Quantum valley Hall effect, orbital magnetism, and anomalous Hall effect in twisted multilayer graphene systems

Jianpeng Liu$^1$ and Xi Dai$^1$

$^1$Department of Physics, Hong Kong University of Science and Technology, Kowloon, Hong Kong

We study the electronic structures and topological properties of $(M + N)$-layer twisted graphene systems. We consider the generic situation that $N$-layer graphene is placed on top of the other $M$-layer graphene, and is twisted with respect to each other by an angle $\theta$. In such twisted multilayer graphene (TMG) systems, we propose that there exists two low-energy bands for each valley emerging from the interface between the $M$ layers and the $N$ layers. These two low-energy bands in the TMG system possess valley Chern numbers that are dependent on both the number of layers and the stacking chiralities, and can be further tuned by an vertical electric field. The valley Chern numbers of the low-energy bands are associated with giant and valley-contrasting orbital magnetizations, which implies that the valley degeneracy can be lifted by weak external magnetic fields or by Coulomb interactions, and a valley-polarized (quantum) anomalous Hall state may be realized.

Twisted bilayer graphene (TBG) has drawn significant attention recently due to the observations of the correlated insulating phases$^{1-5}$ and unconventional superconductivity$^{5,6}$. At small twist angles, the low-energy states of TBG are characterized by four low-energy bands contributed by the nearly decoupled monolayer valleys$^{7,8}$. Around the “magic angles”, the bandwidths of the four low-energy bands become vanishingly small, and these nearly flat bands are believed to be responsible for the correlated insulating and superconducting properties observed in TBG. Numerous theories have been proposed to understand the electronic structures$^{9-20}$, the correlated insulating phase$^{9,21-31}$, and the mechanism of superconductivity$^{9,22,24,25,27,32-37}$.

On the other hand, the four low-energy bands already exhibits interesting properties at the single-particle level. It has been shown that the four low-energy bands are topologically nontrivial in the sense that they are characterized by odd windings of Wilson loops$^{13,18,38}$, which is an example of the fragile topology$^{14}$. The four flat bands have been further proposed to be equivalent to the zeroth pseudo Landau levels (LLs) with opposite Chern numbers and sublattice polarizations$^{18}$, which is the origin of the nontrivial band topology in the TBG system.

Moreover, recently unconventional ferromagnetic superconductivity and correlated insulating phase have been observed in twisted double bilayer graphene$^{39,40}$. This implies that the low-energy flat bands, which are believed to be responsible for the correlated physics in TBG, may also exist in the twisted double bilayer graphene system. Indeed a recent theoretical study has suggested the presence of flat bands in twisted double bilayer graphene$^{41}$. Motivated by these works, in this paper we study the electronic structures and topological properties of twisted multilayer graphene (TMG). In particular, we consider the general situation that the $N$-layer chirally stacked graphene is placed on top of the other $M$-layer chirally stacked graphene, and they are twisted with respect to each other by a non-vanishing angle $\theta$, as schematically shown in Fig. 1(a) (for the case of $M = 2, N = 2$). In such a $(M + N)$-layer TMG system, we propose that there always exists two low-energy bands (for each valley), and that the bandwidths of the two low-energy bands become vanishingly small at the magic angles of twisted bilayer graphene (TBG) for arbitrary numbers of layers $M$ and $N$. The flat bands in the TMG system can be interpreted from the pseudo LL representation of TBG$^{18}$, and is protected by an approximate chiral symmetry in chiral graphene multilayers.

Moreover, we also find that there is a Chern-number hierarchy in the $(M + N)$-layer TMG system. In particular, when the stacking chiralities of the $M$ layers and $N$ layers are the same, the total Chern number of the two low-energy bands for each monolayer valley equals to $\pm(M - N)$ for each spin species$^{49}$. On the other hand, if the stacking chiralities of the $M$ layers and the $N$ layers are opposite, then the total Chern number of the two low-energy bands for each valley is $\pm(M + N - 2)$. The valley Chern numbers can be further tuned by an external electric field, leading to gate-tunable quantum valley Hall effect.

The non-vanishing valley Chern numbers of the low-energy bands are associated with giant and valley-contrasting orbital magnetizations, which implies that the valley degeneracy can be lifted by weak external magnetic fields due to Zeeman couplings, or by Coulomb interactions through spontaneous valley symmetry breaking. In either scenario, a valley-polarized (quantum) anomalous Hall state may be realized. The flat bands at the universal magic angles, together with the Chern-number hierarchy, make the TMG systems a unique platform to study strongly correlated physics with nontrivial band topology, and may have significant implications on the observed ferromagnetic superconductivity and correlated insulating phase in twisted double bilayer graphene$^{39,40}$. 
and \( d_1 = 0.0278 \text{ Å} \) in order to reproduce the interlayer distances in AA- and AB-stacked bilayer graphene. In this paper, the atomic corrugations of the two twisted layers at the interface (between the M layers and the N layers) is also be modeled by Eq. (1). On the other hand, the interlayer distances within the untwisted M layers and the untwisted N layers are set to the interlayer distance of Bernal bilayer graphene \( d_{AB} = 3.35 \text{ Å} \). At a small twist angle \( \theta \), the Brillouin zone (BZ) of the moiré supercell has been significantly reduced compared with those of the untwisted multilayers as shown in Fig. 1(b).

