

Interlayer coupling in commensurate and incommensurate bilayer structures of transition-metal dichalcogenides

Yong Wang,^{1,2} Zhan Wang,¹ Wang Yao,² Gui-Bin Liu,^{3,*} and Hongyi Yu^{2,†}¹*School of Physics, Nankai University, Tianjin 300071, China*²*Department of Physics and Center of Theoretical and Computational Physics, The University of Hong Kong, Hong Kong, China*³*School of Physics, Beijing Institute of Technology, Beijing 100081, China*

(Received 2 November 2016; revised manuscript received 16 February 2017; published 24 March 2017)

The interlayer couplings in commensurate and incommensurate bilayer structures of transition-metal dichalcogenides are investigated with perturbative treatment. The interlayer coupling in $\pm\mathbf{K}$ valleys can be decomposed into a series of hopping terms with distinct phase factors. In H-type and R-type commensurate bilayers, the interference between the three main hopping terms leads to a sensitive dependence of the interlayer coupling strength on the translation that can explain the position dependent local band gap modulation in a heterobilayer moiré superlattice. The interlayer couplings in the Γ valley of valence band and \mathbf{Q} valley of conduction band are also studied, where the strong coupling strengths of several hundred meV can play important roles in mediating the ultrafast interlayer charge transfer in heterobilayers of transition-metal dichalcogenides.

DOI: [10.1103/PhysRevB.95.115429](https://doi.org/10.1103/PhysRevB.95.115429)

I. INTRODUCTION

Monolayer group-VIB transition-metal dichalcogenides (TMDs) have been extensively studied in recent years, mainly due to their exotic physical properties and potential applications in novel two-dimensional (2D) electronics devices [1–5]. Compared with the precedent 2D material graphene, monolayer TMDs have a finite and direct band gap located at the two degenerate but inequivalent hexagonal Brillouin zone (BZ) corners, i.e., the $\pm\mathbf{K}$ valleys, which are essential for the successful operation of transistors and valley-dependent optoelectronics. Furthermore, the strong spin-orbit coupling of the transition metal couples the spin and valley degrees of freedom, making TMDs the ideal platform to develop spintronic and valleytronic devices [1]. Several kinds of electronic and optoelectronic prototype devices have been fabricated with monolayer TMDs, including field-effect transistor, inverter and logic gate, junction and heterostructure, photodetector, solar cell and light-emitting devices, as well as electronic sensors [3,4].

Similar to the monolayers, the natural TMD homobilayers can be obtained from bulk crystals using mechanical exfoliation and have been widely studied. These natural homobilayers mostly exhibit a commensurate 2H (also called AB) stacking where the two layers are 180° rotation of each other [5]. As the two adjacent layers are bound together by the weak van der Waals interaction, the interlayer coupling in $\pm\mathbf{K}$ valleys can be largely suppressed by the giant spin-orbit splitting. The resulted spin-layer locking could lead to various magnetoelectric effects allowing for their quantum manipulations [6–11]. On the other hand, the interlayer couplings in the valence band Γ and conduction band \mathbf{Q} valleys are significantly larger, which strongly shifts their energy positions compared to those of the monolayers and results in a transition from direct to indirect band gap [12,13].

Furthermore, the current technique allows manually stacking two monolayers to form a vertical homo- or heterostructure, with the uncertainty lesser than 1° on their mutual crystallographic alignment [14–16]. This opens up an alternative way to utilize this novel class of 2D materials [17]. For the TMD heterobilayer formed by two different TMD materials, its conduction and valence band edges are located in different layers. Such a type-II band alignment results in the ultrafast interlayer charge transfer which facilitates the photocurrent generation [18–26], and the formation of interlayer excitons [27–29]. Meanwhile, the manually assembled bilayer generally has an incommensurate lattice structure due to the inevitable interlayer twist and/or lattice constant mismatch. This brings anomalous interlayer couplings which have profound effects on the transport [14–16,30,31], optical [32–36], and Raman [37,38] properties of the bilayers. Moreover, recent theoretical studies have shown that the interlayer coupling together with the formation of a large scale moiré superlattice pattern can lead to the emergence of topological orders in a TMD heterobilayer [39,40]. To gain further insights into these interesting phenomena, it is essential to understand the strength and the form of the interlayer coupling in TMD bilayers.

In 2H or other commensurate bilayers, the interlayer coupling can be evaluated by comparing the bilayer band structure to those of the monolayers. The $2\pi/3$ -rotational symmetry of the 2H bilayer is also essential to determine whether the interlayer coupling strength at $\pm\mathbf{K}$ points is zero or not [5]. For the general TMD bilayers, however, it is nontrivial to calculate the interlayer coupling of the incommensurate lattice structures mainly due to the lack of periodic feature. For the limited commensurate cases, the unit cell usually contains too many atoms to be calculated from first principles. Thus some analytical way should be adopted instead of the impractical numerical calculations.

In this paper we investigate the interlayer coupling in general TMD bilayers following the previous studies in twisted bilayer graphene [41–45], by adopting an effective perturbative treatment. The rest of the paper is organized as follows. In Sec. II we show that in general TMD bilayers the interlayer coupling between the $\pm\mathbf{K}$ valley Bloch states can

*gbliu@bit.edu.cn

†yuhongyi@hku.hk

be decomposed into a series of hopping terms with distinct phase factors, which correspond to the Fourier components of the hopping integral between localized atomic orbitals. In Sec. III the symmetry properties of the monolayer TMDs are analyzed and utilized to reveal the relation between the hopping terms. In Sec. IV our perturbative results for the commensurate H- and R-type TMD homobilayers are presented, which show sensitive dependence on the interlayer translation, and are in excellent agreement with the *ab initio* calculations. In Sec. V we apply our perturbative treatment to the lattice-mismatched bilayers, and reveal its connection with the moiré superlattice. In Sec. VI we further study the interlayer coupling of the valence band Γ and conduction band Q valleys, and propose that they play important roles in mediating the ultrafast interlayer charge transfer of TMD heterobilayers. We summarize our results in Sec. VII.

II. EXPRESSION OF INTERLAYER COUPLING IN $\pm K$ VALLEYS

Since the two TMD monolayers are bound by the weak van der Waals force, we can first consider a decoupled bilayer, then add the interlayer coupling as a perturbation. In the vanishing interlayer coupling limit, the monolayer Bloch wave functions in τK valley are denoted as $\psi_{n,\mathbf{k}}(\mathbf{r}) \equiv \langle \mathbf{r} | n, \mathbf{k} \rangle = e^{i(\tau \mathbf{K} + \mathbf{k}) \cdot \mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r})$. Here $n = \{\tau, l\}$ contains both the valley index $\tau = \pm$ and the band index $l = \dots, c+1, c, v, v-1, \dots$. Here c (v) corresponds to the conduction (valence) band, and we use $c+j$ ($v-j$) to denote the j th band above (below) the conduction (valence) band. $u_{n,\mathbf{k}}(\mathbf{r})$ is the periodic part of the Bloch wave functions.

The Bloch wave function $\psi_{n,0}$ can be constructed from the local basis functions as

$$\psi_{n,0}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\tau \mathbf{K} \cdot \mathbf{R}} D_n(\mathbf{r} - \mathbf{R}). \quad (1)$$

Here N is the unit cell number of the corresponding monolayer, and $D_n(\mathbf{r} - \mathbf{R})$ is the linear combination of the atomic orbitals localized near the metal position \mathbf{R} , which depends on the valley index τ and band index l (see Table I). Considering the time reversal relation between the two valleys, D_n in the same band but opposite valleys are related by a complex conjugate. Under the envelope approximation $\psi_{n,\mathbf{k}}(\mathbf{r}) \approx e^{i(\tau \mathbf{K} + \mathbf{k}) \cdot \mathbf{r}} u_{n,0}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} \psi_{n,0}(\mathbf{r})$, one finds

$$\psi_{n,\mathbf{k}}(\mathbf{r}) \approx \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i(\tau \mathbf{K} + \mathbf{k}) \cdot \mathbf{R}} D_n(\mathbf{r} - \mathbf{R}), \quad (2)$$

where $e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{R})} D_n(\mathbf{r} - \mathbf{R}) \approx D_n(\mathbf{r} - \mathbf{R})$ is used since we are interested in low energy electrons and holes with small $|\mathbf{k}|$, and $D_n(\mathbf{r} - \mathbf{R})$ is well localized near \mathbf{R} .

We define a bilayer stacking configuration as the reference one, where the in-plane crystalline axes of the two layers are along the same direction (R-type stacking), and the two metal atoms in different layers horizontally overlap at the in-plane (xy) coordinate origin. Any other stacking configuration can then be obtained from this reference configuration through a θ -angle rotation of the upper layer around the coordinate origin, and followed by a translation of $-\mathbf{r}_0$ for the lower layer [see Fig. 1(a)]. We use the convention that quantities in the upper

TABLE I. The orbital compositions and the corresponding $2\pi/3$ -rotational quantum numbers M (discussed in Sec. III A) of localized function D_n in $+K$ valley of monolayer MoS₂, obtained by following Ref. [5]. The percentage is defined as the overlap probability between the atomic orbital wave function and the K -point Bloch state. $d_0 \equiv d_{z^2}$, $d_{\pm 1} \equiv (d_{xz} \pm i d_{yz})/\sqrt{2}$, $d_{\pm 2} \equiv (d_{x^2-y^2} \pm i d_{xy})/\sqrt{2}$ are the Mo- d orbitals, and $p_0 \equiv p_z$, $p_{\pm 1} \equiv (p_x \pm i p_y)/\sqrt{2}$ are the S- p orbitals. Mo- p_0 denotes the p_z orbital of the Mo atom. Only the two most prominent orbitals are shown.

(Band)	(Major orbital)	(Minor orbital)	(M)
\vdots	\vdots	\vdots	\vdots
$c+3$	d_{+1} (70%)	p_0 (24%)	+1
$c+2$	d_{-2} (78%)	p_0 (19%)	+1
$c+1$	d_{-1} (78%)	p_{+1} (22%)	-1
c	d_0 (88%)	p_{-1} (7%)	0
v	d_{+2} (84%)	p_{+1} (16%)	-1
$v-1$	p_0 (56%)	d_{+1} (38%)	+1
$v-2$	p_{-1} (83%)	Mo- p_0 (17%)	0
$v-3$	p_0 (53%)	d_{-2} (31%)	+1
\vdots	\vdots	\vdots	\vdots

(lower) layer are marked with (without) the prime. The lower layer band edges are located at $\pm \mathbf{K} = \pm \frac{4\pi}{3a}(1,0)$, while those of the upper layer are located at $\pm \mathbf{K}' = \pm \frac{4\pi}{3a'}(\cos \theta, \sin \theta)$, where a (a') is the lower (upper) layer lattice constant.

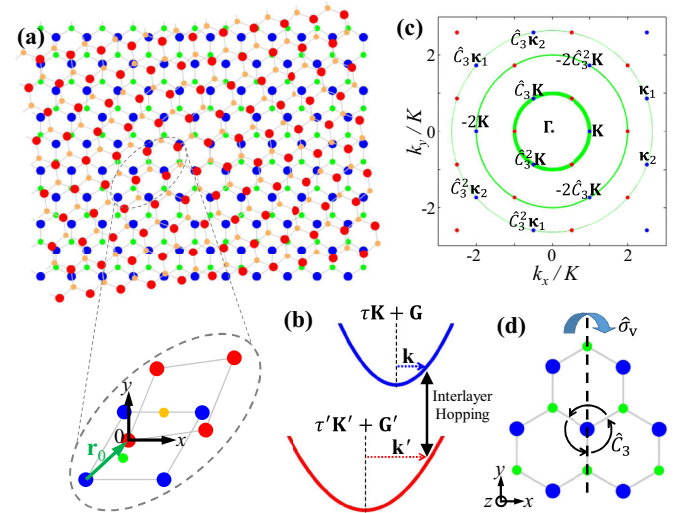


FIG. 1. (a) Illustration of a twisted TMD homo- or heterobilayer. The large red (blue) dots denote the metal atoms in the upper (lower) layer, and the small orange (green) dots denote the chalcogen atoms in the upper (lower) layer. The enlarged view shows two unit cells in the upper and lower layers, respectively. The in-plane (xy) coordinate origin is set on a metal atom in the upper layer. (b) Two wave vectors in different layers must overlap in momentum space to satisfy the momentum conservation of interlayer hopping. (c) The blue dots denote $\kappa \equiv \mathbf{K} + \mathbf{G}$ points, and the red dots are their time reversals. Thicker green circle means smaller $|\kappa|$ thus larger $|t_{nn}(\tau \kappa)|$ [see Eq. (5)]. (d) Illustration of the $2\pi/3$ -rotational (\hat{C}_3) symmetry and the in-plane mirror ($\hat{\sigma}_v$) symmetry of monolayer TMDs.

