

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Designing flat bands by strain

Zhen Bi, Noah F. Q. Yuan, and Liang Fu

Phys. Rev. B **100**, 035448 — Published 31 July 2019

DOI: [10.1103/PhysRevB.100.035448](https://doi.org/10.1103/PhysRevB.100.035448)

# Designing Flat Band by Strain

Zhen Bi,<sup>\*</sup> Noah F. Q. Yuan, and Liang Fu

*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA*

We study the effects of heterostrain on moiré bands in twisted bilayer graphene and bilayer transition metal dichalcogenide (TMD) systems. For bilayer graphene with twist angle near  $1^\circ$ , we show that heterostrain significantly increases the energy separation between conduction and valence bands as well as the Dirac velocity at charge neutrality, which resolves several puzzles in scanning tunneling spectroscopy and quantum oscillation experiments at once. For bilayer TMD, we show that applying small heterostrain generally leads to flat moiré bands that are highly tunable.

## I. INTRODUCTION

Recent experimental discoveries of correlated insulator and superconducting states in two-dimensional moiré materials including twisted bilayer graphene (TBG)[1–3] and graphene-boron nitride heterostructures [4] have stimulated tremendous interest in engineering flat or narrow bands to realize correlated electron phenomena. For this purpose, moiré superlattices provide a unique and highly tunable material platform. With a vast variety of 2d materials and heterostructures available [5], moiré superlattice systems may offer unprecedented advantages for studying many-body physics and realizing exotic states of matter[6–27].

Moiré patterns appear ubiquitously in layered 2d materials with a slight mismatch in the lattice orientation and/or lattice constant of the layers. Aside from twist angles between layers, heterostrain—which refers to relative strains between layers—provides an alternative way to create and modify moiré patterns[28]. Using strain to tune moiré bands may have practical advantages than using twist angle. Strain can be controlled *in situ* via piezoelectric substrate. The possibility of tuning band structure and achieving partially flat band with strain has recently been studied in graphene (see for example [29, 30]) and surface states of topological crystalline insulator [31].

In this work, we systematically study the effect of heterostrain on moiré band structures for homobilayer systems including bilayer graphene and transition metal dichalcogenides (TMD). Our motivation is twofold. First, (unintentional) heterostrain is ubiquitous in TBG samples [32, 33], likely due to the interaction with the substrate. However, the effect of heterostrain on the flat bands in magic-angle TBG is not well understood. We find a small amount of strain dramatically changes the low-energy band structure, which resolves several puzzles in transport and spectroscopy experiments all at once. Second, engineering a *tunable flat band* system by strain provides unprecedented opportunities for future studies. To that end, we focus on bilayer TMDs and show the conditions for flat moiré bands induced by heterostrain

without any twist.

This work consists of two parts. In the first part, we show that heterostrain in TBG significantly increases the energy separation between moiré conduction and valence bands, for instance, to  $\sim 30\text{meV}$  with 0.5% uniaxial heterostrain as commonly observed in STM experiments [32, 33]. The energy separation of the van Hove singularities saturates over a range of twist angles  $\theta$  around  $1^\circ$ , rather than being extremely sensitive to small deviation from the magic angle in the unstrained case. These findings explain the unexpectedly large separation of van Hove singularities observed in the recent STM experiments on TBG with  $\theta \sim 1^\circ$  [32, 33].

Moreover, in the presence of heterostrain, the conduction and valence bands of the same valley remain connected by two Dirac points near charge neutrality. However, due to the lowered symmetry, these two Dirac points are away from mini-Brillouin zone corners and are no longer degenerate in energy. This effect may explain the 4 (instead of 8) fold Landau level degeneracy[52] observed in the transport experiments[1–3]. We also show that heterostrain sets a lower bound for the Dirac velocity, preventing it from vanishing at the magic angle. The Dirac velocity is anisotropic and on the order of  $0.14v_F$  for a small heterostrain of 0.3%-0.6% ( $v_F$  being the bare Dirac velocity of monolayer graphene). This value is comparable to the one inferred from quantum oscillation and capacitance measurements[1].

In the second part, we study bilayer TMDs with volume preserving heterostrain. In contrast to the twisted bilayer graphene, we find that nearly flat moiré bands are generically present near the top of valence band in heterostrained bilayer TMD without fine tuning and without the need of twist. In addition, the moiré band gap and band structure are highly tunable by strain, pressure and displacement field, thus providing an ideal platform for correlation-driven phenomena.

## II. 2-DIMENSIONAL STRAIN AND GENERAL CONTINUUM MODELS

In this section, we consider the general geometrical description of 2-dimensional strain. Mathematically, the coordinate transformation in 2d can be written as

$$\mathbf{r}' = (\mathbb{I} + \mathcal{E})\mathbf{r} + \mathbf{d}_0, \quad (1)$$

---

<sup>\*</sup>Electronic address: [zbi@mit.edu](mailto:zbi@mit.edu)

where  $\mathbf{d}_0$  is a 2-dimensional vector that parametrizes the displacement and  $\mathcal{E}$  is an arbitrary 2-dimensional matrix that contains the strain and rotation. In the small deformation limit,  $\mathcal{E}$  can be written as

$$\mathcal{E} \cong \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} - \theta \\ \epsilon_{xy} + \theta & \epsilon_{yy} \end{pmatrix} = \mathcal{S}(\epsilon) + \mathcal{T}(\theta), \quad (2)$$

where the symmetric part, labeled by  $\mathcal{S}(\epsilon)$ , represents the strain, and the anti-symmetric part, labeled by  $\mathcal{T}(\theta)$ , represents the rotation. We emphasize that strain for 2d materials, different from the rotation angle which is a single parameter, is characterized by a 2 by 2 symmetric matrix, which has 3 free parameters. Therefore, it provides more possibilities to engineer the superlattice structures for layered 2d materials.

Let us review the effects of geometric deformation on the properties of monolayer graphene. For free monolayer graphene, we define  $\mathbf{A}_1 = a(1, 0)$  and  $\mathbf{A}_2 = a(1/2, \sqrt{3}/2)$  to be the primitive lattice vectors. Correspondingly, the reciprocal lattice vectors are  $\mathbf{G}_1 = \frac{2\pi}{a}(1, -1/\sqrt{3})$  and  $\mathbf{G}_2 = \frac{2\pi}{a}(0, 2/\sqrt{3})$ .  $\mathbf{K}_\pm = \mp(2\mathbf{G}_1 + \mathbf{G}_2)/3$  are referred as two valley points. The low energy description of monolayer graphene contains two massless Dirac fermions at  $\mathbf{K}_+$  and  $\mathbf{K}_-$  points with spin degeneracy. Geometrically, a deformation  $\mathcal{E}$  changes the shape of the unit cell as well as the Brillouin zone. Mathematically the rescaled primitive and reciprocal lattice vectors are

$$\mathbf{A}'_i = (\mathbb{I} + \mathcal{E})\mathbf{A}_i, \quad \mathbf{G}'_i \cong (\mathbb{I} - \mathcal{E}^T)\mathbf{G}_i. \quad (3)$$

A generic  $\mathcal{E}$ , including both nonzero strain  $\mathcal{S}$  and rotation  $\mathcal{T}$ , breaks almost all the point group symmetry of the lattice except the inversion  $C_{2z}$ . In addition to the geometric effect, the strain  $\mathcal{S}$  adjusts the distances between atoms, which leads to differences in hopping matrix elements for nearest carbon atoms along different directions. As a result, the locations of the low energy Dirac fermions are shifted away from the rescaled valley points  $\mathbf{K}'_\pm \cong (\mathbb{I} - \mathcal{E}^T)\mathbf{K}_\pm$ . Within a simple two center approximation  $t(r) \sim t_0 e^{\beta(r/a_0-1)}$ , the shift is proportional to the strain in the small strain limit and can be described by an effective gauge connection for the low energy Dirac fermions[34]:

$$\mathbf{A} = \frac{\sqrt{3}}{2a}\beta(\epsilon_{xx} - \epsilon_{yy}, -2\epsilon_{xy}) \quad (4)$$

The two Dirac fermions carry opposite charges under this fictitious gauge field. Therefore, the positions of the two Dirac fermions in momentum space are given by

$$\mathbf{D}_\xi = (\mathbb{I} - \mathcal{E}^T)\mathbf{K}_\xi - \xi\mathbf{A}, \quad (5)$$

where  $\xi = \pm$  labels the two valleys. The hopping modulus factor  $\beta$  is a dimensionless parameter determined by the intrinsic properties of the material. The approximate value of  $\beta$  is estimated by first principle calculation for graphene,  $\beta_g \cong 3.14$ [34].