**B. The effective Hamiltonian**

The low-energy effective Hamiltonian of the twisted \((M + N)\)-layer TMG of the K valley is expressed as

\[
H^K_{\alpha, \alpha'}(M + N) = \left( \begin{array}{c} H^K_M(M) \cup \bar{H}^N_{\alpha}(N) \end{array} \right),
\]

where \( H^K_M(M) \) and \( \bar{H}^N_{\alpha}(N) \) are the effective Hamiltonians for the M-layer and N-layer graphene with stacking chiralities \( \alpha, \alpha' = +/- \). In particular,

\[
H^K_{\alpha}(M) = \left( \begin{array}{cccc} h_0(\hat{k}) & h_\alpha & 0 & 0 \\ h_\alpha^* & h_0(\hat{k}) & h_\alpha & 0 \\ 0 & h_\alpha & h_0(\hat{k}) & h_\alpha \\ & & & \ldots \end{array} \right),
\]

where \( h_0(\hat{k}) = -\hbar v_F (\hat{k} - \mathbf{K}_M) \cdot \sigma \) stands for the low-energy effective Hamiltonian for monolayer graphene near the Dirac point \( \mathbf{K}_M \), and \( h_\alpha \) is the interlayer hopping, with

\[
h_\pm = \left( \begin{array}{cc} 0 & 0 \\ 0 & \pm t_\perp \end{array} \right),
\]

and \( h_- = h^\dagger_+ \).

The off-diagonal term \( \cup \) represents the coupling between the twisted M layers and N layers. Here we assume that there is only the nearest neighbor interlayer coupling, i.e., the topmost layer of the M-layer graphene is only coupled with the bottom-most layer of the N-layer graphene, thus

\[
U = \left( \begin{array}{cccc} 0 & \ldots & 0 \\ \vdots & \ldots & 0 \\ U(\mathbf{r}) e^{-i\Delta \mathbf{k} \cdot \mathbf{r}} & \ldots & 0 \\ 0 & & & 0 \end{array} \right),
\]

where the 2×2 matrix \( U \) describes the tunneling between the Dirac states of the twisted bilayers\(^{8,11}\)

\[
U(\mathbf{r}) = \left( \begin{array}{cc} u_0 \varrho(\mathbf{r}) & u_0' \varrho(\mathbf{r} - \mathbf{r}_{AB}) \\ u_0' \varrho(\mathbf{r} + \mathbf{r}_{AB}) & u_0 \varrho(\mathbf{r}) \end{array} \right),
\]

where \( \mathbf{r}_{AB} = (\sqrt{3}L_s/3, 0) \), \( u_0' \) and \( u_0 \) denote the intersublattice and intrasublattice interlayer tunneling amplitudes, with \( u_0' \approx 0.098 \text{ eV} \), and \( u_0 \approx 0.078 \text{ eV} \).
are almost exactly flat at $\theta$ rated from the other bands. The two low-energy bands energy flat bands marked by the red lines that are separated in Fig. 2(a)-(d) respectively with the same stacking chiralities ($N=2$), ($in Appendix A. It worth to note that in realistic situations of the flat bands in the TMG systems can be found with the other graphene layers due to an (approximate) interface, and they remain flat even after being coupled with less the number of layers. It turns out that the two flat angles of TBG is universal for the TMG systems regardless the number of layers. It turns out that the two flat bands in TMG originates from the twisted bilayer at the interface, and they remain flat even after being coupled with the other graphene layers due to an (approximate) chiral symmetry of Eq. (2). More details about the origin of the flat bands in the TMG systems can be found in Appendix A. It worth to note that in realistic situations there are further neighbor interlayer hoppings in graphene multilayers, which would break the chiral symmetry of the effective Hamiltonian in Eq. (2). The remote interlayer hoppings in graphene multilayers may broaden the flat bands, and we will discuss in detail in Sec. III.

II. THE CHERN-NUMBER HIERARCHY AND QUANTUM VALLEY HALL EFFECT

A. The Chern-number hierarchy

The flat bands at the universal magic angle make the TMG systems a perfect platform to study the strongly correlated physics. In addition to the flat bands and the universal magic angles, the low-energy bands in the TMG systems also exhibit unusual topological properties with non-vanishing valley Chern numbers. To be specific, when the stacking chiralities of the $M$ layers and the $N$ layers are the same, the total Chern number of the two low-energy bands for each monolayer valley equals to $\pm (M-N)$. On the other hand, if the stacking chiralities of the $M$ layers and the $N$ layers are opposite, then the total Chern number of the two flat bands for each valley equals to $\pm (M+N-2)$. Such a Chern-number hierarchy is more concisely summarized in the following equation