Now we add the interlayer coupling \hat{H}_t as a perturbation. We consider the hopping integral between the two wave functions $\psi_{n',\mathbf{k}'}$ and $\psi_{n,\mathbf{k}}$ located in the upper and lower layer, respectively, which can be expressed as

$$\begin{aligned} \langle n, \mathbf{k} | \hat{H}_t | n', \mathbf{k}' \rangle &\equiv \int \psi_{n,\mathbf{k}}^*(\mathbf{r}) \hat{H}_t \psi_{n',\mathbf{k}'}(\mathbf{r}) d\mathbf{r} \\ &= \sum_{\mathbf{R}, \mathbf{R}'} \frac{e^{i(\tau'\mathbf{K}'+\mathbf{k}')\cdot\mathbf{R}'-i(\tau\mathbf{K}+\mathbf{k})\cdot\mathbf{R}}}{\sqrt{NN'}} \langle D_{n,\mathbf{R}} | \hat{H}_t | D_{n',\mathbf{R}'} \rangle. \end{aligned} \quad (3)$$

Here $\langle D_{n,\mathbf{R}} | \hat{H}_t | D_{n',\mathbf{R}'} \rangle \equiv \int D_n^*(\mathbf{r}-\mathbf{R}) \hat{H}_t D_{n'}(\mathbf{r}-\mathbf{R}') d\mathbf{r}$ is the hopping integral between the two localized orbitals $D_{n'}(\mathbf{r}-\mathbf{R}')$ and $D_n(\mathbf{r}-\mathbf{R})$. In the spirit of two-center approximation [41–43], $\langle D_{n,\mathbf{R}} | \hat{H}_t | D_{n',\mathbf{R}'} \rangle$ depends only on the relative position $\mathbf{R}'-\mathbf{R}$. So we can write

$$\begin{aligned} \langle D_{n,\mathbf{R}} | \hat{H}_t | D_{n',\mathbf{R}'} \rangle &= T_{nn'}(\mathbf{R}'-\mathbf{R}) \\ &= \sum_{\mathbf{q}} \frac{e^{-i\mathbf{q}\cdot(\mathbf{R}'-\mathbf{R})}}{\sqrt{NN'}} t_{nn'}(\mathbf{q}). \end{aligned} \quad (4)$$

Here $t_{nn'}(\mathbf{q}) = \frac{1}{\sqrt{\Omega\Omega'}} \int T_{nn'}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$ is the Fourier transform of $T_{nn'}(\mathbf{r})$, with Ω' (Ω) the upper (lower) layer unit cell area.

We denote the in-plane positions of the metal atoms in the upper (lower) layer as $\mathbf{R}' = j'_1\mathbf{a}'_1 + j'_2\mathbf{a}'_2$ ($\mathbf{R} = -\mathbf{r}_0 + j_1\mathbf{a}_1 + j_2\mathbf{a}_2$), where $\mathbf{a}'_{1,2}$ ($\mathbf{a}_{1,2}$) are the corresponding unit lattice vectors and $j'_{1,2}$, $j_{1,2}$ are integers. Substituting Eq. (4) into Eq. (3), we obtain

$$\begin{aligned} \langle n, \mathbf{k} | \hat{H}_t | n', \mathbf{k}' \rangle &= \sum_{\mathbf{q}} t_{nn'}(\mathbf{q}) \sum_{\mathbf{R}, \mathbf{R}'} \frac{e^{i(\tau'\mathbf{K}'+\mathbf{k}')\cdot\mathbf{R}'-i(\tau\mathbf{K}+\mathbf{k})\cdot\mathbf{R}}}{NN'} \\ &= \sum_{\mathbf{q}} t_{nn'}(\mathbf{q}) \sum_{\mathbf{G}, \mathbf{G}'} \delta_{\tau\mathbf{K}+\mathbf{k}-\mathbf{q}, \mathbf{G}} \delta_{\tau'\mathbf{K}'+\mathbf{k}'-\mathbf{q}, \mathbf{G}'} e^{i\mathbf{G}\cdot\mathbf{r}_0} \\ &= \sum_{\mathbf{G}, \mathbf{G}'} \delta_{\tau\mathbf{K}+\mathbf{k}+\mathbf{G}, \tau'\mathbf{K}'+\mathbf{k}'+\mathbf{G}'} t_{nn'}(\tau\mathbf{K}+\mathbf{k}+\mathbf{G}) e^{-i\mathbf{G}\cdot\mathbf{r}_0}. \end{aligned}$$

Here \mathbf{G}' (\mathbf{G}) is the reciprocal lattice vector of the upper (lower) layer. Note that the translation vector \mathbf{r}_0 appear in the phase factor $e^{-i\mathbf{G}\cdot\mathbf{r}_0}$ only. Since the phase factor does not change when we replace \mathbf{r}_0 by $\mathbf{r}_0 + j_1\mathbf{a}_1 + j_2\mathbf{a}_2$, we can restrict \mathbf{r}_0 to be inside a unit cell of the lower layer.

To simplify the above expression, we use the notation $\tau'\mathbf{k}' \equiv \tau'\mathbf{K}' + \mathbf{G}'$ and $\tau\mathbf{k} \equiv \tau\mathbf{K} + \mathbf{G}$, and write

$$\langle n, \mathbf{k} | \hat{H}_t | n', \mathbf{k}' \rangle = e^{i\tau\mathbf{K}\cdot\mathbf{r}_0} \sum_{\kappa'\kappa} \delta_{\mathbf{k}'-\mathbf{k}, \tau\kappa-\tau'\kappa'} t_{nn'}(\tau\mathbf{K}+\mathbf{k}) e^{-i\tau\kappa\cdot\mathbf{r}_0}. \quad (5)$$

Equation (5) is the central result of this paper, closely similar forms also appear in other works for graphene-related van der Waals materials [41–45] and our early paper for heterobilayer TMDs [39]. It implies that the hopping integral between two Bloch functions in different layers is nonzero only when $\mathbf{k}'-\mathbf{k}$ equals one of the discrete values $\tau\kappa-\tau'\kappa'$, as illustrated in Fig. 1(b). Furthermore we expect $t_{nn'}(\mathbf{q})$ to decay fast with the increase of $|\mathbf{q}|$, as $D_n(\mathbf{r})$ and $D_{n'}(\mathbf{r})$ vary smoothly with \mathbf{r} and the integral $\langle D_{n,\mathbf{R}} | \hat{H}_t | D_{n',\mathbf{R}'} \rangle$ is generally a smooth function of

$\mathbf{R}'-\mathbf{R}$. Therefore, in the summation $\sum_{\kappa'\kappa}$ only a few terms of κ' and κ with small magnitudes need to be kept, which greatly reduces the number of $\tau\kappa-\tau'\kappa'$. In Fig. 1(c) we show three groups of κ . \mathbf{K} , $\hat{C}_3\mathbf{K}$, and $\hat{C}_3^2\mathbf{K}$ on the thickest circle are closest to Γ and are expected to have the most pronounced $|t_{nn'}|$; $-2\mathbf{K}$, $-2\hat{C}_3\mathbf{K}$, and $-2\hat{C}_3^2\mathbf{K}$ ($\kappa_{1,2}$, $\hat{C}_3\kappa_{1,2}$, and $\hat{C}_3^2\kappa_{1,2}$) are the second (third) closest to Γ , and the corresponding $|t_{nn'}|$ values are expected to be much weaker.

III. SYMMETRY PROPERTIES OF THE HOPPING TERMS

The monolayer hexagonal lattice structure has both the $2\pi/3$ -rotational (\hat{C}_3) symmetry and the in-plane mirror ($\hat{\sigma}_v$) symmetry [see Fig. 1(d)]. The hopping terms $t_{nn'}(\mathbf{q})$ with the same $|\mathbf{q}|$ values but different \mathbf{q} directions are related by these symmetry operations.

A. $2\pi/3$ -rotational symmetry

We use \hat{C}_3 to denote the in-plane counterclockwise $\frac{2\pi}{3}$ rotation around $\mathbf{r}=0$ when applied on a real space vector (around Γ point when applied on a \mathbf{k} -space vector). As the $\tau\mathbf{K}$ point has a high symmetry, i.e., $\hat{C}_3\tau\mathbf{K} = \tau\mathbf{K} + \mathbf{G}$, the orbital combination $D_n(\mathbf{r})$ should be \hat{C}_3 symmetric: $D_n(\hat{C}_3\mathbf{r}) = e^{i\frac{2\pi}{3}M(n)} D_n(\mathbf{r})$, where the \hat{C}_3 quantum number $M(n) = \tau M(l)$ has opposite value in two valleys because of the time reversal relation. $M(l) = \{0, \pm 1\}$ as a function of the band index l is summarized in Table I. Then the hopping integral satisfies

$$\begin{aligned} T_{nn'}(\hat{C}_3\mathbf{R}'-\hat{C}_3\mathbf{R}) &= \int D_n^*(\mathbf{r}-\hat{C}_3\mathbf{R}) \hat{H}_t D_{n'}(\mathbf{r}-\hat{C}_3\mathbf{R}') d\mathbf{r} \\ &= \int D_n^*(\hat{C}_3\mathbf{r}-\hat{C}_3\mathbf{R}) \hat{H}_t D_{n'}(\hat{C}_3\mathbf{r}-\hat{C}_3\mathbf{R}') d\mathbf{r} \\ &= e^{i\frac{2\pi}{3}[M(n')-M(n)]} T_{nn'}(\mathbf{R}'-\mathbf{R}). \end{aligned}$$

With the equation above, applying Fourier transformation to $T_{nn'}(\mathbf{r})$ results in

$$\begin{aligned} t_{nn'}(\hat{C}_3\mathbf{q}) &= \frac{1}{\sqrt{\Omega\Omega'}} \int T_{nn'}(\mathbf{r}) e^{i\hat{C}_3\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \\ &= \frac{1}{\sqrt{\Omega\Omega'}} \int T_{nn'}(\hat{C}_3\mathbf{r}) e^{i\hat{C}_3\mathbf{q}\cdot\hat{C}_3\mathbf{r}} d\mathbf{r} \\ &= e^{i\frac{2\pi}{3}[M(n')-M(n)]} t_{nn'}(\mathbf{q}). \end{aligned} \quad (6)$$

In the last step in Eq. (6), we have used the relation $\hat{C}_3\mathbf{q}\cdot\hat{C}_3\mathbf{r} = \mathbf{q}\cdot\mathbf{r}$.

B. In-plane mirror symmetry

We use $\hat{\sigma}_v$ to denote the mirror reflection operation on a real space vector $\mathbf{r} = (r_x, r_y)$ over the vertical yz plane, i.e., $\hat{\sigma}_v\mathbf{r} = (-r_x, r_y)$, or on a wave vector $\mathbf{q} = (q_x, q_y)$ as $\hat{\sigma}_v\mathbf{q} = (-q_x, q_y)$. Obviously $\hat{\sigma}_v\mathbf{K} = -\mathbf{K}$, thus under the mirror reflection $\psi_{\tau,0,n}(\hat{\sigma}_v\mathbf{r}) = \psi_{-\tau,0,n}(\mathbf{r}) = \psi_{\tau,0,n}^*(\mathbf{r})$, where the last step comes from the time reversal relation between the two valleys. Together with Eq. (1), the local wave function $D_n(\mathbf{r}-\mathbf{R})$ satisfies the property $D_n(\hat{\sigma}_v\mathbf{r}-\hat{\sigma}_v\mathbf{R}) = D_n^*(\mathbf{r}-\mathbf{R})$. When both the upper and lower layer have the yz -plane mirror

symmetry, i.e., R stacking ($\theta = 0^\circ$) or H stacking ($\theta = 60^\circ$), one gets

$$\begin{aligned} & T_{nn'}(\hat{\sigma}_v \mathbf{R}' - \hat{\sigma}_v \mathbf{R}) \\ &= \int D_n^*(\mathbf{r} - \hat{\sigma}_v \mathbf{R}) \hat{H}_t D_{n'}(\mathbf{r} - \hat{\sigma}_v \mathbf{R}') d\mathbf{r} \\ &= \int D_n^*(\hat{\sigma}_v \mathbf{r} - \hat{\sigma}_v \mathbf{R}) \hat{H}_t D_{n'}(\hat{\sigma}_v \mathbf{r} - \hat{\sigma}_v \mathbf{R}') d\mathbf{r} \\ &= T_{nn'}^*(\mathbf{R}' - \mathbf{R}). \end{aligned}$$

A Fourier transformation of $T_{nn'}(\mathbf{r})$ results in

$$\begin{aligned} t_{nn'}(\hat{\sigma}_v \mathbf{q}) &= \frac{1}{\sqrt{\Omega \Omega'}} \int T_{nn'}(\mathbf{r}) e^{i\hat{\sigma}_v \mathbf{q} \cdot \mathbf{r}} d\mathbf{r} \\ &= \frac{1}{\sqrt{\Omega \Omega'}} \int T_{nn'}(\hat{\sigma}_v \mathbf{r}) e^{i\hat{\sigma}_v \mathbf{q} \cdot \hat{\sigma}_v \mathbf{r}} d\mathbf{r} \\ &= t_{nn'}^*(-\mathbf{q}). \end{aligned} \quad (7)$$

In the last step in Eq. (7), we have used the relation $\hat{\sigma}_v \mathbf{q} \cdot \hat{\sigma}_v \mathbf{r} = \mathbf{q} \cdot \mathbf{r}$. Therefore, $t_{nn'}(\mathbf{q})$ is real when $q_y = 0$ in an R-type or H-type bilayer.