The effects of geometric deformation for monolayer TMDs are similar as for monolayer graphene. For unstrained monolayer TMD, the low energy theory can be modeled by two *massive* Dirac fermions[35, 36] at  $\mathbf{K}_\pm$  points. As a result of the large spin-orbital coupling, the valence bands near two valleys carry opposite spins. Generic deformations reshape the unit cell and the Brillouin zone and break all the point group symmetries. It also shifts the Dirac fermions away from the rescaled valley points  $\mathbf{K}'_\pm$ . The shift is again described by Eq. 4 and 5. First principle calculations suggest for WSe<sub>2</sub>, the hopping modulus factor is  $\beta_{WSe_2} \cong 2.30$ [37].

In the rest of the paper, we consider homobilayers systems starting with *AA stacking* and then apply small twist and heterostrain. The coordinate transformations of the two layers can be described by two deformation matrices  $\mathcal{E}_1$  and  $\mathcal{E}_2$ . The rescaled reciprocal lattice vectors for the two layers are

$$\mathbf{G}'_i \cong (\mathbb{I} - \mathcal{E}_1^T)\mathbf{G}_i, \quad \mathbf{G}''_i \cong (\mathbb{I} - \mathcal{E}_2^T)\mathbf{G}_i. \quad (6)$$

Such geometrical deformation generates a moiré superlattice whose reciprocal lattice vectors  $\mathbf{g}_i$ 's and primitive lattice vectors  $\mathbf{a}_i$ 's are given by

$$\mathbf{g}_i \cong \mathbf{G}'_i - \mathbf{G}''_i = \mathcal{E}^T \mathbf{G}_i, \quad \mathbf{a}_i \cong \mathcal{E}^{-1} \mathbf{A}_i. \quad (7)$$

respectively, where  $\mathcal{E} = \mathcal{E}_2 - \mathcal{E}_1$  is the relative deformation matrix. Physical properties only depend on the relative deformation  $\mathcal{E}$  in the small twist and strain limit. For all the calculations in this paper, we assume that  $\mathcal{E}_2 = -\mathcal{E}_1 = \frac{1}{2}\mathcal{E}$ , namely the two layers are rotated and strained oppositely with the same magnitude.

In the limit where the moiré superlattice constant is much larger than the atomic scale, the low energy electronic structure can be effectively captured by the continuum model[38–40]. The spirit of the continuum model is the same as the nearly free electron approximation for band structures in solid state physics. Essentially, one should take the low energy bare dispersions of the 2d materials and perturb them by the periodic moiré superlattice. A general continuum model for bilayer systems can be schematically written as[6, 7, 38]

$$H = \begin{pmatrix} h_1(\mathbf{k}) + V_1(\mathbf{r}) & T(\mathbf{r}) \\ T^\dagger(\mathbf{r}) & h_2(\mathbf{k}) + V_2(\mathbf{r}) \end{pmatrix}, \quad (8)$$

where  $h_1(\mathbf{k})$  and  $h_2(\mathbf{k})$  are the bare dispersions of layer 1 and 2 respectively. The  $T(\mathbf{r})$  describes the spatial dependent interlayer tunneling and the  $V_1(\mathbf{r})$  and  $V_2(\mathbf{r})$  describe the intralayer potential induced by the moiré superlattice. The  $T(\mathbf{r})$  and  $V(\mathbf{r})$  vary between different materials and are also dependent on extrinsic conditions such as pressure[3]. In the following sections, we will focus our discussions on two cases 1° twisted bilayer graphene with uniaxial heterostrain and 2° bilayer WSe<sub>2</sub> with volume preserving heterostrain.

### III. TWISTED BILAYER GRAPHENE WITH UNIAXIAL HETEROSTRAIN

#### A. The continuum model

In bilayer graphene, the bare low energy dispersion for each layer contains two massless Dirac fermions, namely

$$h_l(\mathbf{k}) = \sum_{\xi=\pm} -\hbar v_F/a[(\mathbb{I} + \mathcal{E}_l^T)(\mathbf{k} - \mathbf{D}_{l,\xi})] \cdot (\xi\sigma^x, \sigma^y), \quad (9)$$

where  $l = 1, 2$  labels the two layers,  $\xi = \pm$  labels the two valleys. The  $\sigma$  matrices act on the pseudospin or sublattice degrees of freedom. Each valley also has two-fold spin degeneracy.  $\mathbf{D}_{l,\xi}$  denotes the location of the Dirac fermion as given in Eq. 5. The Fermi velocity in monolayer graphene is estimated  $v_F \cong 10^6$  m/s[41], which gives the kinetic energy scale  $\hbar v_F/a \cong 2.68$  eV.

In this section, we restrict ourselves to twisted bilayer graphene with uniaxial heterostrain. The uniaxial heterostrain refers to a class of strain where the bilayer system is relatively stressed along one direction and unstressed on the perpendicular direction. Scanning tunneling microscope (STM) experiments have indicated such type of heterostrain in twisted bilayer graphene samples[32, 33, 42, 43]. We anticipate that many generic features of the uniaxial strain should also apply for more general form of strain. Geometrically, uniaxial strain can be described by two parameters, namely the strain magnitude  $\epsilon$  and the strain direction  $\varphi$ . The strain tensor can be written as the following[32]

$$\begin{aligned} \mathcal{S}_{ua} &= R(\varphi)^{-1} \begin{pmatrix} -\epsilon & 0 \\ 0 & \nu\epsilon \end{pmatrix} R(\varphi) \\ &= \epsilon \begin{pmatrix} -\cos(\varphi)^2 + \nu \sin(\varphi)^2 & (1 + \nu) \cos(\varphi) \sin(\varphi) \\ (1 + \nu) \cos(\varphi) \sin(\varphi) & -\sin(\varphi)^2 + \nu \cos(\varphi)^2 \end{pmatrix}, \end{aligned} \quad (10)$$

where  $\nu = 0.16$  is the Poisson ratio for graphene. The combination of twist and strain is described by the relative deformation matrices  $\mathcal{E} = \mathcal{T}(\theta) + \mathcal{S}_{ua}(\epsilon, \varphi)$ .

For graphene system,  $V_l(\mathbf{r})$  is parametrically smaller than the interlayer tunneling, therefore we set it to zero. The interlayer tunneling has the following form (following the convention in, for example, [7]),

$$\begin{aligned} T(\mathbf{r}) &= \begin{pmatrix} u & u' \\ u' & u'' \end{pmatrix} + \begin{pmatrix} u & u' \omega^{-\xi} \\ u' \omega^\xi & u'' \end{pmatrix} e^{i\xi \mathbf{g}_1 \cdot \mathbf{r}} \\ &\quad + \begin{pmatrix} u & u' \omega^\xi \\ u' \omega^{-\xi} & u'' \end{pmatrix} e^{i\xi (\mathbf{g}_1 + \mathbf{g}_2) \cdot \mathbf{r}}, \end{aligned} \quad (11)$$

where  $\omega = e^{i2\pi/3}$ . The effect of lattice corrugation in TBG on moiré band structure can be included in the continuum model by choosing renormalized tunneling amplitudes  $u, u'$  and  $u''$  [7], which depend on heterostrain and twist angle. For simplicity, in our calculation we assume the two graphene layers are unrelaxed and use the tunneling parameters  $u = u' = u'' = 110$  meV[38][53].

