

Theory of optical absorption by interlayer excitons in transition metal dichalcogenide heterobilayers

Fengcheng Wu,¹ Timothy Lovorn,² and A. H. MacDonald²

¹*Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*

²*Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA*

 (Received 9 November 2017; revised manuscript received 7 January 2018; published 22 January 2018)

We present a theory of optical absorption by interlayer excitons in a heterobilayer formed from transition metal dichalcogenides. The theory accounts for the presence of small relative rotations that produce a momentum shift between electron and hole bands located in different layers, and a moiré pattern in real space. Because of the momentum shift, the optically active interlayer excitons are located at the moiré Brillouin zone's corners, instead of at its center, and would have elliptical optical selection rules if the individual layers were translationally invariant. We show that the exciton moiré potential energy restores circular optical selection rules by coupling excitons with different center of mass momenta. A variety of interlayer excitons with both senses of circular optical activity, and energies that are tunable by twist angle, are present at each valley. The lowest energy exciton states are generally localized near the exciton potential energy minima. We discuss the possibility of using the moiré pattern to achieve scalable two-dimensional arrays of nearly identical quantum dots.

DOI: [10.1103/PhysRevB.97.035306](https://doi.org/10.1103/PhysRevB.97.035306)

I. INTRODUCTION

Motivated by the discovery of unprecedented optical [1] and electronic [2,3] properties, the monolayer transition metal dichalcogenides (TMD) class of two-dimensional materials is currently under active study. Because of strong Coulomb interactions between electrons and holes, the optical absorption of monolayer TMD semiconductors like MoS₂ is dominated by excitonic features [4–7]. In TMDs excitons have two-fold valley degeneracy because the material's band extrema are located at inequivalent momenta at the triangular lattice Brillouin zone corners that are related by time-reversal symmetry. Valley-dependent optical selection rules enable optical manipulation of the valley degree of freedom [8–11], providing access to many interesting phenomena, including light-induced Hall effects [12], a valley-selective optical Stark effect [13,14], and a large valley-exclusive Bloch-Siegert shift [15].

Van der Waals heterobilayers composed of different monolayer TMDs provide an even richer platform in which new properties can be engineered. WX₂/MoX₂ (X = S, Se) heterobilayers have been experimentally realized [16–20], and have a type-II band alignment in which the conduction- and valence-band edges are associated with MoX₂ and WX₂, respectively. The lowest-energy excitons are therefore spatially indirect since the constituent electrons and holes are located primarily in different layers. The interlayer excitons have a long lifetime and an electrically tunable energy because of their spatially indirect nature [21], but still possess a sizable electron-hole binding energy. These features make TMD bilayers a strong candidate system to realize spatially indirect exciton condensation [22,23].

Heterobilayers of two-dimensional materials exhibit long-period moiré patterns when they have a small lattice constant mismatch and/or relative orientation angle. Moiré patterns can induce dramatic changes in material properties that can

be altered by controlling the twist angle. In graphene-based heterostructures, the moiré pattern strongly modifies the Dirac spectrum [24–27], leads to Hofstadter-butterfly spectra in a strong magnetic field [24,28,29], and can assist optical spectroscopy in the fractional quantum Hall regime [30]. In TMD heterobilayers, the possibilities are richer because of the wider variety of materials, and they are still relatively unexplored. In a previous paper [31], we have shown that a moiré pattern profoundly alters the spectrum of optical absorption due to *intralayer* excitons by producing satellite excitonic peaks, and enables the design of topological bands of *intralayer* excitons. Moiré pattern in TMDs can also be used to engineer electronic topological insulator nanodots and nanostripes [32].

A comprehensive theory of interlayer excitons in moiré pattern is desirable for interpreting ongoing and future experiments. In a pioneering theoretical paper [33] by Yu *et al.*, it was shown that optically active interlayer excitons are located at the corners of the moiré Brillouin zone (MBZ) instead of its center, and that they have elliptical optical selection rules instead of circular. These predictions are, however, based on a theory that omits one important ingredient, namely, the spatial modulation of exciton energy produced by the moiré pattern. In this paper, we present a theory of interlayer excitons in a moiré pattern that does take account of the exciton potential energy, which has a sizable magnitude according to our *ab initio* calculations. We explain why this potential energy plays an essential role in determining exciton properties.

Our theory predicts that the exciton potential energy has two major effects on interlayer excitons. First, the periodic exciton potential inevitably mixes and splits the three independent exciton center-of-mass eigenstates that are otherwise degenerate at the MBZ corners. The new exciton eigenstates respect the rotational symmetries of the moiré pattern and for this reason circular optical selection rules must be restored. Mixing of center of mass eigenstates by the exciton potential energy splits

the spectral weight associated with interlayer excitons across a variety of distinct states whose energy splittings can be tuned by adjusting twist angles. Interlayer excitons also have a valley degree of freedom, but the optical selection rules are not locked to valley. In particular, different interlayer excitons at a given valley can absorb light with different circular polarization. This can have important experimental consequences, as we will explain. Second, the moiré potential generally results in localized exciton states that are confined near the potential minima. This suggests a Van der Waals heterostructure moiré based strategy for realizing two-dimensional arrays of nearly identical quantum dots with nanoscale. When the moiré potential is fine tuned such that it has an enhanced sixfold rotational symmetry, the exciton moiré band structure has Dirac dispersion within the light cone.

Our paper is organized as follows. In Secs. II and III, we formulate the optical matrix element for interlayer excitons. The derivation closely follows that in Ref. [33]. In Sec. IV, the exciton moiré potential energy is constructed. In Sec. V, we present our main results on the optical absorption spectrum for interlayer excitons in moiré pattern. In Sec. VI, we show the presence of Dirac cones in the exciton band structure when the moiré potential is fine tuned. Finally, in Sec. VII, experimental implications of our work are discussed.

II. THEORY OF MOIRÉ HETEROBILAYER INTERLAYER EXCITONS

We study $WX_2/\text{Mo}X_2$ ($X = \text{S}, \text{Se}$) TMD heterobilayers with a small relative twist angle θ . We choose heterobilayers with a common chalcogen (X) atom, because these have small lattice constant mismatches ($\delta \sim 0.1\%$). The moiré periodicity a_M is given by

$$a_M \approx a_0/\sqrt{\theta^2 + \delta^2}, \quad (1)$$

where the lattice constant mismatch δ is defined as $|a_0 - a'_0|/a_0$, and a_0 and a'_0 are the lattice constants of the two layers. At twist angles larger than $\sim 0.1^\circ$, we can neglect δ . To simplify the discussion, we set δ to be zero below. The low-energy conduction-band states of $WX_2/\text{Mo}X_2$ bilayers are localized in $\text{Mo}X_2$, and the low-energy valence-band holes are localized in the WX_2 layers. Mixing between layers is weak because of the Van der Waals character of the heterobilayer, and because of the heterobilayer band alignment [34].

The bilayer has two distinct stacking orders AA and AB, which are illustrated in Figs. 1(a) and 1(b). Both configurations are experimentally relevant [20,35]. Our choice of stacking convention allows us to study interlayer excitons in AA and AB stacking systems in a unified manner.

We consider interlayer excitons composed of electrons (holes) residing in the $\pm\mathbf{K}_T$ ($\pm\mathbf{K}_B$) valleys of a top (T) $\text{Mo}X_2$ [bottom (B) WX_2] layer. The electron-hole exchange interaction, which couples intralayer $+\mathbf{K}$ and $-\mathbf{K}$ valley excitons [36–39], is extremely weak for interlayer excitons because the electrons and holes are spatially separated. We therefore neglect the exchange interaction and study $\pm\mathbf{K}$ valley excitons separately. Since excitons in opposite valleys are time-reversal partners, we focus on the $+\mathbf{K}$ valley excitons unless otherwise stated. When the weak hybridization between layers is neglected, exciton states can be labeled by center-of-mass