IV. $\pm \mathbf{K}$ -VALLEY COUPLING STRENGTH IN H- AND R-TYPE HOMOBILAYERS

In homobilayer TMDs, the conduction and valence bands of the structures will be twofold degenerate at $\tau \mathbf{K}$ point (without considering the spin-orbit coupling) if there is no interlayer coupling, i.e., $E_{c+j} = E_{c'+j}$, $E_{v-j} = E_{v'-j}$, with $j = 0, 1, 2, \dots$. The presence of the interlayer coupling will cause a finite energy level splitting $\Delta E_{c(v)}$, which contains the information of the hopping terms $t_{nn'}(\mathbf{q})$.

We consider R-type ($\theta = 0^\circ$) or H-type ($\theta = 60^\circ$) TMD homobilayer structures with varying \mathbf{r}_0 . As the two layers are fully commensurate, the interlayer hopping between $\tau \mathbf{K}$ in the lower layer and $\tau' \mathbf{K}'$ in the upper layer is allowed when $\tau' = \tau$ for R stacking, and $\tau' = -\tau$ for H stacking. To simplify the notation, we write $|n'\rangle \equiv |n', 0\rangle$ and $|n\rangle \equiv |n, 0\rangle$. Using Eqs. (6) and (7), the hopping integral of Eq. (5) between $\tau \mathbf{K}$ and $\tau' \mathbf{K}'$ can be written as

$$\begin{aligned} e^{-i\tau \mathbf{K} \cdot \mathbf{r}_0} \langle n | \hat{H}_t | n' \rangle &= \sum_{\kappa} t_{nn'}(\tau \kappa) e^{-i\tau \kappa \cdot \mathbf{r}_0} \\ &\approx (e^{-i\tau \mathbf{K} \cdot \mathbf{r}_0} + e^{-i\tau \hat{C}_3 \mathbf{K} \cdot \mathbf{r}_0} e^{i\frac{2\pi}{3}[M(n')-M(n)]} \\ &\quad + e^{-i\tau \hat{C}_3^2 \mathbf{K} \cdot \mathbf{r}_0} e^{i\frac{4\pi}{3}[M(n')-M(n)]}) t_{nn'}^{(0)} \\ &\quad + (e^{2i\tau \mathbf{K} \cdot \mathbf{r}_0} + e^{2i\tau \hat{C}_3 \mathbf{K} \cdot \mathbf{r}_0} e^{i\frac{2\pi}{3}[M(n')-M(n)]} \\ &\quad + e^{2i\tau \hat{C}_3^2 \mathbf{K} \cdot \mathbf{r}_0} e^{i\frac{4\pi}{3}[M(n')-M(n)]}) t_{nn'}^{(1)} \\ &\quad + (e^{-i\tau \kappa_1 \cdot \mathbf{r}_0} + e^{-i\tau \hat{C}_3 \kappa_1 \cdot \mathbf{r}_0} e^{i\frac{2\pi}{3}[M(n')-M(n)]} \\ &\quad + e^{-i\tau \hat{C}_3^2 \kappa_1 \cdot \mathbf{r}_0} e^{i\frac{4\pi}{3}[M(n')-M(n)]}) t_{nn'}^{(2)} \\ &\quad + (e^{-i\tau \kappa_2 \cdot \mathbf{r}_0} + e^{-i\tau \hat{C}_3 \kappa_2 \cdot \mathbf{r}_0} e^{i\frac{2\pi}{3}[M(n')-M(n)]} \\ &\quad + e^{-i\tau \hat{C}_3^2 \kappa_2 \cdot \mathbf{r}_0} e^{i\frac{4\pi}{3}[M(n')-M(n)]}) (t_{nn'}^{(2)})^*. \end{aligned} \quad (8)$$

Here $t_{nn'}^{(0)} \equiv t_{nn'}(\tau \mathbf{K})$ corresponds to the main hopping term, $t_{nn'}^{(1)} \equiv t_{nn'}(-2\tau \mathbf{K})$ is the first order term, and $t_{nn'}^{(2)} \equiv t_{nn'}(\tau \kappa_1) =$

$t_{nn'}^*(\tau \kappa_2)$ is the second order term. Note that $t_{nn'}^{(0)}$ and $t_{nn'}^{(1)}$ are real due to Eq. (7), while $t_{nn'}^{(2)}$ is complex. We have dropped the other higher order terms with larger $|\kappa|$ whose contributions are expected to be negligible.

Now we analyze the conduction band splitting ΔE_c . Because of the large splitting between two different bands, the hopping between lower layer c band and upper layer n' band with $n' \neq c'$ can be well accounted by a second-order perturbation, which results in an energy shift $\delta E_c(\mathbf{r}_0) \equiv \sum_{n' \neq c'} \frac{|(c|\hat{H}_t|n')|^2}{E_c - E_{n'}}$ to the c band. Similarly, the lower layer n band with $n \neq c$ results in an energy shift $\delta E_{c'}(\mathbf{r}_0) \equiv \sum_{n \neq c} \frac{|(c'|\hat{H}_t|n)|^2}{E_{c'} - E_n}$ to the c' band. So in the subspace spanned by $|c'\rangle$ and $|c\rangle$, the hopping Hamiltonian has a form

$$\begin{aligned} \hat{H}_{cc'} &= [E_c + \delta E_c(\mathbf{r}_0)]|c\rangle\langle c| + [E_c + \delta E_{c'}(\mathbf{r}_0)]|c'\rangle\langle c'| \\ &\quad + \langle c' | \hat{H}_t | c \rangle |c'\rangle\langle c| + \text{H.c.} \end{aligned} \quad (9)$$

ΔE_c is then given by the energy splitting between the eigenstates of $\hat{H}_{cc'}$, which is

$$\Delta E_c = \sqrt{[\delta E_c(\mathbf{r}_0) - \delta E_{c'}(\mathbf{r}_0)]^2 + 4|\langle c' | \hat{H}_t | c \rangle|^2}. \quad (10)$$

The same analysis can be applied to the valence bands, which gives

$$\Delta E_v = \sqrt{[\delta E_v(\mathbf{r}_0) - \delta E_{v'}(\mathbf{r}_0)]^2 + 4|\langle v' | \hat{H}_t | v \rangle|^2}, \quad (11)$$

with $\delta E_v(\mathbf{r}_0) \equiv \sum_{n' \neq v'} \frac{|(v|\hat{H}_t|n')|^2}{E_v - E_{n'}}$ and $\delta E_{v'}(\mathbf{r}_0) \equiv \sum_{n \neq v} \frac{|(v'|\hat{H}_t|n)|^2}{E_{v'} - E_n}$.

For most of the \mathbf{r}_0 values, the corresponding R- or H-type commensurate bilayer structures are unstable and thus do not exist in nature. However, these structures can locally exist in an incommensurate bilayer with a large scale moiré superlattice pattern [39,40,46]. In a local region with a size much larger than the monolayer lattice constant but much smaller than the moiré supercell, the atomic registry between the two layers is locally indistinguishable from an R- or H-type commensurate bilayer, which is characterized by a continuously varying \mathbf{r}_0 . The local band structure of this region is then given by that of the commensurate bilayer with the corresponding \mathbf{r}_0 value [39,40,46]. As \mathbf{r}_0 varies from position to position in a moiré supercell, the \mathbf{r}_0 -dependent conduction/valence band energy shifts $\delta E_{c/v}(\mathbf{r}_0)$ can be responsible for the observed position-dependent local band gap modulation [46].

A. H-type homobilayer

For H-type stacking, the two states with finite hopping strength in different layers have the opposite valley indices $\tau = -\tau'$. Using Eq. (8) together with the $M(n)$ values given in Table I, we find $\delta E_c(\mathbf{r}_0) = \delta E_{c'}(\mathbf{r}_0)$ and $\delta E_v(\mathbf{r}_0) = \delta E_{v'}(\mathbf{r}_0)$. This can be understood from the symmetry consideration. As shown in Fig. 2(a), an H-type homobilayer with an arbitrary \mathbf{r}_0 has a spatial inversion center, which means the two layers are symmetric. So $\delta E_{c/v}(\mathbf{r}_0)$, the lower layer conduction/valence band energy shift induced by the remote bands in the upper layer, is always equivalent to $\delta E_{c'/v'}(\mathbf{r}_0)$ which is the upper layer energy shift induced by the lower layer. The band

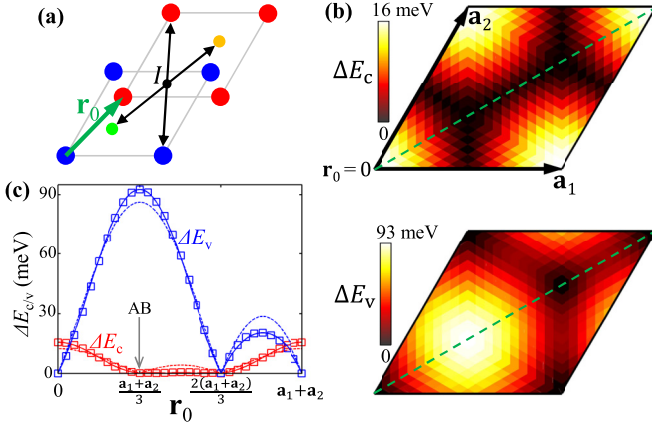


FIG. 2. (a) An H-type TMD homobilayer with any interlayer translation \mathbf{r}_0 always has an inversion center I . Here the large red (blue) dots denote the metal atoms in the upper (lower) layer, and the small orange (green) dots denote the chalcogen atoms in the upper (lower) layer. (b) The *ab initio* results for $\Delta E_{c/v}$ as functions of \mathbf{r}_0 . (c) 2D plots showing the $\Delta E_{c/v}$ line cuts along dashed green lines in (b), where the symbols are the *ab initio* results and the solid curves are the fits using $t_{cc'}^{(0)}$, $t_{cc'}^{(1)}$, and $t_{cc'}^{(2)}$. The dashed curves are the results keeping only the main terms $t_{cc'}^{(0)}$. The natural TMD homobilayers with AB stacking correspond to $\mathbf{r}_0 = (\mathbf{a}_1 + \mathbf{a}_2)/3$.

splittings are then simply given by

$$\Delta E_c = 2|\langle c' | \hat{H}_t | c \rangle|, \quad \Delta E_v = 2|\langle v' | \hat{H}_t | v \rangle|. \quad (12)$$

We have performed *ab initio* calculations for the band structures of MoS₂ H-type homobilayers with different \mathbf{r}_0 . For each given \mathbf{r}_0 , we fix the interlayer distance defined as the vertical distance between the nearest chalcogen atoms of neighboring layers at $d = 2.975 \text{ \AA}$ (the experimental bulk value [47]), and the other lattice parameters are taken from Ref. [48]. The energy splitting values $\Delta E_{c/v}$ are calculated with the projector-augmented wave (PAW) method implemented in the QUANTUM ESPRESSO package [49]. The Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional and scalar relativistic pseudopotential without including the spin-orbit coupling has been exploited, and the cutoff energy for plane wave basis is set as 80 Ry. A $15 \times 15 \times 1$ \mathbf{k} -point sample is generated by the Monkhorst-Pack (MP) approach, and the self-consistent ground state is achieved with the total energy converge criteria 10^{-10} Ry.