An apparent geometric effect of the combination of twist and heterostrain is that the moiré superlattice is no

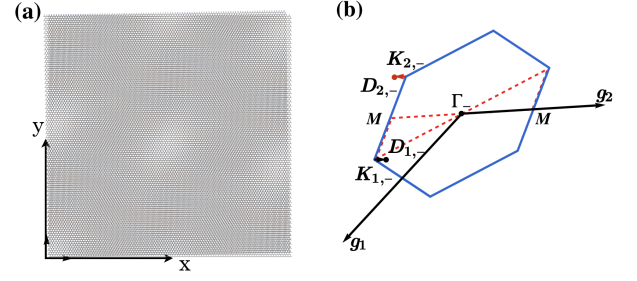


FIG. 1: (a) The moiré pattern of bilayer graphene with twist angle  $\theta = 1.05^\circ$  and uniaxial strain  $\epsilon = 0.7\%$ ,  $\varphi = 0^\circ$ . Notice that the moiré superlattice is not a regular triangular lattice. The moiré dots are elliptical which is commonly observed in STM experiments [32, 33]. (b) The moiré Brillouin zone with the same geometrical parameters. In the plot, for clarity, we only label the rescaled  $\mathbf{K}_-$  points from the two layers. Correspondingly, the red and black dots are the position of the shifted Dirac points from  $\mathbf{K}_-$  valley. The other valley can be obtained by time reversal operation. The red dashed line is the momentum path for the band structure plots.

longer a regular triangular lattice. An example of moiré pattern generated with  $\theta = 1.05^\circ$ ,  $\epsilon = 0.7\%$  and  $\varphi = 0^\circ$  is shown in Fig. 1(a). The uniaxial strain makes the moiré dots elliptical, which is visible in local measurements such as STM[32, 33]. As a consequence, the Brillouin zone of the moiré superlattice is a distorted hexagonal as shown in Fig. 1(b). In order to compare results for different twists and strains, we will always stretch the irregular hexagonal Brillouin zone to be regular.

A typical band structure of twisted bilayer graphene near magic angle with small uniaxial heterostrain is shown in Fig. 2. Three generic features are worth noticing in these plots. First, the energy separation between conduction and valence bands, as seen from the separation of van Hove singularities, is significantly enlarged compared to the case with no strain. Interestingly, these two bands remain rather flat for most parts of the Brillouin zone. This finding agrees with the large splitting of van Hove singularities of the conduction and valence bands observed in the scanning tunneling spectroscopy (STS) [32, 33]. We will study in detail of the evolution of the bandwidth with respect to the geometrical parameters  $\theta$ ,  $\epsilon$  and  $\varphi$ .

Second, the conduction and valence bands remain connected by two Dirac crossings in each valley, which is due to the  $C_{2z}\mathcal{T}$  symmetry. However, these two Dirac crossings are now located at generic points away from Brillouin zone corners because of the lack of three-fold rotational symmetries in the presence of strain. In addition, the energies of the two Dirac points within a valley are shifted by the heterostrain and are no longer degenerate. This may explain the observed 4-fold (instead of 8-fold) Landau level degeneracy in the experiments[1–3] near charge neutrality. We note that the energy offset of the two Dirac points leads to a finite density of electrons

and holes at charge neutrality. Nonetheless, the residual electron/hole density is found to be rather small: around  $\sim 3 \times 10^{10} \text{cm}^{-2}$  (including valley and spin degeneracy) for 0.6% strain. We shall also investigate the dependence of the Dirac fermion energy splitting on the system parameters.

Finally, compared to the unstrained case, the Dirac velocity is found to be greatly enhanced—reaching  $0.14v_F$  for a small strain of 0.6%, and is anisotropic due to the lowered symmetry. Moreover, we find that heterostrain sets a lower bound for the Dirac velocity, preventing it from vanishing around magic angle. The enhanced Dirac velocity found here is comparable to the one inferred from quantum oscillation and capacitance experiments [1].

### B. The bandwidth and the Dirac point shift

Now we study  $1^\circ$  the total bandwidth of the conduction and valence bands and  $2^\circ$  the energy shift of the two Dirac crossings within a valley as a function of the parameters  $\theta$ ,  $\epsilon$  and  $\varphi$ . We are interested in the regime of twist angle  $\theta$  around  $1^\circ$ . Experimental data indicate that the uniaxial heterostrain in twisted bilayer graphene samples can vary from 0.1% to 0.7% [32]. In addition, it is easy to convince ourselves that the system is periodic for  $\varphi \rightarrow \varphi + \pi/3$  because of the symmetry of the unstrained system. Therefore, we can restrict our studies to  $\varphi \in [0, \pi/3]$ .

Firstly, let us fix  $\theta = 1.05^\circ$  and  $\epsilon = 0.3\%$  and plot the density of states the conduction and valence bands for several values of  $\varphi$  as shown in Fig. 3. We observe many subtle changes in the structure of van Hove singularities within each bands as the direction of the strain is varied. For instance,  $\varphi = 0^\circ$  has three density of state peaks in each band. However, for  $\varphi = 20^\circ \sim 30^\circ$ , there is only one prominent peak in each band, which is consistent with the experimental results [32, 33]. Nonetheless, the total bandwidth stays approximately *constant*.

Secondly, we study the total bandwidth measured by the separation of the van Hove singularities in the conduction and valence bands, labeled by  $\Delta$ . Fig. 4 shows the evolution of  $\Delta$  as a function of  $\theta$  for several different  $\epsilon$ 's at fixed strain direction  $\varphi = 25^\circ$ . An interesting feature is that the bandwidth is *insensitive* to the twist angle for angles near the magic angle. We also observe that the saturation value of the bandwidth is linearly proportional to the magnitude of the strain as shown in Fig. 5, which suggests the strain effect is *dominant* near the magic angle. In certain sense, the heterostrain provides an intrinsic kinetic energy cutoff,  $\Lambda \sim \epsilon \hbar v_F$ , for the system and stabilizes the bandwidth.

Finally, we study the energy shift, labeled by  $\delta$ , of two Dirac points within a valley as a function of  $\epsilon$  and  $\varphi$  at fixed twist angle. As shown in Fig. 6, the Dirac point shifts highly depend on the direction of the strain. We also find that the maximal value of the Dirac point shift is proportional to the magnitude of the strain. To our