$$C^K_{\alpha,\alpha'} = -(\alpha(M-1) - \alpha'(N-1)),$$
$$C^{K'}_{\alpha,\alpha'} = +\beta(M-1) - \alpha'(N-1)), \quad (7)$$

where $C^K_{\alpha,\alpha'} (C^{K'}_{\alpha,\alpha'})$ denotes the total Chern number of the two low-energy flat bands for the $K (K')$ valley, and the subscripts $\alpha, \alpha' = \pm$ represent the stacking chiralities of the $M$ layers and $N$ layers. We would like to emphasize
that the total Chern number of the two flat bands (per valley per spin) is a more robust quantity than the Chern number of each individual flat band. This is because the former is protected by the energy gaps between the two flat bands and the other high-energy bands, while the latter is crucially dependent on how the gap between the two flat bands is opened up.

Eq. (7) has been numerically verified using the effective Hamiltonian of TMG shown in Eq. (2). In particular, in Fig. 3(a) we plot the Wilson-loop eigenvalues (denoted as $w(k)$) of the $(2+2)$ TMG ($M=2$, $N=2$) at the first magic angle with the same stacking chirality. The blue circles and red diamonds represent the Wilson loops of the $K$ and $K'$ valleys respectively. As clearly shown in the figure, for each valley the total Chern number of the two flat bands vanishes. In Fig. 3(b) we still plot the Wilson loops of the $(2+2)$ TMG at the first magic angle, but with opposite stacking chiralities. It is clearly seen that for the $K$ valley (blue circles) the two Wilson loops carry the same Chern number $-1$, giving rise to a total Chern number of $-2$ for the $K$ valley (+2 for the $K'$ valley), which is consistent with Eq. (7). In Fig. 3(c) and (d) we plot the Wilson-loop eigenvalues for the $(2+4)$ TMG ($M=2$, $N=4$) at the first magic angle. When the stacking chiralities are the same, the total Chern number of the two flat bands for the $K$ ($K'$) valley equals to $-2$ ($+2$); while if the stacking chiralities are opposite, then the total Chern number of the two flat bands equals to $-4$ ($+4$) for the $K$ ($K'$) valley. Again, this is in perfect agreement with Eq. (7).

In order to understand the origin of the Chern-number hierarchy shown in Eq. (7), we manually tune the strength of the nearest neighbor interlayer hopping $t_\perp$, and inspect the evolution of the band structures of the TMG system. In Fig. 4 we show the band structures of $(2+2)$-TMG for the $K$ valley at $\theta = 1.05^\circ$ with different interlayer hopping strengths $t_\perp$. When $t_\perp = 0$ (Fig. 4(a)), the $(2+2)$-TMG system consists of three mutually decoupled subsystems: the TBG at the interface, one graphene monolayer below the TBG, and one graphene monolayer above the TBG. The TBG at the interface would give rise to two nearly flat bands for each valley at the magic angle (marked by red lines in Fig. 4(a)), whereas the bottom and top graphene monolayers contribute two Dirac cones with giant Fermi velocities around the $K_s$ and $K'_s$ point respectively, as shown in Fig. 4(a). If a small $t_\perp$ is turned on, small gaps would be opened up at the two Dirac points of the bottom and top graphene monolayers, as shown in Fig. 4(b) for $t_\perp = 0.01\text{eV}$. As a result, the valence and/or conduction band of each Dirac cone would acquire a Berry phase of $\pm \pi$ due to the pseudospin winding around the Dirac point, and has a Chern number $\pm 1/2$. If the two Dirac cones were decoupled from the TBG at the interface, then the Chern numbers of the valence band and the conduction band of each Dirac cone must be opposite. However, if the two Dirac cones were coupled with the flat bands from the TBG at the interface, then the conduction and valence bands of each Dirac cone could carry the same Chern number with a total Chern number of $\pm 1$. If the total Chern numbers of the two Dirac cones are the same, say, both are $+1$, then the two flat bands from the interface TBG must have the opposite Chern number $-1$, because the total Chern number of all the three subsystems must vanish. On the other hand, if the total Chern numbers of the two Dirac cones are opposite, then the total Chern number of the two flat bands has to vanish. As $t_\perp$ continues to increase, the gaps of two the two Dirac cones become larger and larger as shown in Fig. 4(c) and (d) for $t_\perp = 0.08\text{eV}$ and $t_\perp = 0.2\text{eV}$. Eventually the Dirac cones from the top and bottom graphene monolayers are pushed to high energies due to the strong interlayer hopping $t_\perp \approx 0.39\text{eV}$. On the other hands, by virtue of the approximate chiral symmetry, the two flat bands originating from the TBG at the interface would remain flat at the magic angles regardless of the magnitude of $t_\perp$.