The calculation results are presented in Fig. 2(b) as surface plots. In the 2D plot of Fig. 2(c) with \mathbf{r}_0 along the long diagonal line of the unit cell, we show both the *ab initio* results (symbols) and the corresponding fits (solid lines) using Eqs. (12) and (8) by keeping the $t_{cc'/vv'}^{(0)}$, $t_{cc'/vv'}^{(1)}$, and $t_{cc'/vv'}^{(2)}$ hopping terms. The two show excellent agreement. The dashed lines are the results keeping only the main hopping terms $t_{cc'}^{(0)}$, which can already reproduce the major features. Thus those $t_{nn'}(\tau\kappa)$ with larger $|\tau\kappa|$ are indeed negligible. The fitting parameters are summarized in Table II, which give $|t_{cc'}^{(0)}| \gg |t_{cc'}^{(1)}|, |t_{cc'}^{(2)}|$ and $|t_{vv'}^{(0)}| \gg |t_{vv'}^{(1)}|, |t_{vv'}^{(2)}|$ as we expected. So in H-type commensurate bilayers, it is a good approximation to write the \mathbf{K} -point conduction/valence band interlayer couplings in the

TABLE II. The obtained hopping strengths for the H-type homobilayer MoS₂ from fitting to the *ab initio* results of band splitting. The main hopping term $t_{vv'}^{(0)}$ is consistent with our previous result [39].

$t_{cc'}^{(0)}$	$t_{cc'}^{(1)}$	$ t_{cc'}^{(2)} $	$t_{vv'}^{(0)}$	$t_{vv'}^{(1)}$	$ t_{vv'}^{(2)} $
2.1 meV	0.4 meV	0.1 meV	14.4 meV	1.2 meV	0.4 meV

forms

$$\begin{aligned} |\langle c | \hat{H}_t | c' \rangle_H| &\approx |e^{i\mathbf{K}\cdot\mathbf{r}_0} + e^{i\hat{C}_3\mathbf{K}\cdot\mathbf{r}_0} + e^{i\hat{C}_3^2\mathbf{K}\cdot\mathbf{r}_0}| t_{cc'}^{(0)}, \\ |\langle v | \hat{H}_t | v' \rangle_H| &\approx |e^{i\mathbf{K}\cdot\mathbf{r}_0} + e^{i(\hat{C}_3\mathbf{K}\cdot\mathbf{r}_0 + \frac{2\pi}{3})} + e^{i(\hat{C}_3^2\mathbf{K}\cdot\mathbf{r}_0 + \frac{4\pi}{3})}| t_{vv'}^{(0)}. \end{aligned} \quad (13)$$

The above equations should also apply to H-type commensurate heterobilayers. Similar forms have been obtained in early papers [39,42,45,50]. Here we would like to point out that in these interlayer coupling forms the $e^{\pm i2\pi/3}$ phase factors have different origins for bilayer TMD and graphene systems. In bilayer TMDs it is from the \hat{C}_3 quantum number $M(n)$ of the atomic orbital combination D_n , as clearly indicated by Eqs. (6) and (8). While in bilayer graphene, it originates from the displacement vectors between the nearest A and B sublattice sites.

From the above equations we get $t_{cc'}^{(0)} \approx \Delta E_c/6$ at $\mathbf{r}_0 = 0$ and $t_{vv'}^{(0)} \approx \Delta E_v/6$ at $\mathbf{r}_0 = (\mathbf{a}_1 + \mathbf{a}_2)/3$ for H-type homobilayers.

B. R-type homobilayer

In contrast to the H stacking, the R-type homobilayer is not inversion symmetric, thus generally the upper and lower layers are not equivalent. We find $\delta E_c(\mathbf{r}_0) \neq \delta E_{c'}(\mathbf{r}_0)$ and $\delta E_v(\mathbf{r}_0) \neq \delta E_{v'}(\mathbf{r}_0)$ for a general R-type stacking. Only for AA staking with $\mathbf{r}_0 = 0$, which has the out-of-plane mirror reflection ($\hat{\sigma}_h$) symmetry [Fig. 3(a)], the two layers become equivalent and $\delta E_{c/v}(\mathbf{r}_0 = 0) = \delta E_{c'/v'}(\mathbf{r}_0 = 0)$.

For R-type stacking, the two states with finite hopping strength in different layers have the same valley indices $\tau = \tau'$. Using Eq. (8) together with the $M(n)$ values given in Table I, we can write the \mathbf{r}_0 dependence of $\delta E_{c/v}(\mathbf{r}_0)$ and $\delta E_{c'/v'}(\mathbf{r}_0)$ as

$$\begin{aligned} \delta E_c(\mathbf{r}_0) - \delta E_{c'}(\mathbf{r}_0) &\approx \delta E_c^{(0)} f(\mathbf{r}_0), \\ \delta E_v(\mathbf{r}_0) - \delta E_{v'}(\mathbf{r}_0) &\approx \delta E_v^{(0)} f(\mathbf{r}_0), \end{aligned} \quad (14)$$

where $\delta E_{c/v}^{(0)}$ are from the main hopping terms:

$$\begin{aligned} \delta E_c^{(0)} &\equiv \frac{|t_{c,v'}^{(0)}|^2}{E_c - E_v} - \frac{|t_{c,c'+1}^{(0)}|^2}{E_{c+1} - E_c} - \frac{|t_{c,v'-1}^{(0)}|^2}{E_c - E_{v-1}} \\ &\quad - \frac{|t_{c,v'-3}^{(0)}|^2}{E_c - E_{v-3}} + \frac{|t_{c,c'+2}^{(0)}|^2}{E_{c+2} - E_c} + \frac{|t_{c,c'+3}^{(0)}|^2}{E_{c+3} - E_c} \dots, \\ \delta E_v^{(0)} &\equiv \frac{|t_{v,v'-1}^{(0)}|^2}{E_v - E_{v-1}} + \frac{|t_{v,v'-3}^{(0)}|^2}{E_v - E_{v-3}} - \frac{|t_{v,c'+2}^{(0)}|^2}{E_{c+2} - E_v} \\ &\quad - \frac{|t_{v,c'+3}^{(0)}|^2}{E_{c+3} - E_v} + \frac{|t_{v,c'}^{(0)}|^2}{E_c - E_v} - \frac{|t_{v,v'-2}^{(0)}|^2}{E_v - E_{v-2}} \dots, \end{aligned} \quad (15)$$

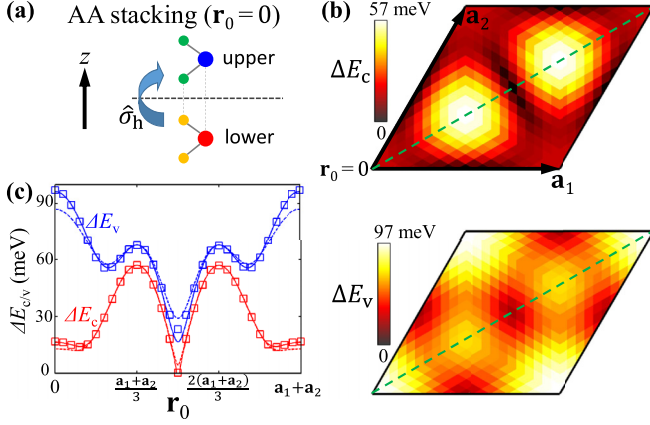


FIG. 3. (a) An AA-type TMD homobilayer has an out-of-plane mirror reflection ($\hat{\sigma}_h$) symmetry. (b) The *ab initio* results for $\Delta E_{c/v}$ as functions of \mathbf{r}_0 for R-type homobilayer MoS₂. (c) 2D plots showing the $\Delta E_{c/v}$ line cuts along dashed green lines in (b), where the symbols are the *ab initio* results and the solid curves are the fits using $\delta E_{c/v}^{(0)}$, $t_{cc'/vv'}^{(0)}$, and $t_{cc'/vv'}^{(1)}$. The dashed curves are the results keeping only the main hopping terms $\delta E_{c/v}^{(0)}$ and $t_{cc'/vv'}^{(0)}$.

and other higher order terms are ignored. So for R-type homobilayer TMDs, the conduction/valence band splitting has a form

$$\begin{aligned} \Delta E_c &= \sqrt{(\delta E_c^{(0)} f(\mathbf{r}_0))^2 + 4|c'|\hat{H}_t|c|^2}, \\ \Delta E_v &= \sqrt{(\delta E_v^{(0)} f(\mathbf{r}_0))^2 + 4|v'|\hat{H}_t|v|^2}, \end{aligned} \quad (16)$$

where

$$\begin{aligned} f(\mathbf{r}_0) &\equiv |e^{i\mathbf{K}\cdot\mathbf{r}_0} + e^{i(\hat{C}_3\mathbf{K}\cdot\mathbf{r}_0 + \frac{2\pi}{3})} + e^{i(\hat{C}_3^2\mathbf{K}\cdot\mathbf{r}_0 + \frac{4\pi}{3})}|^2 \\ &\quad - |e^{i\mathbf{K}\cdot\mathbf{r}_0} + e^{i(\hat{C}_3\mathbf{K}\cdot\mathbf{r}_0 - \frac{2\pi}{3})} + e^{i(\hat{C}_3^2\mathbf{K}\cdot\mathbf{r}_0 - \frac{4\pi}{3})}|^2. \end{aligned}$$

For $\mathbf{r}_0 = 0$, $f(\mathbf{r}_0) = 0$ and $\Delta E_c = 2|c'|\hat{H}_t|c|$, $\Delta E_v = 2|v'|\hat{H}_t|v|$, which agrees with our symmetry analysis that the two layers of AA stacking are related by $\hat{\sigma}_h$ and thus are equivalent. On the other hand, $\langle c'|\hat{H}_t|c\rangle = \langle v'|\hat{H}_t|v\rangle = 0$ for $\mathbf{r}_0 = \pm \frac{\mathbf{a}_1 + \mathbf{a}_2}{3}$, which leads to $\Delta E_c = 9|\delta E_c^{(0)}|$ and $\Delta E_v = 9|\delta E_v^{(0)}|$.

We have also performed *ab initio* calculations for $\Delta E_{c/v}$ in MoS₂ R-type homobilayers with different \mathbf{r}_0 . The calculation details are the same as in the H-type case, and the results are presented in Fig. 3(b). Once gain we show both the *ab initio* results (symbols) and the corresponding fits (solid lines) using Eqs. (16) and (8) in Fig. 3(c) with \mathbf{r}_0 along the long diagonal line of the unit cell. Keeping only the main hopping terms $\delta E_{c/v}^{(0)}$ and $t_{cc'/vv'}^{(0)}$ (dashed lines) can already reproduce the major features, whereas the fits using $\delta E_{c/v}^{(0)}$, $t_{cc'/vv'}^{(0)}$, and $t_{cc'/vv'}^{(1)}$ terms (solid lines) agree almost perfectly with the *ab initio* results. The fitting parameters are summarized in Table III. As a good approximation, the \mathbf{K} point conduction/valence band interlayer hoppings in R-type commensurate bilayers take the forms

$$\begin{aligned} |\langle c|\hat{H}_t|c'\rangle_R| &\approx |e^{i\mathbf{K}\cdot\mathbf{r}_0} + e^{i\hat{C}_3\mathbf{K}\cdot\mathbf{r}_0} + e^{i\hat{C}_3^2\mathbf{K}\cdot\mathbf{r}_0}|t_{cc'}^{(0)}, \\ |\langle v|\hat{H}_t|v'\rangle_R| &\approx |e^{i\mathbf{K}\cdot\mathbf{r}_0} + e^{i\hat{C}_3\mathbf{K}\cdot\mathbf{r}_0} + e^{i\hat{C}_3^2\mathbf{K}\cdot\mathbf{r}_0}|t_{vv'}^{(0)}. \end{aligned} \quad (17)$$

TABLE III. The obtained hopping strengths for the R-type homobilayer MoS₂ from fitting to the *ab initio* results of band splitting. The main hopping term $t_{vv'}^{(0)}$ is consistent with our previous result [39].