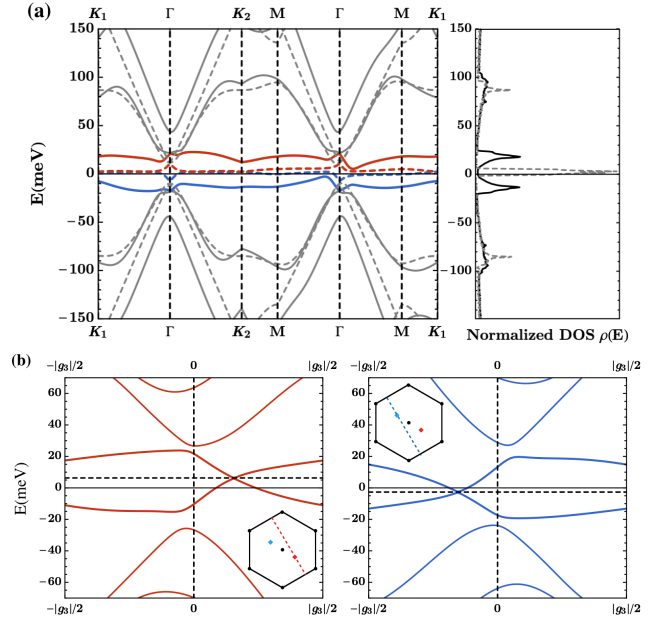


FIG. 2: In (a), we show the band structure (left) and the density of states (right) for the  $\mathbf{K}_-$  valley of twisted bilayer graphene at twist angle  $\theta \cong 1.05^\circ$  with (solid line) and without (dashed line) uniaxial heterostrain  $\epsilon = 0.6\%$ ,  $\varphi = 30^\circ$ . The heterostrain increases the separation between the valence and conduction bands, while keeps them rather flat in most area of the Brillouin zone. In (b), we show that the two bands are still connected to each other through two Dirac points. With heterostrain, the positions of the two Dirac crossings are no longer at the corner of the moiré Brillouin zone, as shown in the inserts of Fig (b). The hexagons represent the stretched Brillouin zone and the blue/red dots are the positions of two Dirac crossings. We plot the band dispersions along the red/blue dashed lines in the Brillouin zone. The conduction and valence bands are rather flat in most part of the momentum space except near the Dirac crossings. The energy shift between the two Dirac points is  $\sim 9 \text{meV}$ . As seen from (b), the Dirac velocity is anisotropic and reaching  $0.15v_F$ , which is much larger than the unstrained case (about  $0.018v_F$  with the current parameters). This velocity is comparable to the one observed in transport and capacitance experiments [1]. Since the Dirac points are shifted in energy, a finite electron/hole Fermi surfaces appear at charge neutrality. Due to the enhanced Dirac velocity, the size of the Fermi pocket is found to be much smaller than the size of the Brillouin zone. The estimated electron/hole density (including the valley and spin degeneracy) is  $\sim 3 \times 10^{10} \text{cm}^{-2}$  for current parameters.

surprise, a quite small heterostrain such as  $\epsilon = 0.1\%$  can create a considerable splitting of the Dirac fermions (on the order of  $5 \text{meV}$ ).

Although a complete analytical understanding of heterostrain effects found in our calculation is currently lacking, we provide some intuition for the observed features in the band structures. In the unstrained twisted bilayer near the magic angle, the nearly flat energy dispersions



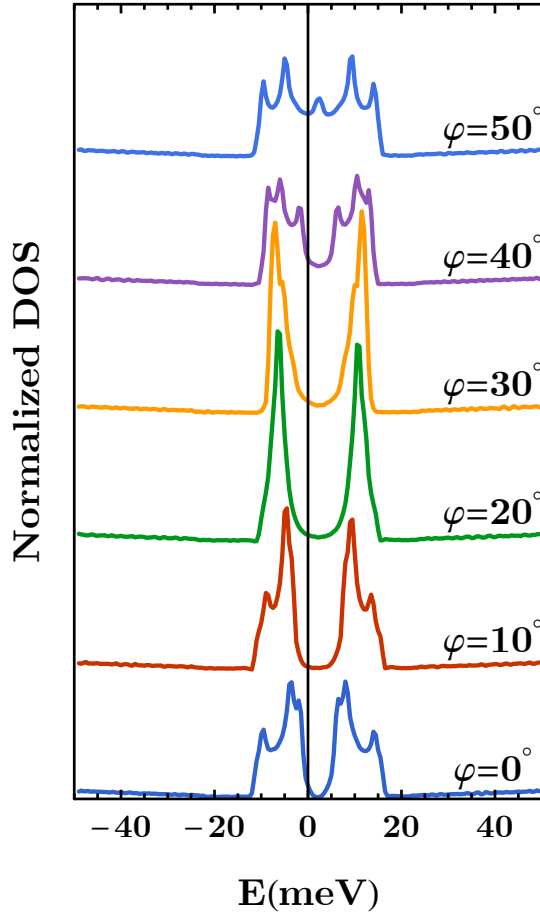


FIG. 3: Normalized density of state plot with fixed  $\theta = 1.05^\circ$ ,  $\epsilon = 0.3\%$  and varying  $\varphi$ . (The curves are relatively shifted to make the plot clear.) We observe many subtle transitions in the structure of the van Hove singularities as the direction of the strain is changes. However, the bandwidth of the middle two bands stays approximately constant. (The density of states has only one prominent peak in each band for  $\varphi$  near  $20^\circ \sim 30^\circ$ .)

arise from subtle interference effects between the mismatched energy dispersion within each layer and electron hopping between layers. For example, the state at  $\mathbf{K}_-$  from layer-1 would be coupled through the moiré inter-layer tunneling to three states from layer-2 with the same energy. Similarly, the state at  $\mathbf{K}_-$  from layer-2 is also coupled to three equal energy states in layer-1. We can crudely think of the magic angle is tuning the bilayer system to a critical situation where the kinetic energy scale vanishes. However, such “interference condition” is explicitly violated by the heterostrain, which introduces a natural kinetic energy cutoff,  $\Lambda \sim \epsilon \hbar v_F$ , to the system. It is easy to see the energies of the three states from layer-2 that can couple to  $\mathbf{K}_-$  state in layer-1 are now shifted precisely by  $\Lambda \sim \epsilon \hbar v_F$  because of the heterostrain. This lifting of perfect interference is presumably the cause for the enlarged bandwidth and the Dirac fermion energy

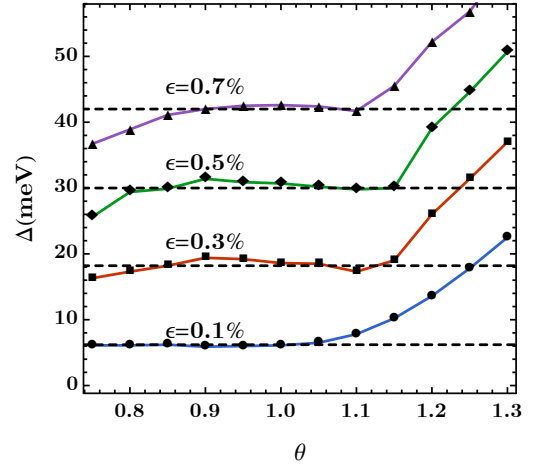


FIG. 4: The splitting of the van Hove singularities,  $\Delta$ , as a function of the twist angle  $\theta$ . We fixed the strain direction  $\varphi = 25^\circ$  and different curves represent different strain magnitudes. The splitting of the van Hove singularities stay approximately constant for nearly magic twist angles.

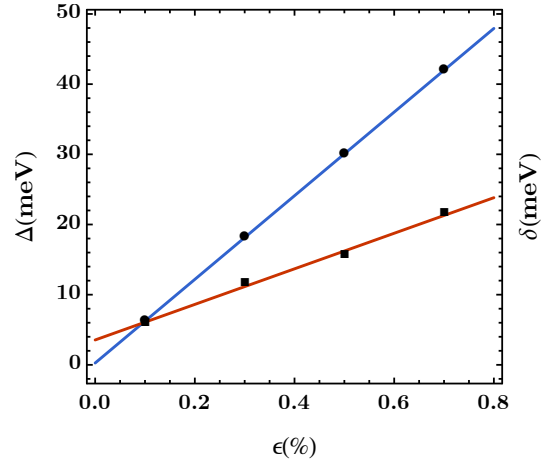


FIG. 5: The blue line shows the saturation values of the van Hove singularity splitting, labeled by  $\Delta$ , as a function of the magnitude of the uniaxial strain. The red line shows the maximal energy shift of the two Dirac crossings within a valley,  $\delta$ , as a function of strain magnitude. The fitting shows that both quantities are linearly proportional to the magnitude of uniaxial strain.

shift.