It is straightforward to generalize the above argument to $(M+N)$-layer TMG with arbitrary number of layers $M$ and $N$. When $t_\perp = 0$, the low-energy band structure of the $(M+N)$-layer TMG system at the magic angle includes the two flat bands contributed by the interface TBG, the $(M-1)$ Dirac cones at $K_s$ from the graphene monolayers below the interface, and $(N-1)$ Dirac cones at $K'_s$ from the graphene monolayers above the interface. As $t_\perp$ increases, the three subsystems are coupled together, and gaps are opened up at the two Dirac points of the $M-1$ and $N-1$ graphene monolayers. As a result, the Chern numbers of the valence and/or conduction bands of the $M-1$ and $N-1$ Dirac cones would be $\pm (M-1)/2$ and $\pm (N-1)/2$. Again, since the two sets of graphene monolayers are coupled with the
TBG at the interface, the conduction and valence bands are allowed to have the same Chern numbers for each set of Dirac cones, such that the total Chern number for the $M - 1$ ($N - 1$) graphene monolayers equals to $\pm(M - 1)$ ($\pm(N - 1)$). On the other hands, the Chern number of the two flat bands originating from the interface TBG has to cancel those of the two sets of Dirac cones, which lead to two situations: (i) if the Chern numbers of the two sets of Dirac cones have the same sign and add together, then the total Chern number of the two flat bands equals to $\pm(M + N - 2)$ (i) if the Chern numbers of the two sets of Dirac cones have opposite signs, then the the total Chern number of the two flat bands equals to $\pm(M - N)$.

This concludes our explanation to the Chern number hierarchy given by Eq. (7).

### B. Gate tunable quantum valley Hall effect

The valley Chern numbers given by Eq. (7) can be further tuned by applying an vertical electric potential $V_{\perp}$. Taking the case of $(2 + 2)$-layer TMG as an example, we study the dependence of the valley Chern numbers on the vertical electric potential $V_{\perp}$ at $\theta = 1.05^\circ$. We consider three different fillings: (i) the chemical potential is in the single-particle gap below the two flat bands, with the Chern number of all the filled bands denoted by $C_{K+1}^0$; (ii) the chemical potential is at the charge neutrality point with one flat band being filled, and the Chern number of the filled bands is denoted by $C_{0}^K$. (iii) the chemical potential is above the two flat bands but below the other high-energy bands, and the Chern number of the filled bands is denoted by $C_{+1}^K$. The superscript $K$ refers to the $K$ valley.

In Table I we show the dependence of $C_{K+1}^{-1}$, $C_{0}^{-1}$, and $C_{+1}^K$ on the vertical electric potential $V_{\perp}$ (in units of meV) for $(2 + 2)$-layer TMG with opposite stacking chiralities at $\theta = 1.05^\circ$. We see that at small electric potential $|V_{\perp}| \lesssim 36$ meV, the Chern numbers of the $K$ valley are unchanged. In particular, the Chern number carried by the two flat bands is always $C_{+1} - C_{-1} = -2$. However, when $V_{\perp} \gtrsim 36$ meV, the Chern numbers at the three different fillings become zeros.

In Table II still we show $C_{K+1}^{-1}$, $C_{0}^{-1}$, and $C_{+1}^K$ vs. $V_{\perp}$ for $(2 + 2)$-layer TMG at $\theta = 1.05^\circ$, but with the same stacking chirality. When the vertical potential $V_{\perp}$ is 0, both $C_{K+1}^{-1}$ and $C_{0}^{-1}$ are vanishing, and $C_{+1}^K$ is ill defined since the two flat bands touch each other at $K_0$ and $K'_0$ points. However, once a small $V_{\perp}$ is applied, the Chern number at the charge neutrality point becomes $C_{0}^K = +2(-2)$ for negative (positive) $V_{\perp}$. As $|V_{\perp}|$ is further increased, $C_{0}^K$ remains as constant $\pm2$, but there are additional topological transitions between the flat bands and the other high-energy bands such that the Chern numbers below and above the flat bands $C_{-1}^K$ and $C_{+1}^K$ are changed from $\pm1$ to $\pm2$.

### III. VERIFICATIONS USING AN EMPIRICAL TIGHT-BINDING MODEL

The flat bands at the first magic angle (Fig. 2) and the Chern-number hierarchy (Eq. (7)) in TMG have been verified using a realistic microscopic tight-binding (TB) model. To be specific, the hopping parameter between two $p_z$ orbitals at different carbon sites $i$ and $j$ in the multilayer system is expressed in the Slater-Koster form

$$t(d) = V_\sigma (\frac{d \cdot \hat{z}}{d_0})^2 + V_\pi (1 - (\frac{d \cdot \hat{z}}{d_0})^2)$$

where $V_\sigma = V_\sigma^0 e^{-(r - d_i)/\delta_0}$, and $V_\pi = V_\pi^0 e^{-(r - a)/\delta_0}$. $d = (d_x, d_y, d_z)$ is the displacement vector between the two carbon sites. $a_0 = d/\sqrt{3} = 1.42$ Å, $d_c = 3.35$ Å is the interlayer distance in AB-stacked bilayer graphene, and $\delta_0 = 0.184$ Å. $V_\sigma^0 = 0.48$ eV and $V_\pi^0 = -2.7$ eV. The atomic corrugations at the interface between the M layers and the N layers in $(M + N)$ TMG are modeled by Eq. (1), and their effects can be taken into account by plugging Eq. (1) into the hopping parameter shown in Eq. (8).