$ \delta E_c^{(0)} $	$t_{cc'}^{(0)}$	$t_{cc'}^{(1)}$	$ \delta E_v^{(0)} $	$t_{vv'}^{(0)}$	$t_{vv'}^{(1)}$
6.3 meV	2.1 meV	0.6 meV	7.5 meV	14.5 meV	1.6 meV

The equations above also apply to R-type commensurate heterobilayers. Similar forms have been obtained in early papers [39,42,45,50].

From the above equations, we get $t_{cc'}^{(0)} \approx \Delta E_c/6$ and $t_{vv'}^{(0)} \approx \Delta E_v/6$ at $\mathbf{r}_0 = 0$ for R stacking.

C. Variation of coupling strength with interlayer distance

As shown in both the theoretical analysis above and the good fit results in Figs. 2(c) and 3(c), $t_{nn'}^{(j)}$ does not directly depend on the interlayer translation \mathbf{r}_0 . However, $t_{nn'}^{(j)}$ should sensitively depend on the interlayer distance d , the equilibrium value of which varies in a large range depending on the stacking pattern in R- or H-type commensurate bilayers characterized by \mathbf{r}_0 [32,33]. A recent scanning tunneling microscopy/spectroscopy experiment has shown that in a single heterobilayer structure with the formation of large scale moiré superlattice, d can vary from position to position due to the variation of local stacking patterns [46].

We have calculated these $\Delta E_{c/v}$ as functions of d , which can be well fitted by exponential functions $\Delta E_n(d) = \Delta E_{0,n}e^{-d/d_n}$. Here $\Delta E_{0,c} = 1.96$ eV, $d_c = 0.62$ Å and $\Delta E_{0,v} = 14.4$ eV, $d_v = 0.59$ Å for R stacking, $\Delta E_{0,c} = 1.77$ eV, $d_c = 0.63$ Å and $\Delta E_{0,v} = 12.7$ eV, $d_v = 0.61$ Å for H stacking.

Considering the similarity of the d orbitals of Mo and W atoms, the hopping strengths for the homobilayers shall provide reasonable estimations to those in the TMD heterobilayers. However, in heterobilayers the $\pm\mathbf{K}$ valleys have much larger conduction/valence band offsets, which leads to negligible layer mixing [51,52]. For example, in MoS₂/WSe₂ heterobilayer, the $\pm\mathbf{K}$ -valley valence (conduction) band offset is found to be 0.83 eV (0.76 eV) [51]. While in MoSe₂/WSe₂ heterobilayer, the valence band offset is 0.3 eV [52]. These values are all much larger than the $\pm\mathbf{K}$ valley coupling strengths which are on the order of several tens of meV. Thus unlike the homobilayers where $t_{cc'/vv'}$ should be treated nonperturbatively, in heterobilayers all $t_{cn'/vn'}$ hopping terms can be treated perturbatively.

D. Interlayer coupling strengths of other bands

Just like the c and v bands, the band splitting values ΔE_{c+j} (ΔE_{v-j}) of other bands in an AA-type homobilayer give the corresponding interlayer coupling strengths $t_{c+j,c'+j}^{(0)}$ ($t_{v-j,v'-j}^{(0)}$). We show the band structure of an AA-type homobilayer MoS₂ in Fig. 4(c). The extracted ΔE_n for n from $c+3$ to $v-3$ are summarized in Table IV. Note that some bands have much larger interlayer coupling than those of the c and v bands, which could be related to their larger p -orbital

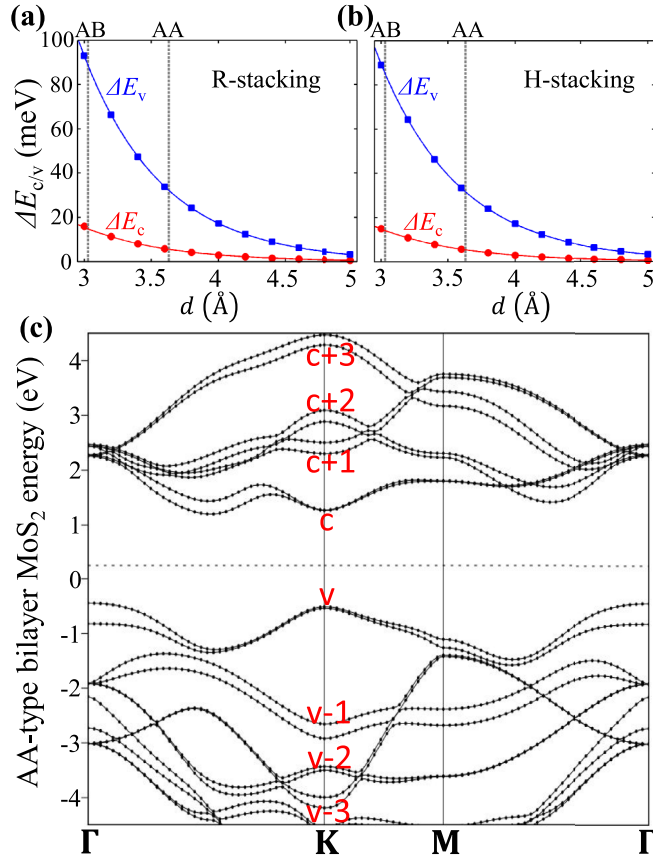


FIG. 4. (a) The solid dots are our *ab initio* results of conduction/valence band splitting at \mathbf{K} point for MoS₂ R-type homobilayers with $\mathbf{r}_0 = 0$ (AA stacking) as functions of interlayer distance d (defined as the vertical distance between the nearest chalcogen atoms of neighboring layers). The solid curves are the exponential fits. These results are also presented in Ref. [39]. The vertical dashed lines show the numerical values of interlayer distance for AB (3.0 Å) and AA (3.6 Å) homobilayer MoS₂, adopted from Ref. [32]. (b) The case for MoS₂ H-type homobilayers. ΔE_c is for $\mathbf{r}_0 = 0$, while ΔE_v is for $\mathbf{r}_0 = (\mathbf{a}_1 + \mathbf{a}_2)/3$ (AB stacking). (c) The *ab initio* band structure of an AA-type MoS₂ homobilayer without spin-orbit coupling, where the band splittings ΔE_n can be clearly seen. The interlayer distance is set as $d_{AA} = 3.72$ Å, and the other calculation details are the same as those in Figs. 2 and 3.

proportions in the corresponding orbital combinations D_n (see Table I). It is then natural to expect that the hopping terms $t_{cn'}^{(0)}$ ($t_{vn'}^{(0)}$) between the remote band n' and the conduction band c (valence band v) are much larger than $t_{cc'}^{(0)}$ ($t_{vv'}^{(0)}$), which can lead to a large range modulation of the c and v band energy with \mathbf{r}_0 . This is consistent with the observed ~ 0.1 eV band gap modulation in a TMD heterobilayer [46].

TABLE IV. The band splittings ΔE_n (in unit of meV) for n from $c+3$ to $v-3$ extracted from Fig. 4(c).

$c+3$	$c+2$	$c+1$	c	v	$v-1$	$v-2$	$v-3$
183	206	203	5	27	265	67	193

V. TWISTED OR LATTICE-MISMATCHED BILAYER STRUCTURES AND MOIRÉ PATTERNS

The interlayer couplings in twisted or lattice-mismatched bilayers can also be described by Eq. (5). Note that the local atomic orbital $D_n(\mathbf{r})$ in the upper layer is rotated by the interlayer twist angle θ with respect to $D_n(\mathbf{r})$ in the lower layer. So in principle the corresponding hopping term $t_{nn'}$ for $\theta \neq 0^\circ$ and 60° should be different from those given in the previous discussion of H-type or R-type commensurate bilayers. However, when considering the cases with close to 0° or 60° twist angle, it is a good approximation to replace $t_{nn'}$ by those of the H-type or R-type commensurate bilayers.

A twisted or lattice-mismatched bilayer can still be commensurate under special conditions, i.e., the two layers form a periodic superlattice structure with the supercell size larger than the monolayer unit cell. As the commensurability is irrelevant to the interlayer translation, we assume a metal atom in the upper layer horizontally overlaps with a metal atom in the lower layer at the xy -plane origin. In the commensurate case, the bilayer supercell is then given by the smallest rhombus with its four vertices located at the overlapping metal atoms, as shown in Fig. 5(a). Notice that in \mathbf{k} space, $\tau\kappa$ in the

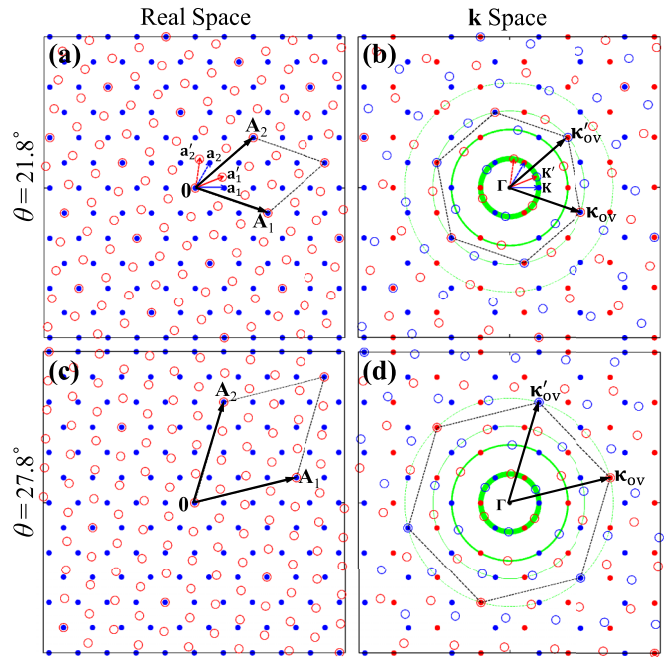


FIG. 5. (a) The real space atomic registry of a lattice-matched commensurate bilayer with $\theta = 21.8^\circ$ twist angle. The solid blue (empty red) dots denote the metal atoms in the lower (upper) layer, with $\mathbf{a}_{1,2}$ ($\mathbf{a}'_{1,2}$) the corresponding unit lattice vectors. The rhombus corresponds to a supercell, with its four vertices located at 0, \mathbf{A}_1 , \mathbf{A}_2 , and $\mathbf{A}_1 + \mathbf{A}_2$, where two metal atoms in opposite layers horizontally overlap. (b) The corresponding \mathbf{k} -space configurations of the two layers. The solid blue (red) dots correspond to κ ($-\kappa$) in the lower layer, and the empty blue (red) dots correspond to κ' ($-\kappa'$) in the upper layer. The six overlapping ($\tau\kappa, \tau'\kappa'$) pairs on the third smallest green circle form a hexagon (dashed lines). The \mathbf{k} -space hexagon corner κ_{ov} (κ'_{ov}) corresponds to \mathbf{A}_1 (\mathbf{A}_2) in the real space. (c) and (d) Another commensurate bilayer with twist angle $\theta = 27.8^\circ$ with larger supercell size and $|\kappa_{ov}|$.

lower layer overlaps with $\tau'\kappa'$ in the upper layer at certain positions κ_{ov} , which means $\tau\mathbf{K}$ and $\tau'\mathbf{K}'$ are coupled through the interlayer hopping $t_{nn'}(\kappa_{ov})$ [see Eq. (5)]. Interestingly, there is one-to-one correspondence between the superlattice unit vector $\mathbf{A}_{1,2}$ and κ_{ov} such that $|\kappa_{ov}| = \frac{4\pi}{3aa'}|\mathbf{A}_{1,2}|$, as shown in Fig. 5(b). So larger supercell size corresponds to larger $|\kappa_{ov}|$ and thus smaller coupling strength $|t_{nn'}(\kappa_{ov})|$ between $\tau\mathbf{K}$ and $\tau'\mathbf{K}'$, which agrees with the findings in a recent work [31]. In fact, in a twisted or lattice-mismatched commensurate bilayer, κ_{ov} always corresponds to second or higher order hopping terms $t_{nn'}^{(j)}$ ($j \geq 2$), which are negligibly small compared to the main terms $t_{nn'}^{(0)}$.

Away from the band edges $\tau\mathbf{K}$ and $\tau'\mathbf{K}'$, the interlayer coupling can be significant. We can always find small wave vectors \mathbf{k}' and \mathbf{k} where $\mathbf{k}' - \mathbf{k}$ equals to $\tau\mathbf{K} - \tau'\mathbf{K}'$ or $\tau\hat{C}_3\mathbf{K} - \tau'\hat{C}_3\mathbf{K}'$ or $\tau\hat{C}_3^2\mathbf{K} - \tau'\hat{C}_3^2\mathbf{K}'$. According to Eq. (5), the coupling between $\tau\mathbf{K} + \mathbf{k}$ and $\tau'\mathbf{K}' + \mathbf{k}'$ is then $\sim t_{nn'}^{(0)}$, which corresponds to the main hopping term. Note that such coupling terms are insensitive to whether the bilayer is commensurate or not. As discussed above, in a twisted or lattice-mismatched bilayer the commensurability only introduces direct coupling between the two band edges $\tau\mathbf{K}$ and $\tau'\mathbf{K}'$, with a negligibly small coupling strength.