### C. Higher order van Hove singularities

As shown in Ref. [8], with a single tuning parameter such as the twisted angle, one can achieve higher order van Hove singularities in bilayer graphene, which are perfect playgrounds for correlation-driven physics. Here we

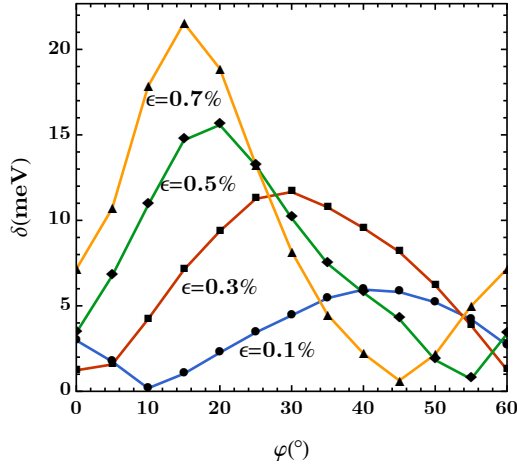


FIG. 6: The energy shift between two Dirac points in a valley, labeled by  $\delta$ , as a function of the direction of the strain. In this plot, we fix  $\theta = 1.05^\circ$  and choose 4 different strain magnitudes  $\epsilon = 0.1\%$  (blue),  $0.3\%$  (red),  $0.5\%$  (green),  $0.7\%$  (yellow). We notice that the energy shift of the Dirac fermions is very sensitive to the direction of the strain. The maximal value of the shift is linear proportional to the strain magnitude.

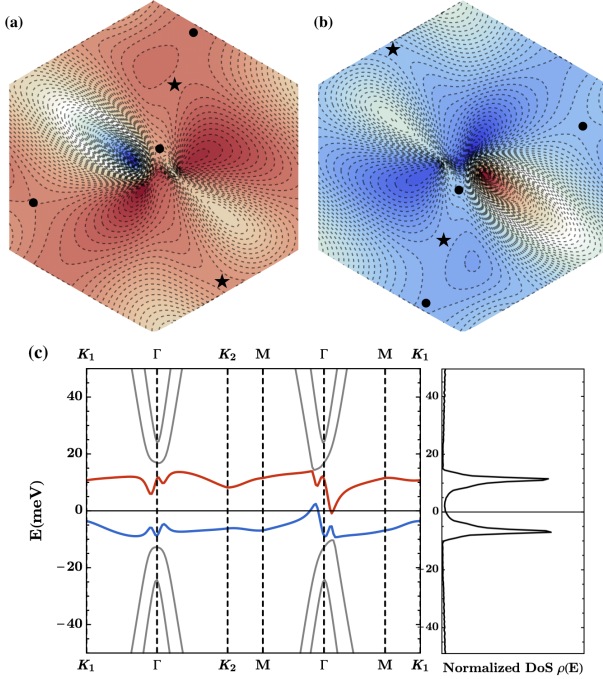


FIG. 7: (a)/(b) The equal energy contour plots for the valence/conduction bands with parameters  $\theta = 1.05^\circ$ ,  $\epsilon = 0.3\%$  and  $\varphi = 25^\circ$ . We can identify multiple ordinary (dot) and higher order (star) van Hove singularities. Since these van Hove singularities happen to be nearby in energy, they contribute to a single enhanced density of state peak in each band as shown in (c).

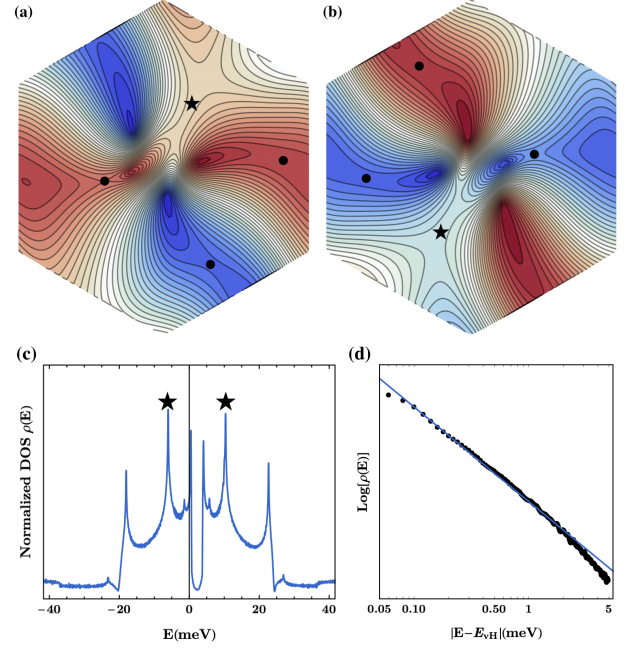


FIG. 8: (a)/(b) The equal energy contour plots for the valence/conduction bands with parameters  $\theta = 1.05^\circ$ ,  $\epsilon = 0.53\%$  and  $\varphi = 50^\circ$ . We observe ordinary and higher order van Hove singularities in both valence and conduction bands. They contribute to the sharp peaks in the density of states in (c). In (d) we zoom in near the density of state peak at  $E_{vH} \cong 10.4\text{meV}$  and make a  $\log - \log$  plot for  $\rho(E)$  vs  $|E - E_{vH}|$  for  $E - E_{vH} > 0$ . The linear fitting indicates that the density of states has a power law divergence near the van Hove singularity, namely  $\rho(E) \sim |E - E_{vH}|^{-\nu}$  with  $\nu \cong 0.255$ . This is the key feature for the higher order van Hove singularities.

show that heterostrain is another effective way to generate higher order van Hove singularities. To demonstrate this, we show two examples in Fig. 7 and 8. We find multiple ordinary and higher order van Hove singularities in both conduction and valence bands in these examples. The energies of the van Hove singularities are not required to be the same in general because of the low symmetry of the heterostrained system. Therefore, we may observe multiple peaks in the density of states for certain conditions as shown in Fig. 8.

## IV. HETEROSTRAIN IN BILAYER TRANSITION METAL DICHALCOGENIDE

### A. Heterostrain induced flat bands

For monolayer TMD, the model hamiltonian for one valley can be written as the following[35–37]

$$h_l(\mathbf{k}) = \sum_{\xi=\pm} -\hbar v_F/a[(\mathbb{I} + \mathcal{E}_l^T)(\mathbf{k} - \mathbf{D}_{l,\xi})] \cdot (\xi \sigma^x, \sigma^y) + \frac{m}{2}(\sigma^z + 1), \quad (12)$$

where  $\sigma$  matrices act on orbital space of  $|d_{z^2}\rangle$  and  $(|d_{x^2-y^2}\rangle + i\xi|d_{xy}\rangle)/\sqrt{2}$  orbitals, and there is no spin degeneracy for the valence band because of the large spin-orbital coupling. This hamiltonian generally applies for a large class of TMD materials[35–37]. In this paper, we take WSe<sub>2</sub> as a representative for TMD materials. For WSe<sub>2</sub>, it is estimated that  $\hbar v_F/a \cong 1.1\text{eV}$  and  $m \cong 1.2\text{eV}$ . Similar as the graphene case, with strains, the Dirac points are shifted and their locations in momentum space are again given by Eq. 5 and 4. It is found from first principle that the drift parameter  $\beta$  is around 2.3 for WSe<sub>2</sub>[37]. Of course, it is hard to pin down the exact value of the  $\beta$  and it also varies for different TMD materials. Therefore, we treat  $\beta$  as a potential tuning parameter to extract general features of the moiré band structure.

In this section, we consider bilayer transition metal dichalcogenide system with only heterostrains for simplicity. It turns out that the bilayer TMD system is a desirable platform for engineering flat bands with strains. We restrict ourselves to a class of heterostrain which is called volume preserving strain. The general form is given by,

$$S_v = \epsilon \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}. \quad (13)$$

Physically,  $S_v$  describes that the material is strained by  $\epsilon$  along  $\varphi$  direction and by  $-\epsilon$  along the normal direction. We consider the bilayer system starting from *AA stacking* configuration, where the inversion symmetry is not present. The *AB stacking* configuration is an interesting case which we will address in a separate paper. Applying a generic  $S_v$  explicitly breaks all the point group symmetries of the system.