The bandstructures calculated using the Slater-Koster TB model at $\theta \approx 1.08^\circ$ for the $(2 + 2)$ TMG are shown in Fig. 4. To be specific, the bandstructure of $(2 + 2)$ TMG with opposite and the same stacking chiralities are shown in Fig. 4(a) and (c) respectively. It is evident that there are four low-energy bands (contributed by the two valleys) which are separated from the other high-energy bands, and the bandwidths are on the order of 10-15 meV, which are greater than those from the continuum model (see Fig. 2(a). This is because in the continuum model only the nearest neighbor interlayer hopping is kept (see Eq. (A3)), which imposes a chiral symmetry to the Hamiltonian of Eq. (2).
width as shown in Fig. 2. However, such a chiral symmetry is broken in the realistic Slater-Koster TB model, and the bandwidths are expected to be enhanced due to the further neighbor interlayer hoppings.

In Fig. 4(b) and (d) we show the Wilson loops of the four low-energy bands of (2+2) TMG at $\theta \approx 1.08^\circ$ calculated using the Slater-Koster TB model. Fig. 4(b) ((d)) denotes the case with opposite (the same) stacking chiralities with the valley Chern number being $\pm 2$ (0). It is evident that the Wilson loops shown in Fig. 4(b) and (d) are topologically equivalent to those calculated by the continuum model as shown in Fig. 3(a)-(b), which provide a strong numerical support to the Chern-number hierarchy given by Eq. (7).

IV. VALLEY-CONTRASTING ORBITAL MAGNETIZATIONS AND VALLEY-POLARIZED STATES

The valley Chern numbers given by Eq. (7) implies opposite orbital magnetizations for the two monolayer valleys $K$ and $K'$. In particular, according to the “modern theory” of orbital magnetizations\(^{44-47}\), the orbital magnetization of the $(M+N)$-layer TMG for either $K$ or $K'$ valley can be expressed as

$$M_z = \frac{e}{2\hbar(2\pi)^2} \sum_n \int_{\epsilon_{nk} \leq \mu} \frac{d \mathbf{k} \text{Im} \{ \partial_{\mathbf{k}} \epsilon_{nk} \}}{\epsilon_{nk} - \epsilon_{nk} - 2\mu} |u_{nk}|$$

where $\epsilon_{nk}$ and $|u_{nk}|$ are the eigenenergies and (the periodic part of) Bloch eigenstates of a $(M+N)$-layer TMG Hamiltonian (denoted by $H$) for either $K$ or $K'$ valley, $\mu$ is the chemical potential, and $H_0 = e^{-i\mathbf{k} \cdot \mathbf{r}} H e^{i\mathbf{k} \cdot \mathbf{r}}$. Since the $K$ and $K'$ valleys are transformed to each other by a time-reversal operation, it is naturally expected that the valley-contrasting Chern numbers shown in Eq. (7) would lead to opposite orbital magnetization for the two valleys. The valley-contrasting orbital magnetizations further suggests that the valley degeneracy may be lifted by external magnetic fields or by Coulomb interactions, leading to a valley-polarized state with possibly quantized anomalous Hall conductivity due to the non-vanishing Chern number.

We have exploited this idea using the continuum model given by Eq. (2). To be explicit, we consider the case of $(2+2)$-layer twisted graphene at the first magic angle $\theta = 1.05^\circ$. We have calculated the orbital magnetizations ($M_z$) of the two flat bands for the $K$ valley using Eq. (9). The dependence of $M_z$ on the chemical potential $\mu$ is plotted in Fig. 5(a). The blue circles in Fig. 5(a) represent the situation that the stacking chiralities are same, with the valley Chern number $-2$ (0). The horizontal axis is the chemical potential, and the vertical dashed lines mark the bulk band edges of the two flat bands. (b) The band structures of the $K$ (solid lines) and $K'$ (dashed lines) valleys under external field $B = 2.5$ T along the negative z direction.
ley degeneracy can be easily lifted by a weak external magnetic field or by Coulomb interactions, leading to a valley-polarized quantum anomalous Hall state. In Fig. 5(b) we plot the Zeeman-splitled bandstructure of (2+2)-layer TMG with opposite stacking chiralities (with valley Chern numbers ±2) at the first magic angle, with an external magnetic field $B = 2.5 \text{T}$ pointing along the $-\hat{z}$ direction (pointing “down”). The chemical potential is set to the charge neutrality point ($\mu = 0$) such that it will not be changed after applying the magnetic field. As clearly shown in Fig. 4(b), both the valley and the spin degrees of freedom are split by the weak external magnetic field, but the valley splitting is much greater than the spin splitting due to the giant orbital magnetization. An external field $B = 2.5 \text{T}$ is already strong enough to open a gap between the states of the $K$ and $K'$ valleys, leading to a quantum anomalous Hall state with Chern number -2 for each spin species. We expect that such a valley-polarized state may also be realized by the spontaneous valley symmetry breaking due to exchange interactions, which we will leave for future study. The same kind of idea can also be applied to other ($M + N$)-layer TMG with non-vanishing valley Chern numbers.