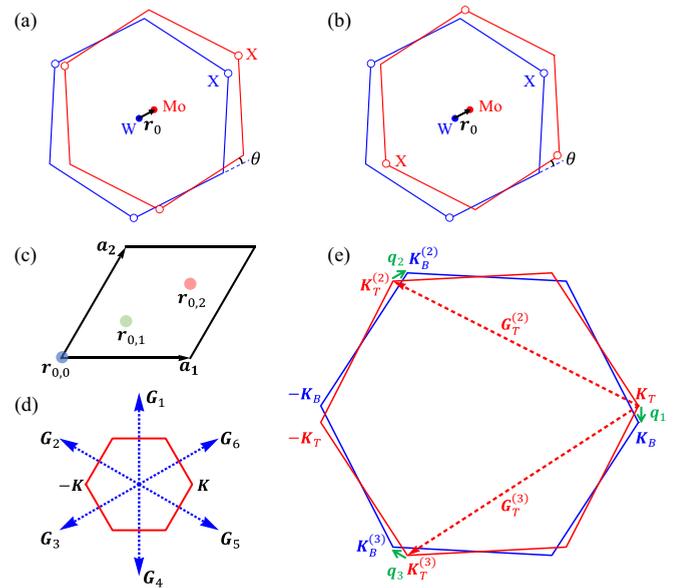


FIG. 1. (a) Illustration of AA stacking with a small twist angle θ and an in-plane displacement \mathbf{r}_0 . The blue and red hexagons mark the Wigner-Seitz cells of the two layers. (b) The corresponding illustration of AB stacking. AA and AB stacking are distinguished by a rotation of the top layer ($\text{Mo}X_2$) by π around the metal (Mo) axis. (c) Unit cell of the bilayer when $\theta = 0$. The vectors $\mathbf{r}_{0,n}$ in this figure are high symmetry points as discussed in the main text. (d) Brillouin zone and the first shell reciprocal lattice vectors for the bilayer with $\theta = 0$. (e) Brillouin zones associated with the top and bottom layers in a twisted bilayer.

momentum \mathbf{Q} . The corresponding exciton wave function is

$$|\mathbf{Q}\rangle = \frac{1}{\sqrt{\mathcal{A}}} \sum_{\mathbf{k}} f_{\mathbf{k}} a_{c(\mathbf{K}_T+\mathbf{k}+\frac{m_e}{M}\mathbf{Q})}^\dagger a_{v(\mathbf{K}_B+\mathbf{k}-\frac{m_h}{M}\mathbf{Q})} |G\rangle, \quad (2)$$

where $|G\rangle$ is the ground state of the heterobilayer, and $a_{c(\mathbf{K}_T+\mathbf{q})}^\dagger$ [$a_{v(\mathbf{K}_B+\mathbf{q})}$] creates an electron (a hole) in the \mathbf{K} -valley conduction (valence) band of the $\text{Mo}X_2$ (WX_2) layer. The meaning of various quantities that enter (2) is as follows: \mathcal{A} is the system area, \mathbf{k} is the electron-hole relative momentum, and \mathbf{Q} is the electron-hole center-of-mass momentum. In making these definitions we have adopted the usual convention of measuring momentum from the band extrema points in each layer. In Eq. (2), m_e and m_h are, respectively, the electron and hole effective masses, $M = m_e + m_h$ is the exciton total mass, and $f_{\mathbf{k}}$ is the electron-hole relative-motion wave function in momentum space. We only study the lowest-energy interlayer excitons, for which $f_{\mathbf{k}}$ has s -wave symmetry, i.e., it depends only on the magnitude of \mathbf{k} . $f_{\mathbf{k}}$ is normalized such that $\langle \mathbf{Q} | \mathbf{Q} \rangle = 1$.

When the bilayer has zero twist angle, the optically active excitons are located at $\mathbf{Q} = 0$ because light carries a vanishingly small momentum. At a finite twist angle, the Brillouin zones associated with the top and bottom layers have a relative rotation, leading to a finite momentum displacement between electron and hole bands [see Fig. 1(e)]. When we adopt the standard and convenient convention of measuring momentum in each layer from its Brillouin-zone corner band extremum, the optically active excitons in the \mathbf{K} valley are shifted to

momenta [40]

$$\mathbf{Q} = (\mathbf{K}_B + \mathbf{G}_B) - (\mathbf{K}_T + \mathbf{G}_T), \quad (3)$$

where \mathbf{G}_B and \mathbf{G}_T are the reciprocal lattice vectors of the bottom and top layers, respectively. By time-reversal symmetry, bright excitons at the $-\mathbf{K}$ valley are located at momenta opposite to those in (3). We prove Eq. (3) below.

The interlayer optical matrix element between the ground state and the exciton state is

$$\mathbf{J}(\theta, \mathbf{r}_0, \mathbf{Q}) = \frac{1}{\sqrt{\mathcal{A}}} \langle G | \hat{\mathbf{j}} | \mathbf{Q} \rangle, \quad (4)$$

where $\hat{\mathbf{j}}$ is the current operator. We evaluate \mathbf{J} using a local approximation valid for long moiré periods. Because the two layers are weakly coupled and have separate lattice translational symmetry, the conduction-band electron and valence-band hole creation operators at a given momentum can be expanded in terms of their Wannier function counterparts:

$$\begin{aligned} a_{v(\mathbf{K}_B + \mathbf{q})} &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{-i(\mathbf{K}_B + \mathbf{q}) \cdot \mathbf{R}} a_{v\mathbf{R}}, \\ a_{c(\mathbf{K}_T + \mathbf{q})}^\dagger &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}'} e^{i(\mathbf{K}_T + \mathbf{q}') \cdot (\mathbf{R}' + \mathbf{r}_0)} a_{c(\mathbf{R}' + \mathbf{r}_0)}^\dagger, \end{aligned} \quad (5)$$

where N is the number of unit cells in each layer, c and v are the Wannier labels identified with the conduction and valence bands, and \mathbf{R} is a triangular lattice Bravais vector. We assume a two-center approximation [33,41] for the optical matrix element and for its dependence on a rigid displacement by \mathbf{r}_0 of one layer with respect to the other. That is to say we assume that

$$\begin{aligned} \langle G | \hat{\mathbf{j}} a_{c(\mathbf{R}' + \mathbf{r}_0)}^\dagger a_{v\mathbf{R}} | G \rangle &\approx \mathbf{j}(\mathbf{R}' + \mathbf{r}_0 - \mathbf{R}) \\ &= \frac{1}{\mathcal{A}} \sum_{\mathbf{p}} e^{-i\mathbf{p} \cdot (\mathbf{R}' + \mathbf{r}_0 - \mathbf{R})} \mathbf{j}(\mathbf{p}), \end{aligned} \quad (6)$$

where $\mathbf{j}(\mathbf{r})$ is a function of displacement. In replacing the first form of Eq. (6) by the second, we have assumed that the matrix element depends only on the difference between electron and hole positions. The third form of Eq. (6) simply expresses $\mathbf{j}(\mathbf{r})$ in terms of its Fourier transform. The two-center approximation preserves all symmetries of the system, and results derived based on this approximation are compatible with all symmetry requirements. Using (2), (5), and (6), we can rewrite \mathbf{J} in the form

$$\begin{aligned} \mathbf{J}(\theta, \mathbf{r}_0, \mathbf{Q}) &= \frac{N}{\mathcal{A}^2} \sum_{\mathbf{k}} \sum_{\mathbf{G}_B, \mathbf{G}_T} f_{\mathbf{k}} \\ &\quad \times \mathbf{j}(\mathbf{K}_B + \mathbf{G}_B + \mathbf{k} - m_h \mathbf{Q} / M) \\ &\quad \times \delta_{\mathbf{Q}, (\mathbf{K}_B + \mathbf{G}_B) - (\mathbf{K}_T + \mathbf{G}_T)} e^{-i\mathbf{G}_T \cdot \mathbf{r}_0}. \end{aligned} \quad (7)$$

Equation (7) establishes the momentum-conservation law in (3). Because of the relatively large separation between layers in the heterobilayer system, we expect that $\mathbf{j}(\mathbf{r})$ should be a smooth function of \mathbf{r} on the scale of the monolayer lattice constant, and hence that $\mathbf{j}(\mathbf{p})$ should decline rapidly on the scale of the monolayer reciprocal lattice vector.

It follows from Eq. (7) that the smallest excitation momenta at which bright excitons appear are $\mathbf{Q} = \mathbf{q}_1, \mathbf{q}_2$, and \mathbf{q}_3 ,

where $\mathbf{q}_1 = \mathbf{K}_B - \mathbf{K}_T$, and \mathbf{q}_2 and \mathbf{q}_3 are the threefold-rotation counterparts of \mathbf{q}_1 as illustrated in Fig. 1(e). The contributions to Eq. (7) from larger momenta allowed by Eq. (3) will be much smaller because $\mathbf{j}(\mathbf{p})$ is a decreasing function of $|\mathbf{p}|$. Below, we retain only the dominant optical response from the \mathbf{q}_n contributions. This yields

$$\mathbf{J}(\theta, \mathbf{r}_0, \mathbf{Q}) \approx \sum_{n=1}^3 \delta_{\mathbf{Q}, \mathbf{q}_n} \mathbf{J}_n e^{-i\mathbf{G}_T^{(n)} \cdot \mathbf{r}_0}, \quad (8)$$

where $\mathbf{G}_T^{(1)}$ is a zero vector, and $\mathbf{G}_T^{(2,3)}$ are reciprocal lattice vectors that connect equivalent \mathbf{K} points in the Brillouin zone of top layer [Fig. 1(e)]. The parameters \mathbf{J}_n are independent of \mathbf{r}_0 .