For bilayer WSe<sub>2</sub> moiré superlattices starting from *AA stacking* configuration, we need to include both the interlayer tunneling and the intralayer potential in Eq. 8. For the tunneling matrix, the inter-orbital tunneling  $u'$  is assumed to be zero because of the large gap between the two orbitals. The intra-orbital tunneling take the same form as in Eq. 11 and it is estimated in Ref.[44] that  $u \cong 1.1\text{meV}$ ,  $u'' \cong 9.7\text{meV}$  for bilayer WSe<sub>2</sub>. The intralayer potential has the following form[44]

$$V_l(\mathbf{r}) = \sum_{i=1,2,3} \begin{pmatrix} V_c e^{i(\mathbf{g}_i \cdot \mathbf{r} + (-1)^l \phi_c)} & 0 \\ 0 & V_v e^{i(\mathbf{g}_i \cdot \mathbf{r} + (-1)^l \phi_v)} \end{pmatrix} + h.c., \quad (14)$$

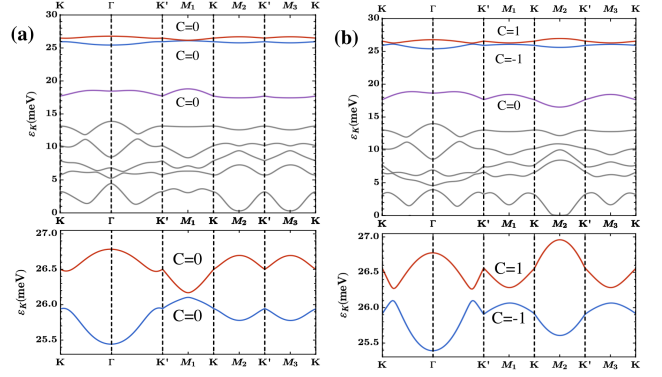


FIG. 9: We focus on the band structures for  $\mathbf{K}_-$  valley of bilayer WSe<sub>2</sub> near the top of the valence band. (a)/(b) show band structures with volume preserving heterostrain at  $\epsilon = 1.5\%$  and  $\varphi = 0^\circ/30^\circ$  respectively. (We take  $\beta_{TMD} \cong 2.3$ .) We notice that the first three bands are quite flat. The top two bands are close in energy and well separated from other bands with gap  $\Delta \sim 8\text{meV}$ . We observe that for  $\varphi = 0$ , the first two valence bands are topologically trivial, while for  $\varphi = 30^\circ$  they carry  $\pm 1$  chern number.

where  $\mathbf{g}_3 = -(\mathbf{g}_1 + \mathbf{g}_2)$ . For WSe<sub>2</sub>, the parameters are  $V_c \cong 6.8\text{meV}$ ,  $V_v \cong 8.9\text{meV}$ ,  $\phi_c \cong 89.7^\circ$  and  $\phi_v \cong 91^\circ$ . These data vary slightly for other TMD materials. Thus, we will focus on generic features of the electron band structure which do not rely much on the precise values of these parameters.

We show the band structure for bilayer WSe<sub>2</sub> with heterostrain calculated from the continuum model in Fig. 9. We find generic flat bands near the top of the valence band for small strain magnitude. The top two valence bands are close in energy and have very small bandwidth compared to the large gap separating them to the rest of the spectrum. The third band is also quite flat and energetically separated from other bands. Increasing strain can enlarge the bandwidth and band gap.

The strain angle  $\varphi$  provides an additional knob for engineering the band structure. Now we focus on the top two valence bands. In Fig. 10 (c), we show the phase diagram as a function of  $\beta$  and  $\varphi$  (the phase diagram is not sensitive to the magnitude of the strain  $\epsilon$  for small  $\epsilon$ ). Interestingly, we find an alternating pattern of topological and trivial bands as a function of  $\varphi$ . The threefold periodicity of the phase diagram in  $\varphi$  is due to the emergent threefold rotational symmetry of interlayer coupling (11) and intralayer potential (14). We show the band structure of a critical point at  $\beta_c \cong 2.10$  and  $\varphi = 0$ . The topological phase transition closes the gap with a Dirac dispersion at one of the  $M$  points in the Brillouin zone.

It is worth to mention the displacement field is a good way to control the relative position of the first three bands. With considerable interlayer bias, we can make one band well separated from other bands on the top of the spectrum. In addition, this band is topologically trivial. An example of such band structure is shown in Fig.



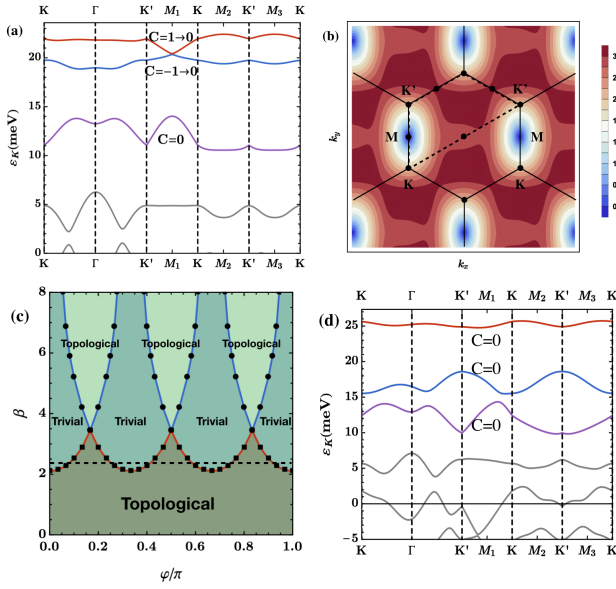


FIG. 10: (a) shows the band structure of  $\mathbf{K}_-$  valley with parameters  $\beta = 2.10$ ,  $\epsilon = 2\%$  and  $\varphi = 0$ . The system is at a topological phase transition. A Dirac crossing is found at one of the  $\mathbf{M}$  points in the Brillouin zone. The three  $\mathbf{M}$  points are not equivalent because of the lack of  $C_3$  rotation symmetry. (b) shows the contour plot of the gap between the first two bands with the same parameters. (c) Phase diagram of the top two valence bands as a function of  $\beta$  and  $\varphi$ . There is an interesting pattern of topological and trivial states (topological state refers to  $\pm 1$  chern number for the top two bands and trivial state refers to 0 chern number) as we tune the angle  $\varphi$  from 0 to  $\pi$ .  $\varphi = 0$  and  $\pi$  are physically equivalent by an exchange of the two layers. The dash line is the estimated value of  $\beta \cong 2.30$  for  $\text{WSe}_2$ . The phase diagram is obtained with  $\epsilon = 2\%$ . However, the phase diagram is *insensitive* to the strain magnitude for small strain  $< 3\%$ . (d) The band structure for  $\epsilon = 2\%$  and  $\varphi = 0$  with an interlayer bias  $\Delta V = 10\text{meV}$ . The displacement field separates a nearly flat band on the top of the spectrum.

10 (d). As expected, the wavefunction of the top band is mostly polarized in one layer.

### B. Pressure tunned topological transition

Pressure on the bilayer systems changes the interlayer distance and in return adjusts the relative ratio between the strength of the interlayer tunneling and intralayer potential energy, which is another way to design the moiré band structure. To demonstrate this effect, we consider twisted or heterostrained bilayer TMD with pressure. Pressure can be modeled by introducing a phenomenological parameter  $p$  in the model Eq. 8,  $T(\mathbf{r}) \rightarrow pT(\mathbf{r})$ , which represents the relative ratio of interlayer tunneling and intralayer potential. For twisted bilayer TMD at twist angle  $\theta \cong 1^\circ$ , we find multiple topological phase transitions between the top three moiré bands as we in-

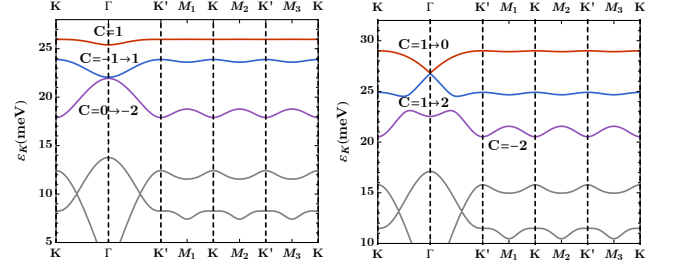


FIG. 11: We consider twisted bilayer  $\text{WSe}_2$  at  $\theta = 1^\circ$  with pressure. As the parameter  $p$  is tuned from 1 to 1.5, there are two successive topological phase transitions. (a) At  $p \cong 1.195$ , the second band and third band touches with quadratic dispersion at  $\Gamma$  point, which changes the Chern number by 2. (b) At  $p \cong 1.410$ , the first and second band close gap with a Dirac dispersion, which changes the Chern number by 1.

crease  $p$  from 1 to 1.5 as shown in Fig. 11. At  $p \cong 1.195$ , there is a quadratic band touching at  $\Gamma$  in moiré Brillouin zone between the second and third band which shifts the Chern numbers of the two bands by  $\pm 2$ . At  $p \cong 1.41$ , the top two bands touch at  $\Gamma$  with a Dirac dispersion. The top band becomes topologically trivial for  $p > 1.41$ . In this system, we can get topological flat bands with chern number  $\pm 2$ . Similar effects also appear in heterostrained cases.