V. SUMMARY

To summarize, we have studied the electronic structures and topological properties of the ($M + N$)-layer TMG system. We have proposed that, with the chiral approximation, there always exists two low-energy bands whose bandwidths become minimally small at the magic angle of TBG. We have further shown that the two flat bands in the TMG system are topologically non-trivial, and exhibits a Chern-number hierarchy. In particular, when the stacking chiralities of the $M$ layers and the $N$ layers are opposite, the total Chern number of the two low-energy bands for each valley equals to $\pm (M + N - 2)$ (per spin). On the other hand, if the stacking chiralities of the $M$ layers and the $N$ layers are the same, then the total Chern number of the two low-energy bands for each valley is $\pm (M - N)$ (per spin). The non-vanishing valley Chern numbers are associated with giant and valley-contrasting orbital magnetizations along $z$ direction, which implies that the valley degeneracy can be lifted by weak external magnetic fields or by Coulomb interactions, leading to a valley-polarized (quantum) anomalous Hall state. Specifically, for the (2+2)-layer TMG with opposite stacking chiralities, we have shown that when the chemical potential is at the charge neutrality point, an external magnetic field $\sim 2.5 \text{T}$ is already strong enough to lift the valley degeneracy, and may lead to a quantum anomalous Hall state. Our work is a step forward in understanding the electronic properties of twisted multilayer graphene. The universal magic angles and the Chern-number hierarchy proposed in this work make the TMG system a perfect platform to study the interplay between electrons’ Coulomb correlations and nontrivial band topology.

Note added: During the preparation of our manuscript, we note a very recently posted work Ref. 48, in which the electronic structures, superconductivity, and correlated insulating phase have been discussed for twisted double bilayer graphene with the same stacking chirality ($AB$-$AB$ stacking).

Acknowledgments

J.L. and X.D. acknowledge financial support from the Hong Kong Research Grants Council (Project No. GRF16300918). We thank Jinhua Gao, Zhen Ma, and Hongming Weng for invaluable discussions.

Appendix A: The flat bands and magic angles in the TMG system

In this appendix we explain the origin of the flat bands and universal magic angles in the ($M + N$)-layer TMG system. After some gauge transformations, the constant wavevectors $\mathbf{K}_M, \mathbf{K}_N$ in Eq. (3) can be removed. Taking the case of (2+2)-layer TMG with the same stacking chiralities as an example, the effective Hamiltonian is explicitly written as (after the gauge transformation)

$$
H^K_{+,+}(2+2) = \begin{pmatrix}
    h_0(\mathbf{k}) & h_+ & 0 & 0 \\
    h_+ & h_0(\mathbf{k}) & U(r) & 0 \\
    0 & U^\dagger(r) & h_0(\mathbf{k}) & h_+ \\
    0 & 0 & h_+ & h_0(\mathbf{k})
\end{pmatrix},
$$

where $h_0(\mathbf{k}) = -\hbar v_F \mathbf{k} \cdot \sigma$, and $h_\alpha$ and $U$ are defined in Eq. (4) and Eq. (5) of the main text. Then we make the following unitary transformation to the basis functions of Eq. (A1) (and more generally, to those of Eq. (2) of the main text)

$$
\begin{align*}
|\tilde{\psi}_{M,s}\rangle &= \frac{1}{2} \left( |\psi_{M,s}\rangle + i |\psi_{(M+1),s}\rangle \right), \\
|\tilde{\psi}_{(M+1),s}\rangle &= \frac{1}{2} \left( |\psi_{M,s}\rangle - i |\psi_{(M+1),s}\rangle \right), \\
|\tilde{\psi}_{l,s}\rangle &= |\psi_{l,s}\rangle, \text{ if } l \neq M, M+1
\end{align*}
$$

, where $|\psi_{l,s}\rangle$ denotes the Bloch state at the $K$ point from the $l$th layer and the $s$ sublattice. Applying the unitary transformation Eq. (A2) to Eq. (A1) (letting $M = 2$), then expanding the interlayer coupling term $U(r)$ to the linear order of $r/L_s$, Eq. (A1) becomes

$$
\tilde{H}^K_{+,+}(2+2) = \begin{pmatrix}
    h_0(\mathbf{k}) & h_+ & \frac{\hbar_i}{\sqrt{2}} & 0 \\
    \frac{h_+}{\sqrt{2}} & h_0(\mathbf{k} - \frac{\pi}{N} \mathbf{A}) & -3iu_0 & -i \frac{\hbar_i}{\sqrt{2}} \\
    \frac{3iu_0}{\sqrt{2}} & h_0(\mathbf{k} + \frac{\pi}{N} \mathbf{A}) & \frac{h_+}{\sqrt{2}} & 0 \\
    0 & -\frac{h_+}{\sqrt{2}} & \frac{h_+}{\sqrt{2}} & h_0(\mathbf{k})
\end{pmatrix}
$$