We extract numerical values for \mathbf{J}_n by comparing with microscopic *ab initio* calculations performed for the commensurate structures obtained by setting θ and \mathbf{Q} to 0. In this limit, Eq. (8) reduces to

$$\mathbf{J}(\mathbf{r}_0) \approx \sum_{n=1}^3 \mathbf{J}_n e^{-i\mathbf{G}_T^{(n)} \cdot \mathbf{r}_0}. \quad (9)$$

In Eq. (9), $\mathbf{G}_T^{(n)}$ is $\lim_{\theta \rightarrow 0} \mathbf{G}_T^{(n)}$. Because the heterobilayer structure is commensurate in this limit, the *ab initio* calculations are readily performed:

$$\begin{aligned} \mathbf{J}(\mathbf{r}_0) &= \frac{1}{\mathcal{A}} \sum_{\mathbf{k}} f_{\mathbf{k}} \langle G | \hat{\mathbf{j}} a_{c(\mathbf{K} + \mathbf{k})}^\dagger a_{v(\mathbf{K} + \mathbf{k})} | G \rangle \\ &\approx \tilde{f}(0) \mathbf{D}(\mathbf{r}_0), \end{aligned} \quad (10)$$

where $\tilde{f}(0) = \sum_{\mathbf{k}} f_{\mathbf{k}} / \mathcal{A}$, and $\mathbf{D}(\mathbf{r}_0) = \langle v\mathbf{K} | \mathcal{J} | c\mathbf{K} \rangle$. $|\tilde{f}(0)|^2$ is the probability that the electron and hole spatially overlap, and \mathcal{J} is the matrix representation of the current operator $\hat{\mathbf{j}}$. In Eq. (10), the final expression is justified by the fact that $f_{\mathbf{k}}$ is peaked around $\mathbf{k} = 0$. For simplicity, we assume $\tilde{f}(0)$ is independent of \mathbf{r}_0 , since the most significant dependence on \mathbf{r}_0 is from $\mathbf{D}(\mathbf{r}_0)$ as we discuss below. Not only the exciton energy but also the strength of optical coupling to the exciton state varies with position within the moiré pattern.

III. INTERLAYER OPTICAL MATRIX ELEMENTS

We performed *ab initio* calculations for aligned layers with identical lattice constants as a function of \mathbf{r}_0 , which is the in-plane displacement shown in Fig. 1. The *ab initio* calculations are performed using fully relativistic density functional theory (DFT) under the local-density approximation as implemented in QUANTUM ESPRESSO [42]; the details of these calculations are the same as those described in Supplemental Material of Ref. [31], except that we use a denser $18 \times 18 \times 1$ k -point sampling, which is required to reduce noise in the individual conduction and valence-band extrema in the AB stacking case. For each value of \mathbf{r}_0 , we fit the bands to a tight-binding model using WANNIER90 [43], and then use the tight-binding model Hamiltonian to evaluate the matrix element of the current operator between the semiconductor ground state and the single exciton state using Eq. (10).

The bilayer has a \hat{C}_3 symmetry when \mathbf{r}_0 has one of three high symmetry displacements: $\mathbf{r}_{0,n} = n(\mathbf{a}_1 + \mathbf{a}_2)/3$ for $n = 0, 1$ and 2 . Here $\mathbf{a}_{1,2}$ are the lattice vectors shown in Fig. 1(c).

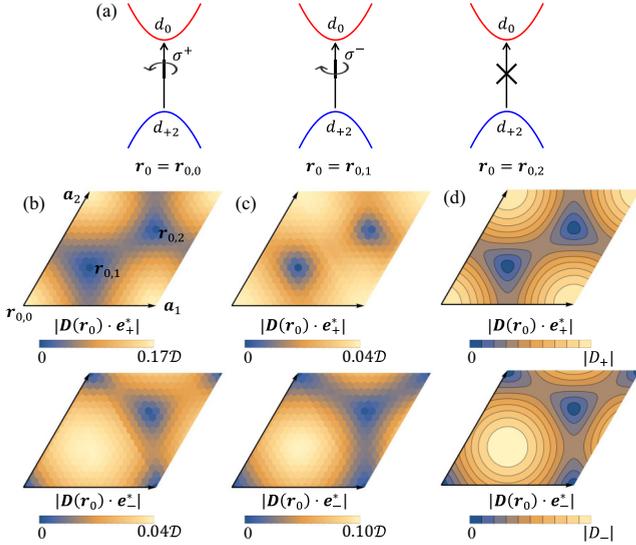


FIG. 2. (a) \mathbf{K} -valley interlayer exciton optical matrix elements for heterobilayers with zero twist angle at special high symmetry displacements. (b) In-plane displacement dependence of the *ab initio* values of $|\mathbf{D}(\mathbf{r}_0) \cdot \mathbf{e}_+^*|$ (top) and $|\mathbf{D}(\mathbf{r}_0) \cdot \mathbf{e}_-^*|$ (bottom) for aligned AA stacking of WS_2/MoS_2 with zero twist angle. Values are given in units of the corresponding intralayer optical response parameter \mathcal{D} . $|\mathbf{D}(\mathbf{r}_0) \cdot \mathbf{e}_-^*|$ at $\mathbf{r}_{0,1}$ is small but nonzero. (c) Same plot as (b) but for AB stacking. (d) Leading harmonic approximation for $|\mathbf{D}(\mathbf{r}_0) \cdot \mathbf{e}_\pm^*|$ explained in the text.

At $\mathbf{r}_{0,0} = 0$, the metal atoms in the two layers are vertically aligned. At $\mathbf{r}_{0,1}$ ($\mathbf{r}_{0,2}$), Mo atoms in the top MoX_2 layer are above the chalcogen (empty) sites of the bottom WX_2 layer. The \hat{C}_3 symmetry puts strong constraints on $\mathbf{D}(\mathbf{r}_{0,n})$. For both AA and AB stackings, the main character of the valence (conduction) band at the \mathbf{K} valley is the W (Mo) d orbital with magnetic quantum number $+2$ (0). The optical selection rule is determined not only by the atomic magnetic quantum numbers, but also by the change in the Bloch phase factor under \hat{C}_3 operations. The Bloch phase shift is different for different $\mathbf{r}_{0,n}$, leading to different optical selection rules. Based on these considerations, we find that the transition at \mathbf{K} satisfies the following circular polarization selection rules:

$$\mathbf{D}(\mathbf{r}_{0,0}) = D_+ \mathbf{e}_+, \mathbf{D}(\mathbf{r}_{0,1}) = D_- \mathbf{e}_-, \mathbf{D}(\mathbf{r}_{0,2}) = 0, \quad (11)$$

where $\mathbf{e}_\pm = (1, \pm i)/\sqrt{2}$. Equation (11) is valid for both AA and AB stacking. Note that the valley-dependent optical selection rule is opposite for $\mathbf{r} = \mathbf{r}_{0,0}$ and $\mathbf{r} = \mathbf{r}_{0,1}$, and that the in-plane transition dipole moment is zero for $\mathbf{r}_0 = \mathbf{r}_{0,2}$. Figure 2(a) summarizes these optical selection rules schematically.

Using Eqs. (9)–(11), we can express \mathbf{J}_n in the form

$$\mathbf{J}_n = \frac{1}{3} [J_+ \mathbf{e}_+ + e^{iG^{(n)} \cdot \mathbf{r}_{0,1}} J_- \mathbf{e}_-], \quad (12)$$

where $J_\pm = \tilde{f}(0) D_\pm$. Because \mathbf{J}_n has both \mathbf{e}_+ and \mathbf{e}_- components, it follows that the optically active exciton states at excitation wave vector \mathbf{q}_n have elliptical optical selection rules, instead of circular [33]. However, as we show below, the exciton potential energy produced by the moiré pattern will mix the three states located at \mathbf{q}_n ($n = 1, 2, 3$). The resulting

eigenstates are coherent superposition of the $|\mathbf{q}_n\rangle$ states for which circular optical selection rules are restored. Substituting Eq. (12) back into Eq. (9), we can parametrize $\mathbf{D}(\mathbf{r}_0)$ in terms of D_\pm :

$$\mathbf{D}(\mathbf{r}_0) \approx \left[\frac{D_+}{3} \sum_{n=1}^3 e^{-iG^{(n)} \cdot \mathbf{r}_0} \right] \mathbf{e}_+ + \left[\frac{D_-}{3} \sum_{n=1}^3 e^{-iG^{(n)} \cdot (\mathbf{r}_0 - \mathbf{r}_{0,1})} \right] \mathbf{e}_-. \quad (13)$$

Equation (13) can be tested numerically. The DFT values of $|\mathbf{D}(\mathbf{r}_0) \cdot \mathbf{e}_\pm^*|$ are presented in Figs. 2(b) and 2(c), respectively for AA and AB stacking of the WS_2/MoS_2 bilayer. In Fig. 2(d), we also plot $|\mathbf{D}(\mathbf{r}_0) \cdot \mathbf{e}_\pm^*|$, as calculated from Eq. (13). The good agreement between analytical and numerical results provides strong justification for Eq. (13). From the *ab initio* results we estimate that $(|D_+|, |D_-|)$ is $(0.17, 0.04)\mathcal{D}$ for AA stacking and $(0.04, 0.10)\mathcal{D}$ for AB stacking, where \mathcal{D} is the matrix element for MoS_2 intralayer excitations in AA stacking with zero twist angle and zero displacement. \mathcal{D} is defined as $|\langle v\mathbf{K} | \mathcal{J} | c\mathbf{K} \rangle|$, where both the conduction and valence states are located in MoS_2 layer. The numerical value of $\hbar\mathcal{D}/e$ is about 5.22 eV \AA , where \hbar is the reduced Planck constant and e is the electron charge. Numerical results of $\text{WSe}_2/\text{MoSe}_2$ bilayer are presented in Ref. [44].