### C. Effective model for the flat bands

With small heterostrains, for instance  $\epsilon \cong 1.5\%$  in Fig. 9, for each valley the two top valence bands which are close in energy and very flat (total bandwidth  $W \cong 1.4\text{meV}$ ), and they are well separated from the rest of the bands (gap  $\Delta \cong 8\text{meV}$ ). One can estimate the Coulomb interaction scale on the moiré superlattice  $V = e^2/(4\pi\epsilon a_M) \cong 4.2\text{meV}$ , with static dielectric constant of  $\text{WSe}_2$   $\epsilon \cong 15.3$ [45] and moiré superlattice constant  $a_M \cong 67a \cong 23.4\text{nm}$  for  $\epsilon_{xy} = 1.5\%$ . The parameters of the system are in the limit  $\Delta \gg V \gg W$ . In addition, the total Chern number for the top two valence bands is trivial. Therefore, we can propose an effective model just for the top two bands. Taking into account of the valley/spin degeneracy, the simplest guess is a 2-band Hubbard model on a triangular lattice, which can be schematically written as the following

$$H_{eff} = \sum_{\alpha=1}^2 \sum_{i,j} \sum_{v=\pm} t_{ij}^\alpha c_{i,\alpha,v}^\dagger c_{j,\alpha,v} + \sum_i \sum_{\alpha=1}^2 U_\alpha n_{i,\alpha} (n_{i,\alpha} - 1) + \dots, \quad (15)$$

where  $n_{i,\alpha} = \sum_{v=\pm} c_{i,\alpha,v}^\dagger c_{i,\alpha,v}$ . The two valleys are found to be degenerate in energy, therefore, have an approximate  $SU(2)$  symmetry in the small intervalley-scattering limit. **The main feature of our model is hopping terms are spatially anisotropic depending on the heterostrain.** Half filling the two bands may lead to interest-

ing correlated states, such as valley/spin polarized state, which spontaneously breaks time reversal symmetry.

Aside from the top two valence bands, we can also focus on the third band which is also quite flat and energetically separated from other bands. The Chern number is again trivial for this band with current parameters. The effective model for this band will be a single band triangular lattice  $SU(2)$  Hubbard model with **anisotropic hopping terms, which is different from the low energy models of twisted bilayer TMD**. Similar anisotropic model also describes the physics of the displacement field separated flat band in Fig. 10 (d). **Triangular lattice  $SU(2)$  Hubbard model model with anisotropy** is previously suggested to be the effective model of the quasi-two-dimensional organic charge transfer salts, which potentially host exotic spin liquid states[46, 47].

Finally, we emphasize that the Coulomb interaction scales with the strain as  $V \sim 1/a_M \sim \epsilon$  while the bandwidth of the moiré bands scales with the strain as  $W \sim (1/a_M)^2 \sim \epsilon^2$ . Therefore, different strains can produce Hubbard models with different  $t/U$  ratio.

## V. DISCUSSION

In this paper, we explore the effects of uniaxial heterostrain in twisted bilayer graphene. It is clear that uniaxial heterostrain generically broadens the bandwidth of the nearly flat bands at magic angle, which could be the reason for the observed large energy separation of van Hove singularities in the STM experiment[32, 33, 48, 49]. Large bandwidth may not be desirable for correlation physics. Therefore, future experiments should try to minimize the heterostrain between layers. On the other hand, heterostrain could be helpful to generate higher order van Hove singularities, which is a good platform to enhance correlation effects. A systematic study of the evolution of the van Hove singularities is an interesting but involved subject. Another natural question to ask is

whether there are situations in which heterostrain helps to flatten the band dispersions rather than to broaden the bands. To answer these questions, we need more comprehensive studies where optimization methods such as machine learning could be helpful. On the other hand, a full analytical understanding of the strain effect is also demanded for future study.

The other effects of heterostrain are shifting the energies of the two Dirac crossings within one valley and greatly enhancing the Dirac velocity. This could potentially explain the observed 4-fold Landau level degeneracy in magneto transport experiments[1–3, 50, 51]. One natural consequence of such scenario is that there will be finite electron/hole fermi surfaces at the charge neutrality point. Nonetheless, due to the enhanced Dirac velocity, the fermi surfaces are estimated within our continuum model to be really small and could be difficult to resolve with current experimental precision. To get a more accurate estimation theoretically, one needs to include various subtle effects such as lattice relaxations, which is beyond the scope of the current paper but a good subject for future investigations.

We also show heterostrain is a good way of creating and tuning flat bands in bilayer TMD systems. There is no requirement of a “magic” strain here. The band structure **and topology** is easily controlled by the strain direction, pressure and displacement field.

## Acknowledgement

We thank Abhay Pasupathy, Cory Dean, Matt Yankowitz, Eva Andrei and Brian LeRoy for stimulating discussions and communications. ZB is supported through Pappalardo fellowship at MIT. NY and LF are supported by DOE Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award DE-SC0018945. LF is partly supported by the David and Lucile Packard Foundation.

- 
- [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 EP (2018), URL <https://doi.org/10.1038/nature26154>.
  - [2] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Nature* **556**, 43 EP (2018), URL <https://doi.org/10.1038/nature26160>.
  - [3] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, *Science* (2019), URL <http://science.sciencemag.org/content/early/2019/01/25/science.aav1910>.
  - [4] G. Chen, L. Jiang, S. Wu, B. Lv, H. Li, K. Watanabe, T. Taniguchi, Z. Shi, Y. Zhang, and F. Wang, arXiv e-prints arXiv:1803.01985 (2018), 1803.01985.
  - [5] K. S. Novoselov, A. Mishchenko, A. Carvalho, and A. H. Castro Neto, *Science* **353**, aac9439 (2016), URL <http://science.sciencemag.org/content/353/6298/aac9439.abstract>.
  - [6] N. F. Q. Yuan and L. Fu, *Phys. Rev. B* **98**, 045103 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.045103>.
  - [7] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, *Phys. Rev. X* **8**, 031087 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevX.8.031087>.
  - [8] N. F. Q. Yuan, H. Isobe, and L. Fu, arXiv e-prints arXiv:1901.05432 (2019), 1901.05432.
  - [9] C. Xu and L. Balents, *Phys. Rev. Lett.* **121**, 087001 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevLett.121.087001>.
  - [10] C.-C. Liu, L.-D. Zhang, W.-Q. Chen, and F. Yang, *Phys. Rev. Lett.* **121**, 217001 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevLett.121.217001>.

