_{vc}^*$. Note that the interlayer tunneling between valence and conduction band states vanishes at $d_0 = 0$.

We determine parameter values by fitting to the first-principles band structure at the three high-symmetry displacements. To simplify the fitting, we set $w_{cv} = w_{vc}$ to be zero, assuming that their effects on the valence bands are weak. For AA stacked WSe$_2$ bilayer, we obtain that $\Delta_\ell = 1.2$ eV, $v_F = 0.55 \times 10^6$ m/s, $(V_0, \psi_\xi, \psi_c) = (8.9$ meV, $91^\circ$, $9.7$ meV) and $(V_c, \psi_c, \psi_{vc}) = (6.8$ meV, $89.7^\circ$, $1.1$ meV). Here $\psi_c$ and $\psi_v$ are close to $90^\circ$, which implies that the scalar potential $(\Delta_{\ell,\xi} + \Delta_{\ell,c})(d_0)$ has an amplitude much smaller than that of the potential difference $(\Delta_{\ell,\xi} - \Delta_{\ell,c})/2$. We note that density-functional theory underestimates the band gap $\Delta_g$, but the exact numerical values should not influence qualitative effects that we discuss.

The band gap $\Delta_g$ is much larger than other energy scales. Therefore, we can integrate out conduction band states using a second-order perturbation theory, and the resulting model is exactly that in Eq. (1) of the main text. The effective mass $m^*$ is then approximated by $\Delta_g/(2\hbar^2 v)$, and the spatial modulation of $m^*$ is of order $V_\xi/\Delta_g \ll 1$ and can be neglected. This provides the first justification that the model with only valence band states is valid.

We continue with the four-band Hamiltonian [5]. When there is a twist between the two layers, the moiré Hamiltonian is:

$$H_\uparrow(\theta, d_0) = \left( \begin{array}{cc} h_\uparrow(k, r) & T(r) \\ T^*(r) & h_\uparrow(k, r) \end{array} \right),$$

(12)

where $r$ is the two-dimensional position operator. $h_\uparrow$ in the moiré pattern has the form

$$h_\uparrow(k, r) = \left( \begin{array}{cc} \Delta_g + \Delta_{\ell,c}(r) & 0 \\ 0 & \Delta_{\ell,v}(r) \end{array} \right)$$

$$+ e^{-i\xi \cdot k} \left[ h_{\varphi}^{\uparrow,\xi} \cdot (\kappa_r + \xi) e^{i \xi \cdot r} \right],$$

(13)

where $\xi$ is the Pauli matrix in the band basis. The moiré

![FIG. 5. Moiré band structure obtained from the four-band model in (12). The twist angle is $1^\circ$. (a) Moiré valence bands, which are nearly identical to those shown in Fig. 3(a) of the main text. (b) Two lowest energy moiré conduction bands in spin up sector at $+K$ valley. Note that the energy is measured relative to $\Delta_g$ in (b). Both bands in (b) have a bandwidth less than 1 meV.](image)
potentials $\Delta_{t,\ell}(r)$ and the tunneling term $T(r)$ are obtained from Eqs. [10]-[11] by replacing $\mathbf{d}_0$ with $\theta \mathbf{z} \times \mathbf{r}$. The moiré Hamiltonian [12] is formally similar to that for twisted bilayer graphene [6-14], and the main differences are the Dirac mass $\Delta_0$ and the moiré potentials $\Delta_{t,\ell}(r)$.

We diagonalize the moiré Hamiltonian [12] by using a plane-wave expansion and parameter values of WSe$_2$ bilayer. The moiré bands and the Chern numbers are summarized in Fig. 6. The moiré valence bands have nearly identical energy spectra as that obtained using the two-band model of the main text, and the Chern numbers are also consistent. Fig. 6(b) shows the two lowest-energy moiré conduction bands in spin up sector at $+K$ valley, which respectively carry Chern number $-1$ and $+1$. The topological character of the moiré conduction bands can be understood in the same way as that for valence bands.

In conclusion, the two-band model given in Eq. (1) of the main text provides a faithful low-energy description of moiré valence bands at small twist angles. On the other hand, the four-band model captures high energy physics of interband transition and is suitable to study exciton and optical properties in the twisted homobilayer, which we leave to a future work.

**KANE-MELE-HUBBARD MODEL**

We provide additional discussion of many-body interactions in the topological moiré bands. As in the main text, we focus on interaction effects within the first two moiré bands with small twist angle, where the tight-binding model $H_{TB}$ is applicable.