IV. INTERLAYER EXCITON POTENTIAL ENERGY

The interlayer twist generates not only a relative shift in momentum space [Fig. 1(e)], but also a moiré pattern in real space [Fig. 3(a)]. The moiré pattern leads to spatial modulation of the exciton energy. The *potential* energy for the interlayer excitons is $E_g - E_b$, where E_g is the band gap between the WX_2 valence band and the MoX_2 conduction band at the \mathbf{K} point, and E_b is the exciton binding energy. We neglect the variation of E_b in the moiré pattern, since it is typically smaller than the variation of E_g . We use a local approximation [31,45,46] to estimate the exciton moiré potential. First, we

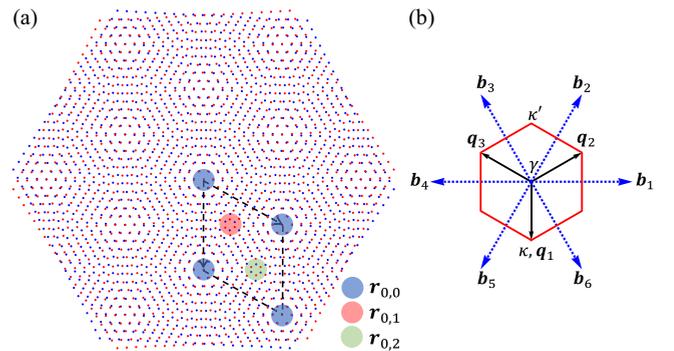


FIG. 3. (a) The small red and blue dots represent Bravais lattices of the two layers with a relative rotation. The dashed lines mark the moiré unit cell. The big dots indicate regions where the local displacement between the two layers is $\mathbf{r}_{0,n}$. Each $\mathbf{r}_{0,1}$ region has three $\mathbf{r}_{0,2}$ neighbors and each $\mathbf{r}_{0,2}$ region has three $\mathbf{r}_{0,1}$ neighbors. (b) Moiré Brillouin zone and the first-shell moiré reciprocal lattice vectors.

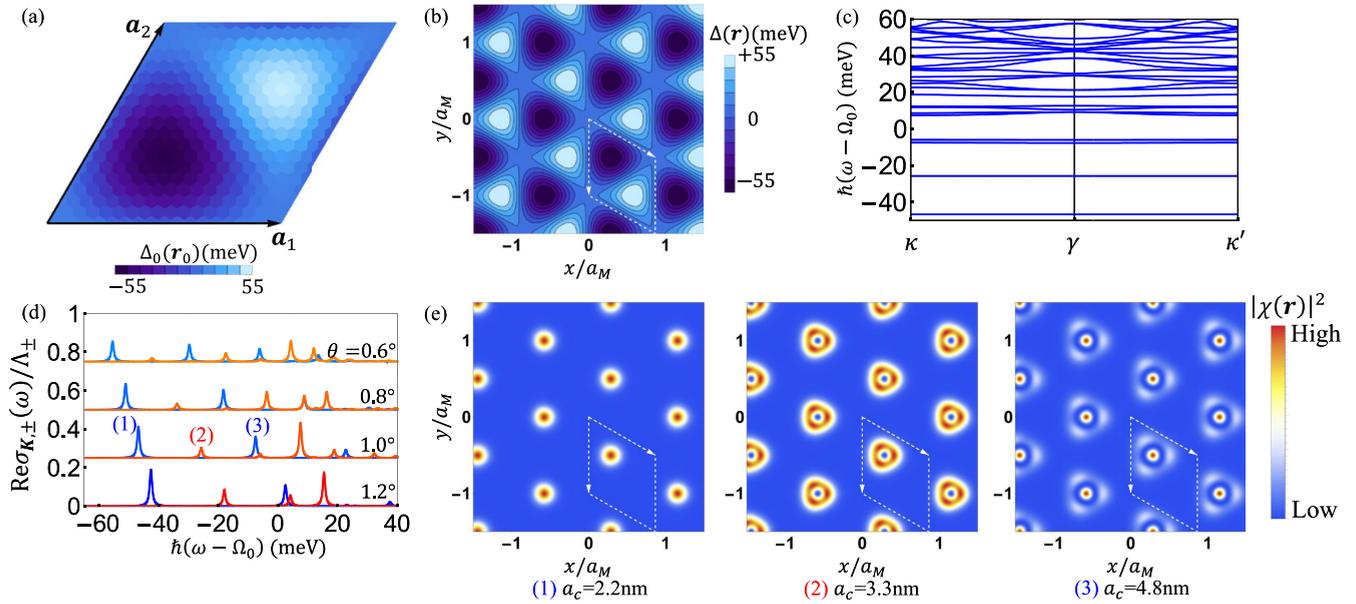


FIG. 4. Results for a WS₂/MoS₂ heterobilayer with AA stacking. (a) Variation of the interlayer band gap as a function of \mathbf{r}_0 when the twist angle is zero. (b) Variation of the band gap in real space when the twist angle is finite. (c) Exciton moiré band structure for $\theta = 1^\circ$. (d) Optical spectrum of \mathbf{K} -valley interlayer excitons for a series of twist angles. The red and blue curves respectively show $\text{Re}\sigma_{\mathbf{K},+}/\Lambda_+$ and $\text{Re}\sigma_{\mathbf{K},-}/\Lambda_-$. Λ_+/Λ_- is about 18. The broadening parameter η is taken to be 0.5 meV. Curves for different θ are shifted vertically. (e) The real-space probability function $|\chi(\mathbf{r})|^2$ of the interlayer excitons responsible for the first three absorption peaks at $\theta = 1^\circ$.

obtain E_g as a function of relative displacement \mathbf{r}_0 at zero twist angle from the DFT band structure. The spatial variation of E_g is shown in Figs. 4(a) and 5(a). Note that the exciton potential energy has its minimum at the $\mathbf{r}_{0,1}$ point for AA stacking and at the $\mathbf{r}_{0,0}$ point for AB stacking. Since the potential is a smooth periodic function of \mathbf{r}_0 in both cases, we can approximate it by the lowest-order harmonic expansion:

$$\Delta_0(\mathbf{r}_0) \equiv E_g(\mathbf{r}_0) - \langle E_g \rangle \approx \sum_{j=1}^6 V_j \exp(i\mathbf{G}_j \cdot \mathbf{r}_0), \quad (14)$$

where $\langle E_g \rangle$ is the average of E_g over \mathbf{r}_0 , and the \mathbf{G}_j are the first-shell reciprocal lattice vectors shown in Fig. 3(c). Because each layer *separately* has \tilde{C}_3 symmetry, Δ_0 is invariant under 120° rotations of \mathbf{r}_0 . This symmetry implies that

$$V_1 = V_3 = V_5, V_2 = V_4 = V_6. \quad (15)$$

Because Δ_0 is always real, $V_1 = V_4^*$ is also required. As a result, all six V_j are determined by $V_1 = V \exp(i\psi)$. We find that for MoS₂/WS₂ heterobilayers (V, ψ) is (12.4 meV, 81.5°) for AA stacking and (1.8 meV, 154.5°) for AB stacking. The variation of the band gap predicted by DFT has been found to agree reasonably well with scanning tunneling microscopy measurement [47].