- [org/doi/10.1103/PhysRevLett.121.217001](https://doi.org/10.1103/PhysRevLett.121.217001).
- [11] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, *Phys. Rev. X* **8**, 031089 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevX.8.031089>.
  - [12] L. Zou, H. C. Po, A. Vishwanath, and T. Senthil, *Phys. Rev. B* **98**, 085435 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.085435>.
  - [13] A. Thomson, S. Chatterjee, S. Sachdev, and M. S. Scheurer, *Phys. Rev. B* **98**, 075109 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.075109>.
  - [14] J. F. Dodaro, S. A. Kivelson, Y. Schattner, X. Q. Sun, and C. Wang, *Phys. Rev. B* **98**, 075154 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.075154>.
  - [15] X. Y. Xu, K. T. Law, and P. A. Lee, *Phys. Rev. B* **98**, 121406 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.121406>.
  - [16] D. M. Kennes, J. Lischner, and C. Karrasch, *Phys. Rev. B* **98**, 241407 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.241407>.
  - [17] Y. Cao, D. Chowdhury, D. Rodan-Legrain, O. Rubies-Bigordà, K. Watanabe, T. Taniguchi, T. Senthil, and P. Jarillo-Herrero, arXiv e-prints arXiv:1901.03710 (2019), 1901.03710.
  - [18] H. C. Po, L. Zou, T. Senthil, and A. Vishwanath, arXiv e-prints arXiv:1808.02482 (2018), 1808.02482.
  - [19] Y.-H. Zhang, D. Mao, Y. Cao, P. Jarillo-Herrero, and T. Senthil, arXiv e-prints arXiv:1805.08232 (2018), 1805.08232.
  - [20] J. Kang and O. Vafek, *Phys. Rev. X* **8**, 031088 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevX.8.031088>.
  - [21] J. Kang and O. Vafek, arXiv e-prints arXiv:1810.08642 (2018), 1810.08642.
  - [22] F. Guinea and N. R. Walet, *Proceedings of the National Academy of Sciences* **115**, 13174 (2018), ISSN 0027-8424, <https://www.pnas.org/content/115/52/13174.full.pdf>, URL <https://www.pnas.org/content/115/52/13174>.
  - [23] F. Wu, A. H. MacDonald, and I. Martin, *Phys. Rev. Lett.* **121**, 257001 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevLett.121.257001>.
  - [24] B. Padhi, C. Setty, and P. W. Phillips, *Nano Letters* **18**, 6175 (2018), pMID: 30185049, <https://doi.org/10.1021/acs.nanolett.8b02033>, URL <https://doi.org/10.1021/acs.nanolett.8b02033>.
  - [25] J. W. F. Venderbos and R. M. Fernandes, *Phys. Rev. B* **98**, 245103 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.245103>.
  - [26] M. Angeli, D. Mandelli, A. Valli, A. Amaricci, M. Capone, E. Tosatti, and M. Fabrizio, *Phys. Rev. B* **98**, 235137 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.235137>.
  - [27] B. L. Chittari, G. Chen, Y. Zhang, F. Wang, and J. Jung, *Phys. Rev. Lett.* **122**, 016401 (2019), URL <https://link.aps.org/doi/10.1103/PhysRevLett.122.016401>.
  - [28] L. Huder, A. Artaud, T. Le Quang, G. T. de Laisardière, A. G. M. Jansen, G. Lapertot, C. Chapelier, and V. T. Renard, *Phys. Rev. Lett.* **120**, 156405 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevLett.120.156405>.
  - [29] F. Guinea, *Solid State Communications* **152**, 1437 (2012), URL <http://www.sciencedirect.com/science/article/pii/S0038109812002256>.
  - [30] R. Frisenda, M. Drüppel, R. Schmidt, S. Michaelis de Vasconcellos, D. Perez de Lara, R. Bratschitsch, M. Rohlfing, and A. Castellanos-Gomez, *npj 2D Materials and Applications* **1**, 10 (2017), URL <https://doi.org/10.1038/s41699-017-0013-7>.
  - [31] E. Tang and L. Fu, *Nature Physics* **10**, 964 EP (2014), URL <https://doi.org/10.1038/nphys3109>.
  - [32] A. Kerelsky, L. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, et al., arXiv e-prints arXiv:1812.08776 (2018), 1812.08776.
  - [33] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, et al., arXiv e-prints arXiv:1901.02997 (2019), 1901.02997.
  - [34] N. N. T. Nam and M. Koshino, *Phys. Rev. B* **96**, 075311 (2017), URL <https://link.aps.org/doi/10.1103/PhysRevB.96.075311>.
  - [35] G.-B. Liu, W.-Y. Shan, Y. Yao, W. Yao, and D. Xiao, *Phys. Rev. B* **88**, 085433 (2013), URL <https://link.aps.org/doi/10.1103/PhysRevB.88.085433>.
  - [36] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, *Phys. Rev. Lett.* **108**, 196802 (2012), URL <https://link.aps.org/doi/10.1103/PhysRevLett.108.196802>.
  - [37] S. Fang, S. Carr, M. A. Cazalilla, and E. Kaxiras, *Phys. Rev. B* **98**, 075106 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.075106>.
  - [38] R. Bistritzer and A. H. MacDonald, *Proceedings of the National Academy of Sciences* **108**, 12233 (2011), ISSN 0027-8424, <https://www.pnas.org/content/108/30/12233.full.pdf>, URL <https://www.pnas.org/content/108/30/12233>.
  - [39] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, *Phys. Rev. B* **86**, 155449 (2012), URL <https://link.aps.org/doi/10.1103/PhysRevB.86.155449>.
  - [40] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, *Phys. Rev. Lett.* **99**, 256802 (2007), URL <https://link.aps.org/doi/10.1103/PhysRevLett.99.256802>.
  - [41] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009), URL <https://link.aps.org/doi/10.1103/RevModPhys.81.109>.
  - [42] L. Huder, A. Artaud, T. Le Quang, G. T. de Laisardière, A. G. M. Jansen, G. Lapertot, C. Chapelier, and V. T. Renard, *Phys. Rev. Lett.* **120**, 156405 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevLett.120.156405>.
  - [43] J.-B. Qiao, L.-J. Yin, and L. He, *Phys. Rev. B* **98**, 235402 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevB.98.235402>.
  - [44] F. Wu, T. Lovorn, E. Tutuc, I. Martin, and A. H. MacDonald, arXiv e-prints arXiv:1807.03311 (2018), 1807.03311.
  - [45] A. Laturia, M. L. Van de Put, and W. G. Vandenberghe, *npj 2D Materials and Applications* **2**, 6 (2018), URL <https://doi.org/10.1038/s41699-018-0050-x>.
  - [46] K. Kanoda and R. Kato, *Annual Review of Condensed Matter Physics* **2**, 167 (2011), <https://doi.org/10.1146/annurev-conmatphys-062910-140521>, URL <https://doi.org/10.1146/annurev-conmatphys-062910-140521>.
  - [47] B. J. Powell and R. H. McKenzie, *Reports on Progress in Physics* **74**, 056501 (2011), URL <http://stacks.iop.org/0034-4885/74/i=5/a=056501>.
  - [48] Y. Xie, B. Lian, B. Jäck, X. Liu, C.-L. Chiu, K. Watan-

- abe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, arXiv e-prints arXiv:1906.09274 (2019), 1906.09274.
- [49] Y. Jiang, J. Mao, X. Lai, K. Watanabe, T. Taniguchi, K. Haule, and E. Y. Andrei, arXiv e-prints arXiv:1904.10153 (2019), 1904.10153.
- [50] S. Liu, E. Khalaf, J. Y. Lee, and A. Vishwanath, arXiv e-prints arXiv:1905.07409 (2019), 1905.07409.
- [51] Y.-H. Zhang, H. C. Po, and T. Senthil, arXiv e-prints arXiv:1904.10452 (2019), 1904.10452.
- [52] The possibility of the lifted degeneracy by heterostrain is brought to our attention by Matthew Yankowitz and Cory Dean.
- [53] **With our choice of parameters, the first magic angle without strain locates around  $\theta_c \cong 0.95^\circ$ , slightly different from the BM model. However, our results are not sensitive to this detail as shown in Fig. 4**