where the pseudo vector potential $A = (2\pi u_0)/(L_x e V_F) (y, -x)^{18}$. Note that the diagonal blocks $h_{0}(k \pm \xi A) = -h_{F}(k \pm e A/h)$ are equivalent to the Dirac fermions coupled with opposite pseudo magnetic fields, which would generate pseudo LLs of opposite Chern numbers $\pm 1^{18}$. In particular, the two zeroth pseudo LLs have opposite sublattice polarizations, thus the intrasublattice coupling term $\pm 3\eta_{0}$ in Eq. (A3) vanishes in the subspace of the zeroth pseudo LLs$^{18}$. Therefore, if we re-write Eq. (A3) in a mixed basis consisted of the Dirac states from the first and fourth layers, and the zeroth pseudo LLs from the second and third layers, then $\tilde{H}_{\pm \pm}(2 + 2)$ becomes

$$\tilde{H}_{\pm \pm}(2 + 2) \approx \left(\begin{array}{cc} h_{b} & 0 \\ 0 & h_{t} \end{array}\right).$$

Each element of Eq. (A4) is a $3 \times 3$ matrix. In particular,

$$h_{b} = \left(\begin{array}{ccc} 0 & h_{F}(k_{x} + i k_{y}) & 0 \\ h_{F}(k_{x} - i k_{y}) & 0 & 2 \sqrt{h_{F} \eta} \eta^{*} \\ 0 & 2 \sqrt{h_{F} \eta} & 0 \end{array}\right),$$

and

$$h_{t} = \left(\begin{array}{ccc} 0 & i \eta \frac{\xi}{\sqrt{2}} & 0 \\ -i \eta \frac{\xi}{\sqrt{2}} & 0 & h_{F}(k_{x} + i k_{y}) \\ 0 & h_{F}(k_{x} - i k_{y}) & 0 \end{array}\right),$$

where $\eta \xi$ denotes the coupling between the zeroth pseudo LL and the Dirac states of the first (fourth) layer, which can be expressed as some integral over the eigenfunctions of the zeroth pseudo LLs. Note that we have dropped the higher pseudo LLs in Eq. (A4).

The Diagonalizations of Eq. (A6) and Eq. (A5) yield two decoupled zero modes. These two zero modes originate from the two zeroth pseudo LLs contributed by the $M$th and $(M + 1)$th twisted bilayer ($M = 2$ for Eq. (A4)), and they stay at zero energy even after being coupled with the other layers due to the chiral symmetry of the effective Hamiltonian of TMG: all the matrix elements in Eq. (2) and Eq. (A1)-Eq. (A4) are intersublattice. As a consequence, if we apply the gauge transformation such that all the basis functions at the $B$ sublattice changes sign, then the total Hamiltonian and eigenenergies would change sign as well. However, the eigenenergies are supposed to be invariant under such a gauge transformation, which thus enforces that both $E(k)$ and $-E(k)$ have to be the eigenenergies of the Hamiltonian. Therefore, a zero mode would stay at zero energy as long as the chiral symmetry is preserved. Similar argument applies to any $(M + N)$-layer TMG systems with either opposite or the same stacking chiralities. As long as the chiral symmetry is preserved, the zeroth pseudo LLs emerging from the interface between the $M$ layers and $N$ layers would be pinned to zero energy.

On the other hands, it is well known that at the magic angles of TBG the bandwidth of the two low-energy bands for each valley is minimal. From the perspective of the pseudo LLs$^{18}$, it means that at the magic angles, the states within the zeroth pseudo LLs are minimally coupled with each other (and to the higher pseudo LLs), thus they are almost exactly flat. As discussed above, by virtue of the chiral symmetry, these zeroth pseudo LLs that are maximally flat at the magic angles would remain flat even after being coupled with the other layers in the TMG systems. It follows that the magic angles in TBG should be universal in the TMG systems by virtue of the chiral symmetry in Eq. (2).