The interaction strength can be tuned by the three-dimensional dielectric environment, and the range of interaction can be controlled by a nearby metallic gate. Here we only consider the on-site repulsive interaction $U_0$ and assume that longer-range interactions are suppressed due to the nearby gate. Effects of long-range Coulomb interaction on the topological bands are certainly also interesting, which we leave to future exploration. $U_0$ scales as $e^2/(\epsilon a_W)$, where $\epsilon$ is the effective dielectric constant. $a_W$ is the spatial extent of the Wannier orbitals, which can be estimated based on a harmonic oscillator potential approximation of the moiré potential $\Delta_{t,\ell}(r)$ near one of its maximum position [13]. At $\theta = 1^\circ$, $a_W$ is about 3 nm; for comparison, the corresponding moiré period $a_M$ is 19 nm. If we take $\theta$ to be 10, then $U_0 \approx e^2/(\epsilon a_W) \approx 47$ meV, which is a hundred times larger than the hopping parameter $t_0 \approx 0.45$ meV at $\theta = 1^\circ$. Therefore, the ratio $U_0/t_0$, which depends on the dielectric environment and the twist angle, can be much greater than one.

By combining the tight-binding model and the on-site interaction, we arrive at the Kane-Mele-Hubbard model:

$$H = H_{TB} + U_0 \sum_{\mathbf{R}R} n_{\mathbf{R}\uparrow} n_{\mathbf{R}\downarrow}. \quad (14)$$

The tight-binding model is given by:

$$H_{TB} = \sum_{\ell,s} \sum_{\mathbf{R}\mathbf{R}'} t_0 c_{\mathbf{R}\ell s}^\dagger c_{\mathbf{R}'(-\ell)s} + \sum_{\ell,s} \sum_{\mathbf{R}} t_1 e^{i \kappa_0 a M (R + a_M) s} c_{\mathbf{R}\ell s}^\dagger c_{\mathbf{R}\ell s}, \quad (15)$$

where $\sigma_{x,y,z}$ and $s_{x,y,z}$ are Pauli matrices in the sublattice and spin spaces, and $f_{0,x,y,z}$ are functions of the momentum $k$. The tight binding model has two energy branches $f_0 \pm \sqrt{f_x^2 + f_y^2 + f_z^2}$ and each branch is doubly degenerate. The term $f_0$ is not important for the topological properties and is absent in the original Kane-Mele model of Refs. [27, 28]. At the filling factor where there are two electrons per unit cell (equivalent to one electron per sublattice site), there is a competition between the quantum spin Hall insulator and the antiferromagnetic Mott insulator in the Kane-Mele-Hubbard model [32]. Here we focus on a different filling factor where there are three electrons (equivalent to one hole) per unit cell. At this filling factor, the upper two bands are half filled in the non-interacting limit. Because the upper two bands carry spin contrast Chern numbers, the isolated upper band in either spin sector does not have exponentially localized Wannier states. These nearly flat topological bands are similar to Landau levels which are completely flat but have no exponentially localized Wannier states. It is well established that Landau levels with internal degeneracies (for example spin) develop ferromagnetism when they are partially filled with an integer number of electrons per cyclotron orbit. By making an analogy to the Landau level problems, we anticipate that ferromagnetism is likely at the filling factor of one hole per moiré unit cell.

We now use a mean-field theory to demonstrate the Ising anisotropy in the ferromagnetic state. We perform a Hartree-Fock approximation for the Hubbard interaction on each site:

$$n_{\uparrow} n_{\downarrow} \approx \frac{n_0^2 + m^2}{4} - \frac{1}{2} c^\dagger (m \cdot s) c \quad (16)$$

where $n_0 = \langle n_{\uparrow} + n_{\downarrow} \rangle$ and $m = \langle c^\dagger s c \rangle$. Here $\langle \ldots \rangle$ represents the mean-field average value, and $n_0 = 3/2$. The full mean-field Hamiltonian is:

$$H_{MF} = 2N \frac{M^2}{U_0} + \sum_k c_k^\dagger [f_0 + f_x \sigma_z s_z + f_x \sigma_x + f_y \sigma_y + M \cdot s] c_k. \quad (17)$$
where $M = -U_0 m / 2$, and $N$ is the number of unit cells. $H_{MF}$ generally has four non-degenerate energy bands. We consider zero temperature, and assume that the magnetization is strong enough such that the highest energy band is completely empty and other three bands are fully occupied by electrons. The total energy is then given by:

$$
\mathcal{E}(M) = \text{constant} + 2N \frac{M^2}{U_0} - \sum_k \sqrt{M^2 + f^2} \frac{2|M|\sqrt{f_+^2 + f_\perp^2 \cos^2 \theta_M}}{2},
$$

where $f^2 = f_+^2 + f_\perp^2$, $f_\perp^2 = f_+^2 + f_\parallel^2$, and the vector $M$ is parametrized by the spherical coordinates $M(\sin \theta_M \cos \phi_M, \sin \theta_M \sin \phi_M, \cos \theta_M)$. The saddle point equation $\partial \mathcal{E} / \partial M = 0$ correctly reproduces the self-consistent equation $m = \langle c^\dagger s c \rangle$. The energy functional $\mathcal{E}(M)$ is independent of the azimuthal angle $\phi_M$, which is a result of the spin $U(1)$ symmetry of the Hamiltonian (14). $\mathcal{E}(M)$ does depend on the polar angle $\theta_M$, and is minimized by placing $M$ along $\hat{z}$ axis. Thus, there is an Ising anisotropy that favors the out-of-plane spin polarization.

**AB STACKING CONFIGURATION**

In TMD bilayers with long period moiré pattern, there is another distinct stacking configuration that we label as AB [22]. AA and AB stacking, which are sometimes respectively referred to as $R$ and $H$ stacking [31], are distinguished by a $180^\circ$ rotation of the top layer. Below we discuss moiré bands in the AB stacking configuration.

In AB stacking with zero twist angle and an arbitrary in-plane displacement $d_0$, there is always an inversion symmetry $\hat{P}$, and the inversion center is at the midpoint between the two shifted metal atoms located in different layers [33], as illustrated in Fig. 6. The combined symmetry $\hat{P} \hat{T}$ enforces a double degeneracy of the band structure at every momentum. We again focus on $+K$ valley, and electronic properties at $-K$ valley can be inferred by using the $\hat{T}$ symmetry. At $+K$ valley, the topmost valence bands associated with the bottom and top layer respectively have spin up and down characters, and they are degenerate due to the $\hat{P} \hat{T}$ symmetry, leading to the following $k.p$ Hamiltonian:

$$
\mathcal{H}_{+K}(\theta = 0, d_0) = \begin{pmatrix}
-\frac{\hbar^2 k^2}{2m^2} + \Delta(d_0) & 0 \\
0 & -\frac{\hbar^2 k^2}{2m^2} + \Delta(d_0)
\end{pmatrix},
$$

where the two components refer to bottom (spin up) and top (spin down) layers, $\Delta(d_0)$ is the variation of the valence band maximum energy as a function of $d_0$, and the $2 \times 2$ Hamiltonian is proportional to an identity matrix in the layer pseudospin space because of the the $\hat{P} \hat{T}$ symmetry. In this low-energy Hamiltonian (19), the two layers are effectively decoupled.

In twisted bilayer, we obtain the moiré Hamiltonian by replacing $d_0$ with $\theta \hat{z} \times r$ and taking into account the momentum shift between the two layers:

$$
\mathcal{H}_{+K} = \begin{pmatrix}
-\frac{\hbar^2 (k - \kappa_\perp)^2}{2m^2} + \Delta(r) & 0 \\
0 & -\frac{\hbar^2 (k - \kappa_\perp)^2}{2m^2} + \Delta(r)
\end{pmatrix},
$$

where the two layers are again effectively decoupled and can be studied separately in the single particle Hamiltonian, and the momentum shift $\kappa_\perp$ can therefore be removed by a layer-dependent gauge transformation. In each layer electrons move in the layer-independent potential $\Delta(r)$ and form moiré bands that are topologically trivial. The topmost moiré valence bands can be used to simulate the Hubbard model on a triangular lattice with two orbitals that are centered in different layers but at the same in-plane positions. This two-orbital Hubbard model has an approximate $SU(4)$ symmetry, which would become exact if the vertical separation between the two orbitals were neglected. We will report a detailed study of this two-orbital Hubbard model in a separate work. A vertical electric field generates an energy difference between the two orbitals, and can lead to the realization of the conventional single-orbital Hubbard model when the field is strong enough [19].