The interlayer coupling between $\tau\mathbf{K} + \mathbf{k}$ and $\tau'\mathbf{K}' + \mathbf{k}'$ discussed above is especially important for bilayers with $\tau\mathbf{K}$ and $\tau'\mathbf{K}'$ close to each other, in which $|\mathbf{k}|$ and $|\mathbf{k}'|$ can be small enough that low energy carriers in different layers are efficiently coupled. On the other hand, it is known that in such a bilayer with $|\tau\mathbf{K} - \tau'\mathbf{K}'| \ll 4\pi/3a$, a moiré superlattice pattern with large scale periodicity will form [14–16,30,50], as shown in Fig. 6(a). Below we show that the moiré superlattice picture is fully consistent with our theoretical analysis in Sec. II.

We note that the moiré pattern is not a rigorous periodic structure but a good approximation, whose emergence can be understood as follows. Any quantity involving the periodicity of both layers [e.g., $\psi_{n,\mathbf{k}}^* \psi_{n',\mathbf{k}'}$ which appears in the hopping integral in Eq. (3)] can be written as the sum of all $e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}}$ terms by a Fourier transformation. Here $\mathbf{G} = j_1\mathbf{b}_1 + j_2\mathbf{b}_2$ ($\mathbf{G}' = j'_1\mathbf{b}'_1 + j'_2\mathbf{b}'_2$) are the lower (upper) layer reciprocal lattice vectors, with $\mathbf{b}_{1,2}$ ($\mathbf{b}'_{1,2}$) the corresponding primitive reciprocal lattice vectors and $j_{1,2}$, $j'_{1,2}$ integers. Those terms with large $|\mathbf{G}|$ or $|\mathbf{G}'|$ are related to the fast oscillating components in $\psi_{n,\mathbf{k}}$ or $\psi_{n',\mathbf{k}'}$ with periods much smaller than the lattice constant, and can be dropped. Then the remaining slowly oscillating terms always have $\mathbf{G} - \mathbf{G}' = j_1(\mathbf{b}_1 - \mathbf{b}'_1) + j_2(\mathbf{b}_2 - \mathbf{b}'_2)$. Thus the large scale moiré period is characterized by the primitive reciprocal lattice vectors $\mathbf{B}_1 \equiv \mathbf{b}_1 - \mathbf{b}'_1$ and $\mathbf{B}_2 \equiv \mathbf{b}_2 - \mathbf{b}'_2$. The above analysis requires $|\mathbf{B}_{1,2}| \approx \frac{4\pi}{\sqrt{3}a} \sqrt{\delta^2 + \delta\theta^2} \ll \frac{4\pi}{3a}$, with $\delta = a/a' - 1$ and $\delta\theta$ the twist angle deviation to 0 or $\pi/3$. The moiré superlattice constant is then $A \approx a/\sqrt{\delta^2 + \delta\theta^2} \gg a$, with $|\delta| \ll 1$ and $|\delta\theta| \ll 1$ the prerequisites for the existence of a moiré pattern.

The moiré superlattice mini Brillouin zone (BZ) has its corners located at $\tau\mathbf{K} - \tau'\mathbf{K}'$ and its $\pi/3$ rotations [see Fig. 6(b)][42,50,53]. The mini BZ forms a complete basis in \mathbf{k} space for the bilayer structure. Inside the mini BZ, the original monolayer bands are folded into a series of closely spaced mini

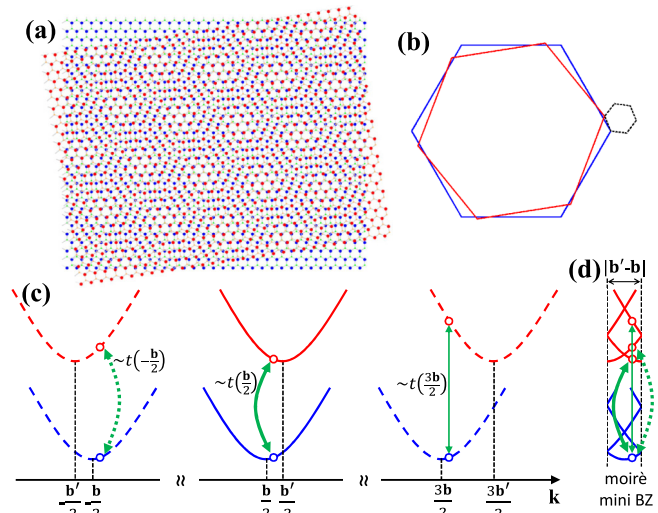


FIG. 6. (a) A typical hexagonal bilayer moiré pattern with lattice constant mismatch $\delta = 0.05$ and twist angle $\theta = 9^\circ$. (b) The monolayer BZs (solid blue and red hexagons) and the moiré superlattice mini BZ (dashed black hexagon). (c) The band dispersions of two 1D systems (denoted as blue and red colors). Only those near the band edges located at $\mathbf{b}/2$ and $\mathbf{b}'/2$ are shown. The double arrows indicate the $t(j\mathbf{b} + \mathbf{b}/2 + \mathbf{k})$ hopping terms with $j = 0, \pm 1$ [Eq. (19)], and the arrow thickness corresponds to the hopping strength. (d) The corresponding interlayer hopping terms (double arrows) between different mini bands in the 1D moiré mini BZ.

bands, and a state with wave vector \mathbf{k} in one layer can hop to various mini bands in the other layer with the same \mathbf{k} . We note that for small $|\kappa|$ and $|\kappa'|$, the delta function in Eq. (5) can be written as $\delta_{\mathbf{k}-\mathbf{k}', \tau\kappa-\tau'\kappa'} = \delta_{(\tau\mathbf{K}+\mathbf{k})-(\tau'\mathbf{K}'+\mathbf{k}'), j_1\mathbf{B}_1+j_2\mathbf{B}_2}$, which is just the momentum conservation condition in the mini BZ picture. The effect of the interlayer coupling is to open gaps between the mini bands.

For convenience we use two 1D systems to illustrate the correspondence between the individual BZs and the moiré mini BZ in Figs. 6(c) and 6(d). The band edges of the individual 1D systems are assumed to be located at $\mathbf{b}/2$ and $\mathbf{b}'/2$, where \mathbf{b} and \mathbf{b}' are the primitive reciprocal lattice vectors of the corresponding systems and $|\mathbf{b}' - \mathbf{b}| \ll |\mathbf{b}|$. Following Eq. (2) the Bloch states can be written as

$$\begin{aligned} \psi_{\mathbf{b}/2+\mathbf{k}}(\mathbf{r}) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i(\mathbf{b}/2+\mathbf{k})\cdot\mathbf{R}} D(\mathbf{r} - \mathbf{R}), \\ \psi_{\mathbf{b}'/2+\mathbf{k}'}(\mathbf{r}) &= \frac{1}{\sqrt{N'}} \sum_{\mathbf{R}'} e^{i(\mathbf{b}'/2+\mathbf{k}')\cdot\mathbf{R}'} D(\mathbf{r} - \mathbf{R}'). \end{aligned} \quad (18)$$

Here $D(\mathbf{r} - \mathbf{R})$ is the atomic orbital combination localized near \mathbf{R} . Analogous to Eq. (5), we write the hopping integral between the two 1D systems as

$$\begin{aligned} &\int \psi_{\mathbf{b}/2+\mathbf{k}}^*(\mathbf{r}) \hat{H}_t \psi_{\mathbf{b}'/2+\mathbf{k}'}(\mathbf{r}) d\mathbf{r} \\ &= \sum_{jj'} \delta_{(j+\frac{1}{2})\mathbf{b}+\mathbf{k}, (j'+\frac{1}{2})\mathbf{b}'+\mathbf{k}'} t(j\mathbf{b} + \mathbf{b}/2 + \mathbf{k}) e^{-ij\mathbf{b}\cdot\mathbf{r}_0}. \end{aligned} \quad (19)$$

Here $t(j\mathbf{b} + \mathbf{b}/2 + \mathbf{k})$ with $j = 0, \pm 1, \dots$ are the Fourier transformations of the hopping integral between the two localized orbitals, which are indicated as double arrows near $(j + 1/2)\mathbf{b}$ in Fig. 6(c). These terms with different j have one-to-one correspondence with those between different mini bands in the moiré mini BZ, as shown in Fig. 6(d).

Although the individual BZ picture is equivalent to the moiré mini BZ as discussed above, we find that it is more convenient to extract the hopping strength using the former picture. Considering that the magnitude of $t(\mathbf{q})$ decays fast with the increase of $|\mathbf{q}|$, in the individual BZ picture we can just focus on the hopping terms $t(\mathbf{q})$ with \mathbf{q} inside the first BZs. Whereas in the moiré mini BZ picture we cannot directly get which two mini bands have a strong hopping strength.

On the other hand, a local picture becomes more convenient for describing large scale moiré superlattices [39,40,46]. We can consider a local region with a size much larger than the monolayer unit cell, but at the same time much smaller than the moiré supercell. The corresponding atomic registry is locally indistinguishable from an R- or H-type commensurate bilayer, thus we can discuss its local band structure which is given by that of the corresponding commensurate bilayer. Different local regions are characterized by different \mathbf{r}_0 , which results in a periodic modulation of the local band structure. In TMD heterobilayers where the $\pm\mathbf{K}$ valleys have negligible layer mixing, the interlayer coupling appears as a local band structure modulation, which is equivalent to applying band-dependent external superlattice potentials on two decoupled layers [40].

VI. INTERLAYER COUPLING IN Γ_V AND Q_C VALLEYS

In 2H homobilayer TMDs, the Γ_V and Q_C energies are strongly shifted away from the corresponding monolayer positions as evidenced by the photoluminescence and ARPES measurements [32,33,54], which is a signature of the strong interlayer coupling near these positions [5]. Here Γ_V denotes the Γ point of valence band, and Q_C denotes the six conduction band extrema near the middle of the Γ - $\tau\mathbf{K}$ lines [Fig. 7(a)]. From the *ab initio* results of homobilayer band structures, we estimate that the interlayer hopping strengths in the Γ_V and Q_C valleys are on the order of several hundred meV [Fig. 7(b)].

Note that all Q_C points are located on a ring with radius $\sim |\mathbf{K}|/2$ [Fig. 7(a)], while Fig. 7(b) indicates a strong interlayer coupling near the conduction band $M/2$ point (the middle of the Γ - \mathbf{M} line). Thus we speculate that all conduction band \mathbf{k} points on this ring have strong interlayer couplings. Furthermore, for an arbitrary interlayer twist angle, the Γ positions are not affected and the Q_C valleys are always on this ring. Therefore, we expect that the interlayer twist does not change the strong coupling nature of Γ_V and Q_C valleys.

We have also calculated the band splitting ΔE_Γ (ΔE_Q) at Γ_V (Q_C) point for AA-type MoS_2/WS_2 heterobilayers, which is found to depend sensitively on the interlayer distance d . The band splitting can be approximated as $\Delta E_{\Gamma/Q} = \sqrt{(\Delta E_{0,\Gamma/Q})^2 + 4t_{\Gamma/Q}^2}$, when ignoring the coupling with other bands. For a large enough d , i.e., under the vanishing interlayer coupling limit, the interlayer hopping strength $t_{\Gamma/Q} \rightarrow 0$, from which we get the band offset values $\Delta E_{0,Q} = 0.3$ eV and

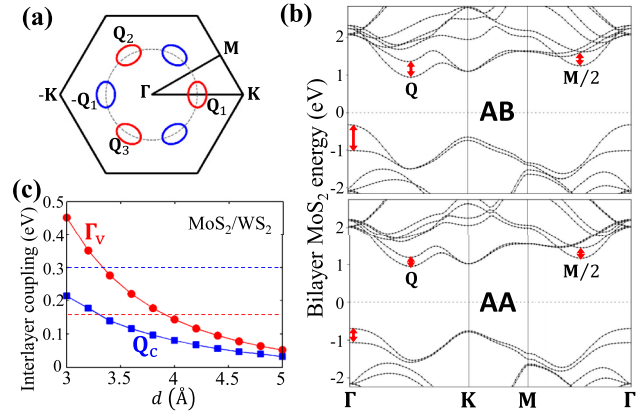


FIG. 7. (a) The red and blue pockets illustrate the energy contours of the six Q_C valleys. The dashed circle corresponds to a ring-shaped region with strong conduction band interlayer coupling. (b) The *ab initio* band structures of AB and AA homobilayer MoS_2 with interlayer distances $d_{AB} = 2.975$ Å and $d_{AA} = 3.72$ Å, respectively. The spin-orbit coupling is not considered. The band splittings at Γ_V , Q_C , and $M/2$ are denoted by red arrows. (c) The obtained interlayer coupling strength at Γ_V and Q_C for AA-type MoS_2/WS_2 heterobilayers. The red (blue) dashed line shows the interlayer band offset $\Delta E_{0,\Gamma} = 0.16$ eV ($\Delta E_{0,Q} = 0.3$ eV).