---

1. Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, et al., Nature 556, 80 (2018).
2. A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. Kastner, and D. Goldhaber-Gordon, arXiv preprint arXiv:1901.03520 (2019).
3. Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, et al., arXiv preprint arXiv:1901.02997 (2019).
4. A. Keralsky, L. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Home, C. Dean, et al., arXiv preprint arXiv:1802.0776 (2018).
5. E. Codecido, Q. Wang, R. Koester, S. Che, H. Tian, R. Lv, S. Tran, K. Watanabe, T. Taniguchi, F. Zhang, et al., arXiv preprint arXiv:1902.05151 (2019).
6. Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature 556, 43 (2018).
7. J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. Lett. 99, 256802 (2007).
8. R. Bistritzer and A. H. MacDonald, Proceedings of the National Academy of Sciences 108, 12233 (2011).
9. H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Phys. Rev. X 8, 031089 (2018).
10. N. F. Q. Yuan and L. Fu, Phys. Rev. B 98, 045103 (2018).
11. M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Phys. Rev. X 8, 031087 (2018).
12. J. Kang and O. Vafek, Phys. Rev. X 8, 031088 (2018).
13. Z. Song, Z. Wang, W. Shi, G. Li, C. Fang, and B. A. Bernevig, arXiv preprint arXiv:1807.10676 (2018).
14. H. C. Po, L. Zou, T. Senthil, and A. Vishwanath, arXiv preprint arXiv:1808.02482 (2018).
15. K. Hejazi, C. Liu, H. Shapourian, X. Chen, and L. Balents, arXiv preprint arXiv:1808.01568 (2018).
16. G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, arXiv
preprint arXiv:1808.05250 (2018).
17 H. K. Pal, S. Spitz, and M. Kindermann, arXiv preprint arXiv:1803.07060 (2018).
18 J. Liu, J. Liu, and X. Dai, arXiv preprint arXiv:1810.03103v3 (2018).
19 B. Lian, F. Xie, and A. Bernevig, B, arXiv preprint arXiv:1811.11786 (2018).
20 Y.-H. Zhang, D. Mao, and T. Senthil, arXiv preprint arXiv:1901.08209 (2019).
21 A. Sboychakov, A. Rozhkov, A. Rakhmanov, and F. Nori, arXiv preprint arXiv:1807.08190 (2018).
22 H. Isobe, N. F. Q. Yuan, and L. Fu, Phys. Rev. X 8, 041041 (2018).
23 X. Y. Xu, K. T. Law, and P. A. Lee, Phys. Rev. B 98, 121406 (2018).
24 T. Huang, L. Zhang, and T. Ma, arXiv preprint arXiv:1804.06096 (2018).
25 C.-C. Liu, L.-D. Zhang, W.-Q. Chen, and F. Yang, Phys. Rev. Lett. 121, 217001 (2018).
26 L. Rademaker and P. Mellado, Phys. Rev. B 98, 235158 (2018).
27 J. W. F. Venderbos and R. M. Fernandes, Phys. Rev. B 98, 245103 (2018).
28 J. Kang and O. Vafek, arXiv preprint arXiv:1810.08642 (2018).
29 M. Xie and A. H. MacDonald, arXiv preprint arXiv:1812.04213 (2018).
30 C.-M. Jian and C. Xu, arXiv preprint arXiv:1810.03610 (2018).
31 N. Bultinck, S. Chatterjee, and M. P. Zaletel, arXiv preprint arXiv:1901.08110 (2019).
32 C. Xu and L. Balents, Phys. Rev. Lett. 121, 087001 (2018).
33 F. Wu, A. H. MacDonald, and I. Martin, Phys. Rev. Lett. 121, 257001 (2018).
34 X.-C. Wu, K. A. Pawlak, C.-M. Jian, and C. Xu, arXiv preprint arXiv:1805.06906 (2018).
35 B. Lian, Z. Wang, and B. A. Bernevig, arXiv preprint arXiv:1807.04382 (2018).
36 V. Kozii, H. Isobe, J. W. Venderbos, and L. Fu, arXiv preprint arXiv:1810.04159 (2018).
37 F. Wu, arXiv preprint arXiv:1811.10620 (2018).
38 J. Ahn, S. Park, and B.-J. Yang, arXiv preprint arXiv:1808.05375 (2018).
39 X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, arXiv preprint arXiv:1903.08130 (2019).
40 Y. Cao, D. Rodan-Legrain, O. Rubies-Bigord, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, arXiv preprint arXiv:1903.08596 (2019).
41 N. R. Chebrolu, B. L. Chittari, and J. Jung, arXiv preprint arXiv:1901.08420 (2019).
42 J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. B 86, 155449 (2012).
43 J.-K. Lee, S.-C. Lee, J.-P. Ahn, S.-C. Kim, J. I. Wilson, and P. John, The Journal of chemical physics 129, 234709 (2008).
44 T. Thonhauser, D. Ceresoli, D. Vanderbilt, and R. Resta, Physical review letters 95, 137205 (2005).
45 D. Xiao, J. Shi, and Q. Niu, Physical review letters 95, 137204 (2005).
46 D. Ceresoli, T. Thonhauser, D. Vanderbilt, and R. Resta, Physical Review B 74, 024408 (2006).
47 J. Shi, G. Vignale, D. Xiao, and Q. Niu, Physical review letters 99, 197202 (2007).
48 J. Y. Lee, E. Khalaf, S. Liu, X. Liu, Z. Hao, P. Kim, and A. Vishwanath, arXiv preprint arXiv:1903.08685 (2019).
49 Unless otherwise specified, in the rest of the paper the Chern number refers to the Chern number for each physical spin species.
50 The bandstructures with opposite stacking chiralities are very similar to those shown in Fig. 2.
51 The Dirac points of the $M$ graphene monolayers $K_M$ and of the $N$ graphene monolayers $K_N$ are respectively mapped to $K_s$ and $K'_s$ points of the moiré Brillouin zone, see Fig. 1(b)