When the two layers have a relative twist angle θ , the lattice sites associated with top and bottom layers are related by $\mathbf{R}_T = \mathcal{R}(\theta)\mathbf{R}_B + \mathbf{r}_0$, where $\mathcal{R}(\theta)$ is the rotation matrix. The local displacement of the top layer relative to the bottom layer near position \mathbf{R}_T is given by

$$\begin{aligned} \mathbf{d}(\mathbf{R}_T) &= \hat{T}_B(\mathbf{R}_T) = \hat{T}_B(\mathbf{R}_T - \mathbf{R}_B) \\ &\approx \hat{T}_B[\theta \hat{z} \times (\mathbf{R}_T - \mathbf{r}_0) + \mathbf{r}_0], \end{aligned} \quad (16)$$

where the operator \hat{T}_B reduces a vector to the Wigner-Seitz cell of the \mathbf{R}_B lattice. When θ is small, the displacement varies smoothly in real space and gives rise to the moiré pattern. We assume that the variation of the band gap locally follows the variation of the displacement:

$$\begin{aligned} \Delta(\mathbf{r}) &= \Delta_0[\mathbf{d}(\mathbf{r})] \approx \sum_{j=1}^6 V_j \exp[i\mathbf{G}_{B,j} \cdot \mathbf{d}(\mathbf{r})] \\ &\approx \sum_{j=1}^6 V_j \exp[i\mathbf{b}_j \cdot (\mathbf{r} + \mathbf{r}_0 \times \hat{z}/\theta - \mathbf{r}_0/2)], \end{aligned} \quad (17)$$

where $\mathbf{G}_{B,j}$ is the first-shell reciprocal lattice vectors associated with the bottom layer. $\mathbf{b}_j = \theta \mathbf{G}_j \times \hat{z}$ is the reciprocal lattice vector of the moiré pattern, illustrated in Fig. 3(b). The moiré periodicity is inversely proportional to the twist angle: $a_M \approx a_0/\theta$. The last expression in (17) implies that the global displacement \mathbf{r}_0 just leads to a spatial translation of the $\Delta(\mathbf{r})$ potential when θ is finite. Without loss of generality, we take \mathbf{r}_0 to be zero for the twisted bilayer in the following.

Including both kinetic and potential energy contributions the exciton Hamiltonian is

$$H = \hbar\Omega_0 + \frac{\hbar^2 \mathbf{Q}^2}{2M} + \Delta(\mathbf{r}), \quad (18)$$

where $\hbar\Omega_0$ is a constant, $\hbar^2 \mathbf{Q}^2/(2M)$ is the exciton center-of-mass kinetic energy. The periodic potential $\Delta(\mathbf{r})$ is plotted in Figs. 4(b) and 5(b). We have numerically diagonalized this Hamiltonian using a plane-wave expansion. The exciton band structure in the MBZ is shown in Figs. 4(c) and 5(c), respectively, for AA and AB stackings with a 1° twist angle. The moiré potential mixes excitons with different momenta,

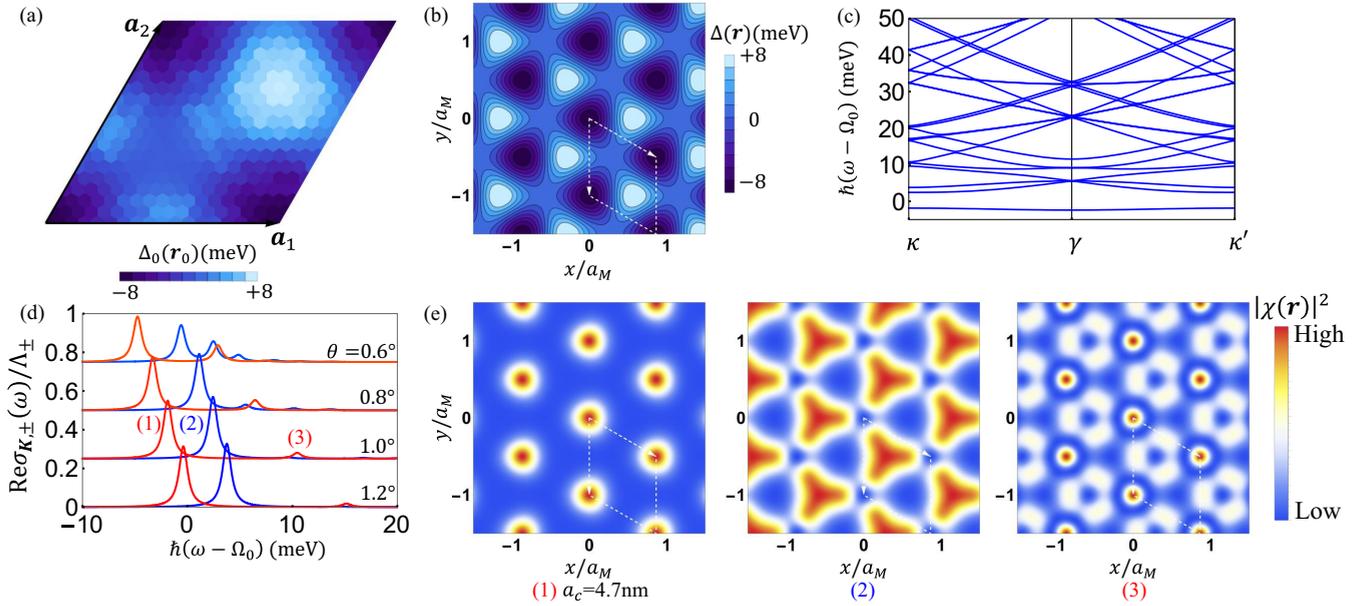


FIG. 5. Same plots as those in Fig. 4 but for the WS₂/MoS₂ heterobilayer in AB stacking. In (d), Λ_+/Λ_- is about 6.

leading to a dramatic change in the interlayer absorption spectrum as we explain in the next section.

V. INTERLAYER EXCITON OPTICAL ABSORPTION SPECTRUM

To study the effects of the moiré potential on the optical response, it is instructive to first consider the weak potential limit and use perturbation theory. As shown in Eq. (8), there are three optically active exciton states, which are at momentum \mathbf{q}_n , i.e., at the corners of the MBZ [Fig 3(b)]. These three states have the same kinetic energy, and would be energetically degenerate if there were no moiré potential energy. However, the three states are coupled by the moiré potential because their momenta differ by moiré reciprocal lattice vectors. Therefore the potential lifts the degeneracy. In degenerate perturbation theory, the potential projected to the $|\mathbf{q}_n\rangle$ has the matrix form

$$\mathcal{V} = \begin{pmatrix} 0 & V_5 & V_6 \\ V_2 & 0 & V_1 \\ V_3 & V_4 & 0 \end{pmatrix}. \quad (19)$$

The eigenstates of \mathcal{V} have the form

$$|\Phi_\lambda\rangle = \frac{1}{\sqrt{3}}(|\mathbf{q}_1\rangle + e^{i2\lambda\pi/3}|\mathbf{q}_2\rangle + e^{-i2\lambda\pi/3}|\mathbf{q}_3\rangle), \quad (20)$$

where λ takes the values 0 and ± 1 . The corresponding eigenvalue is $2V \cos(\psi + 2\lambda\pi/3)$. As a result, different $|\Phi_\lambda\rangle$ generally have different energies unless ψ is fine tuned to multiples of $\pi/3$. The optical matrix element for $|\Phi_\lambda\rangle$ is given by

$$\frac{1}{\sqrt{\mathcal{A}}}\langle G|\hat{\mathbf{j}}|\Phi_\lambda\rangle = \frac{1}{\sqrt{3}}(\mathbf{J}_1 + e^{i2\lambda\pi/3}\mathbf{J}_2 + e^{-i2\lambda\pi/3}\mathbf{J}_3), \quad (21)$$

which is $\mathbf{J}_+ \mathbf{e}_+/\sqrt{3}$ for $\lambda = 0$, $\mathbf{J}_- \mathbf{e}_-/\sqrt{3}$ for $\lambda = +1$, and vanishes for $\lambda = -1$. Since we have set the global displacement \mathbf{r}_0 to be zero, the \mathbf{J}_n in (21) are the optical matrix elements for

$|\mathbf{q}_n\rangle$. We see from Eq. (21) that circular optical selection rules are restored by scattering off the moiré potential. In particular, the exciton states $|\Phi_0\rangle$ and $|\Phi_{+1}\rangle$ couple, respectively, to σ_+ and σ_- circularly polarized light, while $|\Phi_{-1}\rangle$ is optically dark. When ψ is fine tuned to $2\pi/3$, $|\Phi_0\rangle$ and $|\Phi_{+1}\rangle$ becomes energetically degenerate, which gives rise to Dirac cone in the exciton band structure near κ point. We will discuss this special situation in the next section.