$\Delta E_{0,\Gamma} = 0.16$ eV. $t_{\Gamma/Q}$ for intermediate values of d are then derived from the relation above and shown in Fig. 7(c).

The Γ_V and Q_C valley Bloch functions can be approximated similar to Eq. (2) for the $\pm\mathbf{K}$ valley

$$\begin{aligned} \psi_{\Gamma,\mathbf{k}}(\mathbf{r}) &\approx \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} D_{\Gamma}(\mathbf{r} - \mathbf{R}), \\ \psi_{\tau Q_j,\mathbf{k}}(\mathbf{r}) &\approx \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i(\tau Q_j + \mathbf{k})\cdot\mathbf{R}} D_{\tau Q}(\mathbf{r} - \mathbf{R}). \end{aligned} \quad (20)$$

Here we use τQ_j with $\tau = \pm$ and $j = 1, 2, 3$ to distinguish the six degenerate but inequivalent Q_C [Fig. 7(a)], which are related by \hat{C}_3 or time reversal operations. $D_{\Gamma}(\mathbf{r} - \mathbf{R})$ and $D_{\tau Q}(\mathbf{r} - \mathbf{R})$ are the linear combinations of atomic orbitals localized around \mathbf{R} for the corresponding valleys. Following the derivation of Eq. (5), the hopping strength can be written as

$$\begin{aligned} \langle \Gamma, \mathbf{k} | \hat{H}_t | \Gamma', \mathbf{k}' \rangle &\approx t_{\Gamma}(\mathbf{k}) \delta_{\mathbf{k}', \mathbf{k}}, \\ \langle \tau Q_j, \mathbf{k} | \hat{H}_t | \tau' Q'_j, \mathbf{k}' \rangle &\approx t_Q(\tau Q_j + \mathbf{k}) \delta_{\tau' Q'_j + \mathbf{k}', \tau Q_j + \mathbf{k}}. \end{aligned}$$

In the last step above, we have used the fact that $\mathbf{G} + \mathbf{k}$ and $\tau Q_j + \mathbf{G} + \mathbf{k}$ are well outside the monolayer first BZ when $\mathbf{G} \neq 0$, and the corresponding $t_{\Gamma}(\mathbf{G} + \mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}_0}$ and $t_Q(\tau Q_j + \mathbf{G} + \mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}_0}$ terms have much smaller magnitudes than those at $\mathbf{G} = 0$ and can be ignored. Thus, unlike the $\pm\mathbf{K}$ valleys [Eqs. (13) and (17)] discussed previously, the Γ_V and Q_C valley interlayer couplings are nearly independent of the interlayer translation \mathbf{r}_0 .

The interlayer coupling strengths of Γ_V and Q_C valleys are comparable to the corresponding band offsets in TMD heterobilayers [Fig. 7(c)], which is distinct from the $\pm\mathbf{K}$ valleys. The strong interlayer couplings of Γ_V and Q_C valleys originate from: (1) the nonignorable p_z orbital of chalcogen

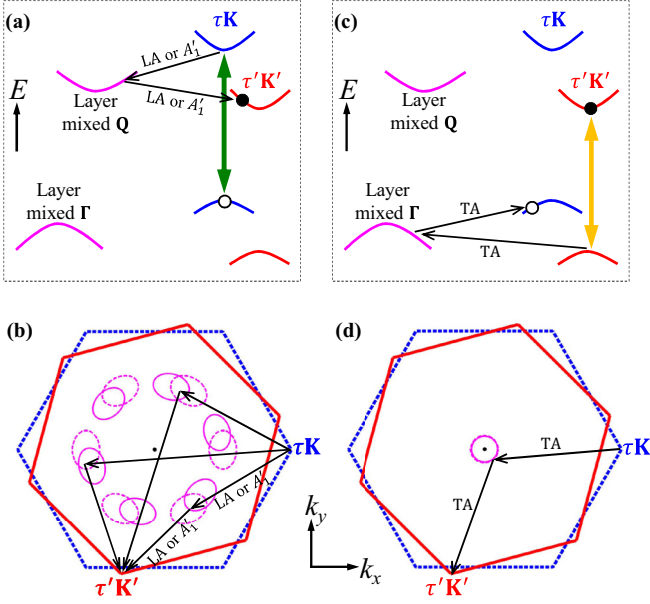


FIG. 8. (a) Schematic illustration of the electron interlayer charge transfer process in the energy space. The blue (red) curves are the lower (upper) layer $\tau\mathbf{K}$ ($\tau'\mathbf{K}'$) valley bands, and the purple curves are the strongly layer mixed Γ_v and \mathbf{Q}_c valleys. The double arrow illustrates the optical generation of electron-hole pairs in the $\tau\mathbf{K}$ valley. The single arrows correspond to the electron relaxation pathways. (b) The electron interlayer charge transfer process in the momentum space. The dashed blue (solid red) hexagon is the lower (upper) layer BZ. The electron can be scattered to three \mathbf{Q}_c valleys through emitting a phonon with wave vector \mathbf{M} . (c) and (d) Schematic illustration of the hole interlayer charge transfer.

atoms [5] in D_Γ and $D_{\tau\mathbf{Q}}$; and (2) the fact that they correspond to the $t(\mathbf{q})$ Fourier components with $|\mathbf{q}| < |\mathbf{K}|$. The resulting strong layer mixing can play an important role in the interlayer charge transfer processes of TMD heterobilayers with type-II band alignments [18–27].

Experiments have found that the charge transfer process is ultrafast (<50 fs) and independent on the interlayer twist [18–20], both of which cannot be explained by the weak interlayer coupling strength of the $\pm\mathbf{K}$ valleys. Here we propose the following electron (hole) interlayer charge transfer mechanism mediated by the \mathbf{Q}_c (Γ_v) valley. For a type-II heterobilayer with the conduction (valence) band edge located at $\tau'\mathbf{K}'_c$ ($\tau\mathbf{K}_v$), a high energy electron in $\tau\mathbf{K}_c$ valley can relax to one of the \mathbf{Q}_c valleys through scattering with phonons, other carriers, or impurities/defects. As \mathbf{Q}_c valleys are strongly layer mixed, this electron can further relax to the $\tau'\mathbf{K}'_c$ valley, as shown in Figs. 8(a) and 8(b). A high energy hole in $\tau'\mathbf{K}'_v$ can relax to the strongly layer mixed Γ_v valley and then to $\tau\mathbf{K}_v$, see Figs. 8(c) and 8(d). We expect that such interlayer charge transfer rate is close to the $\pm\mathbf{K}$ valley carrier relaxation rate in few-layer or bulk TMDs, since they both involve intervalley relaxation from $\pm\mathbf{K}$ to \mathbf{Q}_c or Γ_v . Actually the measured intervalley relaxation time in few-layer MoS₂ is ~ 20 fs [55], which indeed agrees well with the interlayer charge transfer time (<50 fs) in heterobilayer TMDs [18–20].

DFT calculations suggest that electrons in \mathbf{K}_c valley couple strongly with LA and A'_1 phonons with wave vectors in the

vicinity of \mathbf{M} [56], which leads to scatterings between \mathbf{K}_c and $-\mathbf{Q}_{1,2,3}$ valleys. On the other hand, holes in \mathbf{K}_v valley couple strongly with TA phonons with wave vectors in the vicinity of $-\mathbf{K}$ [57], which leads to scatterings between \mathbf{K}_v and Γ_v valleys. Using the Fermi golden rule, we can estimate the phonon emission assisted electron/hole intervalley scattering rates as

$$\frac{1}{\tau_e} = \frac{2\pi}{\hbar} \sum_{\mathbf{q}} \frac{|g_{e,\mathbf{q}}|^2}{N} \delta(E_{c,\mathbf{Q},\mathbf{q}} + \hbar\omega - E_{c,\mathbf{K}}),$$

$$\frac{1}{\tau_h} = \frac{2\pi}{\hbar} \sum_{\mathbf{q}} \frac{|g_{h,\mathbf{q}}|^2}{N} \delta(E_{v,\Gamma,\mathbf{q}} - \hbar\omega - E_{v,\mathbf{K}'}). \quad (21)$$

Here $\frac{1}{\sqrt{N}}g_{e/h,\mathbf{q}}$ are the electron-phonon coupling matrix elements with N the lattice number. In monolayer MoS₂, DFT calculation gives $g_{e,\mathbf{q}}^0 \sim 0.11$ eV (0.08 eV) for LA (A'_1) phonons with wave vectors in the vicinity of \mathbf{M} [56], and $g_{h,\mathbf{q}}^0 \sim 0.1$ eV for TA phonons with wave vectors in the vicinity of $-\mathbf{K}$ [57]. From the interlayer coupling strength and band offset values given in Fig. 7(c), we assume 20% (50%) of the involved \mathbf{Q}_c (Γ_v) valley in the heterobilayer is in the layer of the initial \mathbf{K} electron (\mathbf{K}' hole), which then leads to $g_{e,\mathbf{q}} \sim \sqrt{0.2}g_{e,\mathbf{q}}^0$ ($g_{h,\mathbf{q}} \sim \sqrt{0.5}g_{h,\mathbf{q}}^0$). We also use the effective mass approximation for the band dispersions $E_{c,\mathbf{Q},\mathbf{q}} \approx E_{c,\mathbf{Q}} + \frac{\hbar^2 q^2}{2m_{\mathbf{Q}}^*}$ and $E_{v,\Gamma,\mathbf{q}} \approx E_{v,\Gamma} - \frac{\hbar^2 q^2}{2m_{\Gamma}^*}$. Using the value $m_{\mathbf{Q}}^* \sim m_0$ [2], and taking into account both the LA, A'_1 phonons, and the three possible pathways shown in Fig. 8(b), we get $\tau_e \sim 50$ fs. For the hole we use $m_{\Gamma}^* \sim 2m_0$ [2], which results in $\tau_h \sim 50$ fs. They agree well with the experimental value (<50 fs) for the interlayer charge transfer process [18–20].

The interlayer charge transfer mechanism proposed above is also consistent with the insensitivity to the interlayer twist, because the strong layer mixing nature of Γ_v and \mathbf{Q}_c valleys is not affected. This is obvious for Γ_v , where the interlayer coupling strength and band offset are not affected by the twist angle. For \mathbf{Q}_c valleys, they are always on the ring region with strong interlayer coupling for any twist angle. Meanwhile, considering the large \mathbf{Q}_c valley effective mass in the direction perpendicular to the Γ - $\tau\mathbf{K}$ line [2], the twist angle does not change the interlayer band offset much. Therefore the strong layer mixing of \mathbf{Q}_c valleys are unaffected by the interlayer twist. For TMD heterobilayers with arbitrary stacking, the interlayer charge transfer can efficiently happen through emitting two intralayer phonons, one in the upper layer and the other in the lower layer.

VII. CONCLUSION

In conclusion, the interlayer couplings in $\pm\mathbf{K}$, Γ_v , and \mathbf{Q}_c valleys of commensurate and incommensurate TMD bilayer structures are studied. The coupling strengths in $\pm\mathbf{K}$ valleys depend sensitively on the interlayer translation for R- and H-type commensurate bilayers, which can explain the observed band gap modulation in TMD heterobilayers with large scale moiré pattern. The coupling strengths for Γ_v and \mathbf{Q}_c valleys are huge and insensitive to both the interlayer translation and twist angle. The resulted strong layer mixing of Γ_v and \mathbf{Q}_c can mediate the twist-insensitive and ultrafast interlayer charge transfer in TMD heterobilayers. We expect that the results presented

in this paper would be meaningful and illuminating for further exploring the rich physics and potential applications in various commensurate and incommensurate TMD bilayer structures.