The above analysis assumes that the potential is weak. We now argue that circular optical selection rules are generally expected for the twisted bilayer. When $\mathbf{r}_0 = 0$, the bilayer with any twist angle θ has a \hat{C}_3 symmetry. Exciton states can be classified by the symmetry: $\hat{C}_3|\chi\rangle = \exp[i2\pi l_\chi/3]|\chi\rangle$, where integer l_χ is 0 or ± 1 . Due to the \hat{C}_3 symmetry, the excitation from ground state to exciton $|\chi\rangle$ is coupled to σ_\pm circularly polarized light if $l_\chi = \pm 1$, and optically forbidden under normal incident light if $l_\chi = 0$. In Appendix B, we explicitly show that the finite \mathbf{r}_0 and zero \mathbf{r}_0 problems are related by a unitary transformation.

Our *ab initio* results indicate that the amplitude of the moiré potential can be as large as 110 meV, which is not weak at all. Due to the moiré potential, the $|\mathbf{q}_n\rangle$ are redistributed to different exciton states $|\chi_\alpha\rangle$ at the κ point. We assume that only the $|\mathbf{q}_n\rangle$ components in $|\chi_\alpha\rangle$ contribute to the optical response. The optical conductivity can then be expressed as follows:

$$\begin{aligned} \text{Re}\sigma_{\kappa,\pm}(\omega) &= \frac{1}{\omega\mathcal{A}} \sum_{\alpha} |\langle \chi_\alpha | \mathbf{j} \cdot \mathbf{e}_\pm | G \rangle|^2 \Gamma_1(\omega - \omega_\alpha) \\ &\approx \frac{1}{\omega} \sum_{\alpha} \left| \sum_{n=1}^3 \langle \mathbf{q}_n | \chi_\alpha \rangle \mathbf{J}_n \cdot \mathbf{e}_\pm^* \right|^2 \Gamma_1(\omega - \omega_\alpha) \\ &\approx \Lambda_\pm \sum_{\alpha} \left| \frac{1}{3} \sum_{n=1}^3 e^{i\phi_\pm^{(n)}} \langle \mathbf{q}_n | \chi_\alpha \rangle \right|^2 \Gamma_2(\omega - \omega_\alpha), \quad (22) \end{aligned}$$

where $\Gamma_m(\omega - \omega_\alpha) = \eta^m / [\hbar^2(\omega - \omega_\alpha)^2 + \eta^2]$, $\hbar\omega_\alpha$ is the energy of state $|\chi_\alpha\rangle$, and η is a broadening parameter. In (22), $\Lambda_\pm = |J_\pm|^2 / (\eta\Omega_0)$, and the phases are, respectively, $\phi_+^{(n)} = 0$ and $\phi_-^{(n)} = \mathbf{G}^{(n)} \cdot \mathbf{r}_{0,1}$. $\text{Re}\sigma_{\mathbf{K},\pm}(\omega)$ measures the optical absorption by \mathbf{K} -valley excitons in response to σ_\pm polarized light at frequency ω . The optical response from $-\mathbf{K}$ valley can be obtained by time-reversal symmetry: $\text{Re}\sigma_{-\mathbf{K},\mp} = \text{Re}\sigma_{\mathbf{K},\pm}$.

Theoretical optical conductivities calculated for small twist angles are shown in Figs. 4(d) and 5(d). The optical spectrum has the following features. (i) both $\text{Re}\sigma_{\mathbf{K},+}$ and $\text{Re}\sigma_{\mathbf{K},-}$ have peaks around energy $\hbar\Omega_0$, but their peaks are located at distinct energies. This is a manifestation of the circular optical selection rule. For example, the \mathbf{K} -valley exciton state in peak (1) of Fig. 4(d) can be excited by σ_- polarized light, but not by σ_+ polarized light. (ii) There are a series of peaks instead of a single peak in $\text{Re}\sigma_{\mathbf{K},+}$ ($\text{Re}\sigma_{\mathbf{K},-}$) near energy $\hbar\Omega_0$. Umklapp scattering off the moiré potential makes finite-momentum excitons optically active. (iii) The optical spectrum has a strong dependence on the twist angle because it tunes the moiré periodicity in real space and the MBZ in momentum space, leading to changes in the exciton moiré band structure and optical response.

To gain a deeper insight, we study the exciton center-of-mass wave function $\chi_\alpha(\mathbf{r})$, which is defined in real space as follows:

$$\chi_\alpha(\mathbf{r}) = \sum_{\mathbf{Q}} \langle \mathbf{Q} | \chi_\alpha \rangle e^{i\mathbf{Q}\cdot\mathbf{r}}. \quad (23)$$

Figure 4(e) [Fig. 5(e)] plots $|\chi_\alpha(\mathbf{r})|^2$ for excitons in the first three peaks of Fig. 4(d) [Fig. 5(d)]. The amplitude of the moiré potential is about 110 meV for AA stacking. In such a strong potential, the exciton wave function is composed of Wannier states that are strongly confined to the potential minimum positions, as shown in Fig. 4(e). The energy dispersion of the first few exciton bands is almost flat [Fig. 4(c)], indicating that tunneling between neighboring Wannier states is negligibly weak. For AB stacking, the moiré potential is weaker, but a similar confinement is found for excitons in peak (1) [see Fig. 5(e)].

To characterize the localization length of low-energy excitons, we define a parameter a_c as

$$a_c = \sqrt{\int d\mathbf{r} |\delta\mathbf{r}|^2 |\chi_\alpha(\mathbf{r})|^2 / \int d\mathbf{r} |\chi_\alpha(\mathbf{r})|^2}, \quad (24)$$

where the spatial integral is restricted to one moiré Wigner-Seitz cell centered at a potential minimum position, and $|\delta\mathbf{r}|$ is the distance relative to the center. For the exciton state in peak (1) of Fig. 4(d), a_c is 2.8 nm at $\theta = 0.6^\circ$, and decreases to 2.3 nm at $\theta = 1.0^\circ$. For comparison, the moiré periodicity a_M is about 30.5 and 18.3 nm, respectively, for $\theta = 0.6^\circ$ and 1.0° . Values of a_c for other exciton states are listed in Figs. 4(e) and 5(e). Here, a_c of the localized exciton states is not significantly larger than the exciton internal radius a_X (a_X is about 1.3 nm as estimated in Appendix A). The relative motion between electron and hole will be affected by the moiré potential. This effect is neglected in our theory. A more accurate approach is to solve the exciton problem for a supercell structure. Such a theory will still lead to localized exciton states due to the moiré

potential, so we expect that the qualitative picture of our theory remains valid.

The localized Wannier states can be understood using a model of particles in a two-dimensional parabolic potential. Near its minima, the moiré potentials can be approximated as parabolic: $\beta(\delta\mathbf{r}/a_M)^2/2$, where β is a constant independent of a_M . A particle with mass M in the parabolic potential has energies $\sqrt{\beta\hbar^2/(Ma_M^2)}(n_x + n_y + 1)$, where $n_{x,y}$ are non-negative integers. The quantized energy levels explain why peaks (1), (2), and (3) in Fig. 4(e) have nearly equal energy spacing. Wave function Ψ_{00} of the lowest-energy level is proportional to $\exp[-\delta\mathbf{r}^2/(2\ell^2)]$, where $\ell = [\hbar^2 a_M^2 / (\beta M)]^{1/4}$. Ψ_{00} has an s -wave symmetry and provides a good approximation for the exciton in peak (1) of both AA and AB stackings. The localization length a_c estimated based on Ψ_{00} is equal to ℓ , which captures the dependence of a_c on the moiré periodicity a_M . The exciton in peak (2) of AA stacking corresponds to a chiral p -wave state of the first excited level in the parabolic potential, while the exciton in peak (3) mimics a zero-angular-momentum state associated with the second excited level [see Fig. 4(e)].

Finally, we reinterpret the optical response in terms of the real-space wave function. The optical matrix element of state $|\chi_\alpha\rangle$ is

$$\begin{aligned} \frac{1}{\sqrt{\mathcal{A}}} \langle G | \hat{J} | \chi_\alpha \rangle &= \sum_{n=1}^3 \mathbf{J}_n \langle \mathbf{q}_n | \chi_\alpha \rangle \\ &= \frac{1}{\mathcal{A}} \int d\mathbf{r} \left(\sum_{n=1}^3 \mathbf{J}_n e^{-i\mathbf{q}_n \cdot \mathbf{r}} \right) \chi_\alpha(\mathbf{r}) \\ &= \frac{1}{\mathcal{A}} \int d\mathbf{r} \tilde{\mathbf{J}}(\mathbf{r}) u_\alpha(\mathbf{r}), \end{aligned} \quad (25)$$

where $\tilde{\mathbf{J}}(\mathbf{r})$ represents $\mathbf{J}_1 + \mathbf{J}_2 \exp(-i\mathbf{b}_2 \cdot \mathbf{r}) + \mathbf{J}_3 \exp(-i\mathbf{b}_3 \cdot \mathbf{r})$, and $u_\alpha(\mathbf{r})$, the periodic part of χ_α , is equal to $\exp(-i\mathbf{q}_1 \cdot \mathbf{r}) \chi_\alpha(\mathbf{r})$. Both the magnitude and polarization of $\tilde{\mathbf{J}}(\mathbf{r})$ change in real space. The variation follows that of the local displacement, because $\tilde{\mathbf{J}}(\mathbf{r})$ is equivalent to $\mathbf{J}[\mathbf{d}(\mathbf{r})]$, where functions $\mathbf{J}(\mathbf{r}_0)$ and $\mathbf{d}(\mathbf{r})$ are respectively defined in (9) and (16). $\tilde{\mathbf{J}}(\mathbf{r})$ generally has both \mathbf{e}_+ and \mathbf{e}_- components, but becomes fully circularly polarized at high-symmetry positions where the local displacement $\mathbf{d}(\mathbf{r})$ is $\mathbf{r}_{0,0}$ or $\mathbf{r}_{0,1}$.

Equation (25) indicates that the optical response is an average of $\tilde{\mathbf{J}}$ over space weighted by the exciton center-of-mass wave function $u_\alpha(\mathbf{r})$. Different forms of $u_\alpha(\mathbf{r})$ give rise to distinct responses. For example, excitons in both peak (1) and (2) [Fig. 4(d)] are localized around the potential minimum positions. However, they couple to light with different circular polarization, because their localized Wannier states have respectively s and p wave symmetries. Equation (25) can be generalized to the case where a heterobilayer has a finite system size.

VI. DIRAC CONES IN EXCITON BAND STRUCTURE

As mentioned above, there is a degeneracy between the two lowest energy states at κ point when ψ is fine tuned to $2\pi/3$. At $\psi = 2\pi/3$, the exciton potential $\Delta(\mathbf{r})$ has an additional two-fold rotational symmetry besides the \hat{C}_3 symmetry,

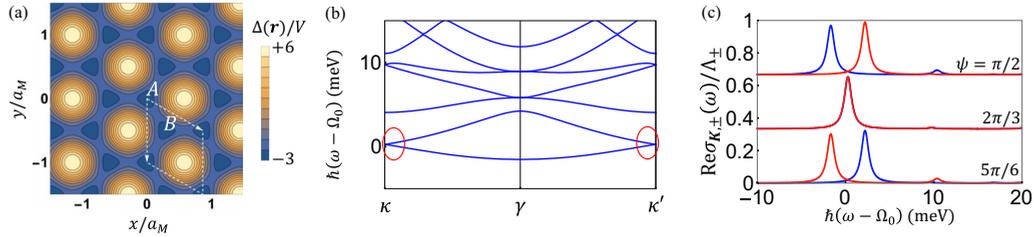


FIG. 6. (a) Exciton potential with $\psi = 2\pi/3$. A and B are potential minimum positions that form a honeycomb lattice. (b) Exciton band structure when $\psi = 2\pi/3$. The red circles highlight the Dirac cones. (c) Optical spectrum for ψ below, at, and above $2\pi/3$. The red and blue curves, respectively, show $\text{Re}\sigma_{K,+}/\Lambda_+$ and $\text{Re}\sigma_{K,-}/\Lambda_-$. Curves for different ψ are shifted vertically. In (b) and (c), the twist angle is 1° . Other parameters take their values in AB stacked WS_2/MoS_2 bilayers.

and creates an effective honeycomb lattice for excitons, as illustrated in Fig. 6(a). It is well known from the study of graphene that a honeycomb lattice harbors Dirac cones in band structure. By explicit calculation using our continuum model, we do find Dirac cones located around κ and κ' points [see Fig. 6(b)].

The two sublattices of the honeycomb lattice are located at positions A and B [Fig. 6(a)]. When $\psi = 2\pi/3$, A and B positions are related by the two-fold rotational symmetry, and thus have the same potential energy $\Delta_A = \Delta_B$. On the other hand, the local optical selection rules are opposite for A and B [see Eq. (25)]. Because of the additional twofold rotational symmetry, $\text{Re}\sigma_{K,+}/\Lambda_+$ and $\text{Re}\sigma_{K,-}/\Lambda_-$ have energetically degenerate response peaks, as shown in Fig. 6(c).

When ψ is away from $2\pi/3$, the twofold rotational symmetry is broken, leading to difference between Δ_A and Δ_B . The staggered potential ($\Delta_{AB} = \Delta_A - \Delta_B$) gaps out the Dirac cone. The sign of Δ_{AB} is controlled by the parameter ψ . For ψ slightly below (above) $2\pi/3$, Δ_{AB} is positive (negative). The optical spectrum can have a strong variation when ψ crosses $2\pi/3$, as indicated in Fig. 6(c).

We emphasize again that optically active excitons of valley \mathbf{K} ($-\mathbf{K}$) are located at κ (κ') point in our convention. Therefore the Dirac cones, if present, are directly relevant to optical responses. The Dirac cones are generally gapped out, which can still lead to interesting effects. For example, a domain wall, which separates two domains with opposite potential difference Δ_{AB} , will host excitonic valley-momentum locked helical states. We leave a detailed study of the helical state to future work.

VII. DISCUSSION

Our theory applies to small twist angle in the range between 0.5° and 2° , where moiré pattern has a strong influence on electronic and optical properties. The twist angle can be controlled to 0.1° accuracy in the case of bilayer graphene using a technique developed in Ref. [29]. Similar controlled technique could be generalized to TMD bilayers. We hope our work can stimulate systematic experimental study of TMD bilayers with small twist angle.

We briefly discuss some experimental implications of our work. One of our important findings is that interlayer excitons have optical selection rules that are not locked to valleys, as illustrated in Figs. 4(d) and 5(d). To demonstrate the

experimental consequences, we consider AA stacking for definiteness. In AA stacking, the *intralayer* A excitons of both MoX_2 and WX_2 layers at the \mathbf{K} valley are coupled to σ_+ circularly polarized light. In a photoluminescence experiment, σ_+ circularly polarized light in resonance with MoS_2 A exciton will excite the intralayer excitons at the \mathbf{K} valley. The A exciton will then relax to form interlayer excitons through hole transfer across the MoS_2/WS_2 bilayers. In the transfer process, the carrier spin and valley indices are expected to be conserved [48]. Therefore interlayer excitons are formed predominantly in valley \mathbf{K} . Reversed circular polarization emission occurs for certain interlayer excitons. For example, the interlayer exciton in peak (1) of Fig. 4(d) will emit σ_- circularly polarized light, which is opposite to the polarization of the excitation light. The physics of the processes that control the steady state probability distribution of interlayer excitons under continuous intralayer excitation is clearly extremely rich, and this will ultimately control the polarization distribution of the outgoing light. Because many of the moiré bands have minima away from the optically active moiré Brillouin-zone corners, the moiré pattern may cause more excitons to accumulate in optically dark states.

The moiré potential can lead to localized exciton states that are confined to the potential minimum positions, as illustrated in Figs. 4(e) and 5(e). This opens the door to scalable engineering of a two-dimensional array of quantum dots using twisted TMD heterobilayers. The effective radius of the quantum dot can be estimated in terms of the parameter a_c , while the interdot separation is given by a_M . Both a_c and a_M are tunable by twist angle. As discussed in Sec. V, a_c and a_M are, respectively, about 3 and 30 nm for AA stacking with $\theta = 0.6^\circ$. The nanoscale quantum dots could be studied using single molecule localization microscopy, which has a nanometer resolution [49]. When fine tuned, the moiré potential gives rise to Dirac excitons, which could be utilized to design one-dimensional excitonic channels. The twisted heterobilayers provide a new platform to study excitons, polaritons, and their condensate in a triangular lattice, where the Bose-Hubbard model physics can be explored [50].

Scanning tunneling microscopy measurement [47] has identified a spatial modulation with an amplitude of 150 meV in the local band gap of a $\text{MoS}_2/\text{WSe}_2$ heterobilayer with rotational alignment and lattice mismatch. Our theory can be generalized to this case, and similar properties of interlayer excitons are expected.

In conclusion, optical absorption by interlayer excitons in TMDs is split into subfeatures with both senses of circular selectivity in a given valley. The overall spectrum and its optical selectivity are both sensitive to spatial variation of the exciton potential energy and the local absorption strength. Photoluminescence and absorption studies of these systems are expected to provide a rich characterization of twisted heterobilayers.

Note added. When finalizing the manuscript, we became aware of a related work by Yu *et al.* [51]. While a different approach was taken in that paper, our results appear to agree where they overlap. We have explicitly worked out the optical absorption spectrum in this manuscript.