ACKNOWLEDGMENTS

We thank T. Cao for helpful discussions and P. Rivera for proofreading. Y.W. and Z.W. were supported by NSFC

with Grants No. 11604162 and No. 61674083. Y.W. was also supported by UGC of Hong Kong (AoE/P-04/08). G.-B.L. was supported by NSFC with Grant No. 11304014 and the China 973 Program with Grant No. 2013CB934500. H.Y. and W.Y. were supported by the RGC and UGC of Hong Kong (HKU17312916, AoE/P-04/08), and the HKU ORA.

-
- [1] J. R. Schaibley, H. Yu, G. Clark, P. Rivera, J. S. Ross, K. L. Seyler, W. Yao, and X. Xu, *Nat. Rev. Mater.* **1**, 16055 (2016).
- [2] A. Kormányos, G. Burkard, M. Gmitra, J. Fabian, V. Zólyomi, N. D. Drummond, and V. Fal'ko, *2D Mater.* **2**, 022001 (2015).
- [3] K. F. Mak and J. Shan, *Nat. Photon.* **10**, 216 (2016).
- [4] D. Jariwala, V. K. Sangwan, L. J. Lauhon, T. J. Marks, and M. C. Hersam, *ACS Nano* **8**, 1102 (2014).
- [5] G.-B. Liu, D. Xiao, Y. Yao, X. Xu, and W. Yao, *Chem. Soc. Rev.* **44**, 2643 (2015).
- [6] Z. Gong, G.-B. Liu, H. Yu, D. Xiao, X. Cui, X. Xu, and W. Yao, *Nat. Commun.* **4**, 2053 (2013).
- [7] S. Wu, J. S. Ross, G.-B. Liu, G. Aivazian, A. Jones, Z. Fei, W. Zhu, D. Xiao, W. Yao, D. Cobden, and X. Xu, *Nat. Phys.* **9**, 149 (2013).
- [8] A. M. Jones, H. Yu, J. S. Ross, P. Klement, N. J. Ghimire, J. Yan, D. G. Mandrus, W. Yao, and X. Xu, *Nat. Phys.* **10**, 130 (2014).
- [9] H. Yuan, M. S. Bahramy, K. Morimoto, S. Wu, K. Nomura, B.-J. Yang, H. Shimotani, R. Suzuki, M. Toh, C. Kloc, X. Xu, R. Arita, N. Nagaosa, and Y. Iwasa, *Nat. Phys.* **9**, 563 (2013).
- [10] B. Zhu, H. Zeng, J. Dai, Z. Gong, and X. Cui, *Proc. Natl. Acad. Sci. USA* **111**, 11606 (2014).
- [11] T. Jiang, H. Liu, D. Huang, S. Zhang, Y. Li, X. Gong, Y.-R. Shen, W.-T. Liu, and S. Wu, *Nat. Nanotechnol.* **9**, 825 (2014).
- [12] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, *Phys. Rev. Lett.* **105**, 136805 (2010).
- [13] A. J. Bradley, M. M. Ugeda, F. H. da Jornada, D. Y. Qiu, W. Ruan, Y. Zhang, S. Wickenburg, A. Riss, J. Lu, S.-K. Mo, Z. Hussain, Z.-X. Shen, S. G. Louie, and M. F. Crommie, *Nano Lett.* **15**, 2594 (2015).
- [14] L. A. Ponomarenko, R. V. Gorbachev, G. L. Yu, D. C. Elias, R. Jalil, A. A. Patel, A. Mishchenko, A. S. Mayorov, C. R. Woods, J. R. Wallbank, M. Mucha-Kruczynski, B. A. Piot, M. Potemski, I. V. Grigorieva, K. S. Novoselov, F. Guinea, V. I. Fal'ko, and A. K. Geim, *Nature (London)* **497**, 594 (2013).
- [15] C. R. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, P. Moon, M. Koshino, T. Taniguchi, K. Watanabe, K. L. Shepard, J. Hone, and P. Kim, *Nature (London)* **497**, 598 (2013).
- [16] 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).
- [17] A. K. Geim and I. V. Grigorieva, *Nature (London)* **499**, 419 (2013).
- [18] X. Hong, J. Kim, S.-F. Shi, Y. Zhang, C. Jin, Y. Sun, S. Tongay, J. Wu, Y. Zhang, and F. Wang, *Nat. Nanotechnol.* **9**, 682 (2014).
- [19] Y. Yu, S. Hu, L. Su, L. Huang, Y. Liu, Z. Jin, A. A. Puzos, D. B. Geohegan, K. W. Kim, Y. Zhang, and L. Cao, *Nano Lett.* **15**, 486 (2015).
- [20] A. F. Rigosi, H. M. Hill, Y. Li, A. Chernikov, and T. F. Heinz, *Nano Lett.* **15**, 5033 (2015).
- [21] H. Fang, C. Battaglia, C. Carraro, S. Nemsak, B. Ozdol, J. S. Kang, H. A. Bechtel, S. B. Desai, F. Kronast, A. A. Unal, G. Conti, C. Conlon, G. K. Palsson, M. C. Marting, A. M. Minor, C. S. Fadley, E. Yablonovitch, R. Maboudian, and A. Javey, *Proc. Natl. Acad. Sci. USA* **111**, 6198 (2014).
- [22] M.-H. Chiu, M.-Y. Li, W. Zhang, W.-T. Hsu, W.-H. Chang, M. Terrones, H. Terrones, and L.-J. Li, *ACS Nano* **8**, 9649 (2014).
- [23] C.-H. Lee, G.-H. Lee, A. M. van der Zande, W. Chen, Y. Li, M. Han, X. Cui, G. Arefe, C. Nuckolls, T. F. Heinz, J. Guo, J. Hone, and P. Kim, *Nat. Nanotechnol.* **9**, 676 (2014).
- [24] M. M. Furchi, A. Pospischil, F. Libisch, J. Burgdörfer, and T. Mueller, *Nano Lett.* **14**, 4785 (2014).
- [25] R. Cheng, D. Li, H. Zhou, C. Wang, A. Yin, S. Jiang, Y. Liu, Y. Chen, Y. Huang, and X. Duan, *Nano Lett.* **14**, 5590 (2014).
- [26] F. Ceballos, M. Z. Bellus, H.-Y. Chiu, and H. Zhao, *ACS Nano* **8**, 12717 (2014).
- [27] P. Rivera, J. R. Schaibley, A. M. Jones, J. S. Ross, S. Wu, G. Aivazian, P. Klement, K. Seyler, G. Clark, N. J. Ghimire, J. Yan, D. G. Mandrus, W. Yao, and X. Xu, *Nat. Commun.* **6**, 6242 (2015).
- [28] H. Yu, Y. Wang, Q. Tong, X. Xu, and W. Yao, *Phys. Rev. Lett.* **115**, 187002 (2015).
- [29] P. Rivera, K. L. Seyler, H. Yu, J. R. Schaibley, J. Yan, D. G. Mandrus, W. Yao, and X. Xu, *Science* **351**, 688 (2016).
- [30] J. Jung *et al.*, *Nat. Commun.* **6**, 6308 (2015).
- [31] K. Zhou, D. Wickramaratne, S. Ge, S. Su, A. De, and R. K. Lake, *arXiv:1606.03682*.
- [32] K. Liu, L. Zhang, T. Cao, C. Jin, D. Qiu, Q. Zhou, A. Zettl, P. Yang, S. G. Louie, and F. Wang, *Nat. Commun.* **5**, 4966 (2014).
- [33] A. M. van der Zande, J. Kunstmann, A. Chernikov, D. A. Chenet, Y. M. You, X. X. Zhang, P. Y. Huang, T. C. Berkelbach, L. Wang, F. Zhang, M. S. Hybertsen, D. A. Muller, D. R. Reichman, T. F. Heinz, and J. C. Hone, *Nano Lett.* **14**, 3869 (2014).
- [34] W.-T. Hsu, Z.-A. Zhao, L.-J. Li, C.-H. Chen, M.-H. Chiu, P.-S. Chang, Y.-C. Chou, and W.-H. Chang, *ACS Nano* **8**, 2951 (2014).
- [35] H. Heo, J. H. Sung, S. Cha, B.-G. Jang, J.-Y. Kim, G. Jin, D. Lee, J.-H. Ahn, M.-J. Lee, J. H. Shim, H. Choi, and M.-H. Jo, *Nat. Commun.* **6**, 7372 (2015).
- [36] S. Huang, X. Ling, L. Liang, J. Kong, H. Terrones, V. Meunier, and M. S. Dresselhaus, *Nano Lett.* **14**, 5500 (2014).
- [37] A. A. Puzos, L. B. Liang, X. F. Li, K. Xiao, B. G. Sumpter, V. Meunier, and D. B. Geohegan, *ACS Nano* **10**, 2736 (2016).
- [38] C. H. Lui, Z. Ye, C. Ji, K.-C. Chiu, C.-T. Chou, T. I. Andersen, C. Means-Shively, H. Anderson, J.-M. Wu, T. Kidd, Y.-H. Lee, and R. He, *Phys. Rev. B* **91**, 165403 (2015).

- [39] Q. Tong, H. Yu, Q. Zhu, Y. Wang, X. Xu, and W. Yao, *Nat. Phys.* (2016), doi:[10.1038/nphys3968](https://doi.org/10.1038/nphys3968).
- [40] F. Wu, T. Lovorn, and A. H. MacDonald, [arXiv:1610.03855](https://arxiv.org/abs/1610.03855).
- [41] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, *Phys. Rev. Lett.* **99**, 256802 (2007).
- [42] R. Bistritzer and A. H. MacDonald, *Proc. Natl. Acad. Sci. USA* **108**, 12233 (2011).
- [43] S. Shallcross, S. Sharma, E. Kandelaki, and O. A. Pankratov, *Phys. Rev. B* **81**, 165105 (2010).
- [44] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, *Phys. Rev. B* **86**, 155449 (2012).
- [45] P. San-Jose, A. Gutiérrez-Rubio, M. Sturla, and F. Guinea, *Phys. Rev. B* **90**, 115152 (2014).
- [46] C. Zhang, C.-P. Chuu, X. Ren, M.-Y. Li, L.-J. Li, C. Jin, M.-Y. Chou, and C.-K. Shih, *Sci. Adv.* **3**, e1601459 (2017).
- [47] Th. Böker, R. Severin, A. Müller, C. Janowitz, R. Manzke, D. Voß, P. Krüger, A. Mazur, and J. Pollmann, *Phys. Rev. B* **64**, 235305 (2001).
- [48] S. Bhattacharyya and A. K. Singh, *Phys. Rev. B* **86**, 075454 (2012).
- [49] P. Giannozzi *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [50] J. Jung, A. Raoux, Z. Qiao, and A. H. MacDonald, *Phys. Rev. B* **89**, 205414 (2014).
- [51] M.-H. Chiu, C. Zhang, H.-W. Shiu, C.-P. Chuu, C.-H. Chen, C.-Y. S. Chang, C.-H. Chen, M.-Y. Chou, C.-K. Shih, and L.-J. Li, *Nat. Commun.* **6**, 7666 (2015).
- [52] N. R. Wilson, P. V. Nguyen, K. L. Seyler, P. Rivera, A. J. Marsden, Z. P. L. Laker, G. C. Constantinescu, V. Kandyba, A. Barinov, N. D. M. Hine, X. Xu, and D. H. Cobden, *Sci. Adv.* **3**, e1601832 (2017).
- [53] P. San-Jose, A. Gutiérrez-Rubio, M. Sturla, and F. Guinea, *Phys. Rev. B* **90**, 075428 (2014).
- [54] P.-C. Yeh *et al.*, *Nano Lett.* **16**, 953 (2016).
- [55] Z. Nie, R. Long, L. Sun, C.-C. Huang, J. Zhang, Q. Xiong, D. W. Hewak, Z. Shen, O. V. Prezhdo, and Z.-H. Loh, *ACS Nano* **8**, 10931 (2014).
- [56] X. Li, J. T. Mullen, Z. Jin, K. M. Borysenko, M. Buongiorno Nardelli, and K. W. Kim, *Phys. Rev. B* **87**, 115418 (2013).
- [57] Z. Jin, X. Li, J. T. Mullen, and K. W. Kim, *Phys. Rev. B* **90**, 045422 (2014).