ACKNOWLEDGMENTS

F.W. thanks Xiaoqin Li for valuable discussions. The work of F.W. at Argonne was supported by Department of Energy, Office of Basic Energy Science, Materials Science and Engineering Division. Work at Austin was supported by the US Army Research Office under MURI award W911NF-17-1-0312, and by the Welch Foundation under Grant No. TBF1473. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper.

APPENDIX A: ELECTRON-HOLE RELATIVE MOTION WAVE FUNCTION

The electron and hole in one exciton are bound by the attractive Coulomb interaction. In a parabolic band approximation, which is accurate when the binding energy is small compared to the band gap, the electron-hole relative motion wave function $f(\mathbf{k})$ is determined by

$$\sum_{\mathbf{k}'} \left[\frac{\hbar^2 \mathbf{k}^2}{2\mu} \delta_{\mathbf{k}\mathbf{k}'} - \frac{1}{\mathcal{A}} U(\mathbf{k} - \mathbf{k}') \right] f(\mathbf{k}') = -E_b f(\mathbf{k}), \quad (\text{A1})$$

where $\mu = m_e m_h / (m_e + m_h)$ is the reduced mass and E_b is the electron-hole binding energy. In (A1), $U(\mathbf{q}) = [2\pi e^2 / (\epsilon q)] \exp(-qd)$ is the interlayer Coulomb energy, where ϵ is the effective dielectric constant and d is the vertical distance between W layer and Mo layer.

The effective mass m_e and m_h vary as a function of \mathbf{r}_0 . *Ab initio* calculations show that the variation in μ (reduced mass) and M (total mass) is only about $\pm 1\%$, which we will neglect for simplicity. For WS_2/MoS_2 bilayers in both AA and AB stackings, $m_e \approx 0.42m_0$, $m_h \approx 0.34m_0$, and $d \approx 6.15 \text{ \AA}$, where m_0 is the free electron mass. The dielectric constant ϵ has a strong dependence on the environment. Assuming that the heterobilayer is put on a hexagonal boron nitride substrate and exposed to vacuum, ϵ is about $(5 + 1)/2 = 3$. Equation (A1) is solved numerically. We find that the binding energy

E_b is about 300 meV for the lowest energy exciton, of which $f(\mathbf{k})$ depends only on the magnitude of \mathbf{k} . For comparison, the binding energy of the interlayer exciton in $\text{MoS}_2/\text{WSe}_2$ heterobilayer on sapphire substrate was determined to be 260 meV in a previous experimental work [52]. We approximate the exciton internal radius a_X by the root of mean square of the electron-hole separation. a_X can be estimated based on the wave function $f(\mathbf{k})$. We obtain a numerical value of a_X about 1.3 nm. a_X is much smaller compared to the moiré periodicity when the twist angle is smaller than 1° . Therefore it is a good approximation to assume that the energy of the exciton follows the local band gap.

For a $\text{WSe}_2/\text{MoSe}_2$ bilayer, $m_e \approx 0.49m_0$, $m_h \approx 0.35m_0$, and $d \approx 6.47 \text{ \AA}$. Still using $\epsilon \approx 3$, we find E_b is about 300 meV and a_X is about 1.3 nm. The lattice constant a_0 is about 3.19 \AA for WS_2 and 3.32 \AA for WSe_2 .

APPENDIX B: DEPENDENCE OF THE OPTICAL SPECTRUM ON \mathbf{r}_0

The global displacement \mathbf{r}_0 just leads to a spatial translation of the moiré potential. Therefore, the energy spectrum is independent of \mathbf{r}_0 . Here we further show that the optical spectrum is also independent of \mathbf{r}_0 .

The moiré potential has the form

$$\Delta(\mathbf{r}) = \sum_{j=1}^6 V_j \exp[i\mathbf{b}_j \cdot (\mathbf{r} + \mathbf{r}'_0)], \quad (\text{B1})$$

where $\mathbf{r}'_0 = \mathbf{r}_0 \times \hat{z} / \theta - \mathbf{r}_0 / 2$. The wave function at a finite \mathbf{r}_0 is related to that at $\mathbf{r}_0 = 0$ by applying a unitary transformation to the plane-wave state:

$$\begin{aligned} |\chi\rangle &= e^{-i\mathbf{q}_1 \cdot \mathbf{r}'_0} (c_1 e^{i\mathbf{q}_1 \cdot \mathbf{r}'_0} |\mathbf{q}_1\rangle + c_2 e^{i\mathbf{q}_2 \cdot \mathbf{r}'_0} |\mathbf{q}_2\rangle + c_3 e^{i\mathbf{q}_3 \cdot \mathbf{r}'_0} |\mathbf{q}_3\rangle + \dots) \\ &= c_1 |\mathbf{q}_1\rangle + c_2 e^{i\mathbf{b}_2 \cdot \mathbf{r}'_0} |\mathbf{q}_2\rangle + c_3 e^{i\mathbf{b}_3 \cdot \mathbf{r}'_0} |\mathbf{q}_3\rangle + \dots \\ &\approx c_1 |\mathbf{q}_1\rangle + c_2 e^{i\mathbf{G}_T^{(2)} \cdot \mathbf{r}_0} |\mathbf{q}_2\rangle + c_3 e^{i\mathbf{G}_T^{(3)} \cdot \mathbf{r}_0} |\mathbf{q}_3\rangle + \dots, \end{aligned} \quad (\text{B2})$$

where $|\chi\rangle$ is the wave function at κ point, and \dots represents other components in $|\chi\rangle$. In (B2), the approximation $\mathbf{G}_T \approx \mathbf{G} + \theta \hat{z} \times \mathbf{G} / 2$ is used. The phase factors $e^{i\mathbf{G}_T^{(j)} \cdot \mathbf{r}_0}$ capture the dependence on \mathbf{r}_0 , while the coefficient c_j is independent of \mathbf{r}_0 .

The optical matrix element for $|\chi\rangle$ is

$$\begin{aligned} \frac{1}{\sqrt{\mathcal{A}}} \langle G | \hat{\mathbf{j}} | \chi \rangle &= \frac{1}{\sqrt{\mathcal{A}}} \sum_{n=1}^3 \langle \mathbf{q}_n | \chi \rangle \langle G | \hat{\mathbf{j}} | \mathbf{q}_n \rangle \\ &= c_1 \mathbf{J}_1 + [c_2 e^{i\mathbf{G}_T^{(2)} \cdot \mathbf{r}_0}] [\mathbf{J}_2 e^{-i\mathbf{G}_T^{(2)} \cdot \mathbf{r}_0}] \\ &\quad + [c_3 e^{i\mathbf{G}_T^{(3)} \cdot \mathbf{r}_0}] [\mathbf{J}_3 e^{-i\mathbf{G}_T^{(3)} \cdot \mathbf{r}_0}] \\ &= c_1 \mathbf{J}_1 + c_2 \mathbf{J}_2 + c_3 \mathbf{J}_3, \end{aligned} \quad (\text{B3})$$

where we have used Eq. (9). Equation (B3) proves that the optical spectrum is independent of \mathbf{r}_0 for small but finite twist angle θ .

[1] X. Xu, W. Yao, D. Xiao, and T. F. Heinz, *Nat. Phys.* **10**, 343 (2014).

[2] X. Xi, Z. Wang, W. Zhao, J.-H. Park, K. T. Law, H. Berger, L. Forró, J. Shan, and K. F. Mak, *Nat. Phys.* **12**, 139 (2016).

- [3] Z. Fei, T. Palomaki, S. Wu, W. Zhao, X. Cai, B. Sun, P. Nguyen, J. Finney, X. Xu, and D. H. Cobden, *Nat. Phys.* **13**, 677 (2017).
- [4] Z. Ye, T. Cao, K. O'Brien, H. Zhu, X. Yin, Y. Wang, S. G. Louie, and X. Zhang, *Nature (London)* **513**, 214 (2014).
- [5] K. He, N. Kumar, L. Zhao, Z. Wang, K. F. Mak, H. Zhao, and J. Shan, *Phys. Rev. Lett.* **113**, 026803 (2014).
- [6] A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. B. Aslan, D. R. Reichman, M. S. Hybertsen, and T. F. Heinz, *Phys. Rev. Lett.* **113**, 076802 (2014).
- [7] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Phys. Rev. Lett.* **111**, 216805 (2013).
- [8] T. Cao, G. Wang, W. Han, H. Ye, C. Zhu, J. Shi, Q. Niu, P. Tan, E. Wang, B. Liu, and J. Feng, *Nat. Commun.* **3**, 887 (2012).
- [9] H. Zeng, J. Dai, W. Yao, D. Xiao, and X. Cui, *Nat. Nanotechnol.* **7**, 490 (2012).
- [10] K. F. Mak, K. He, J. Shan, and T. F. Heinz, *Nat. Nanotechnol.* **7**, 494 (2012).
- [11] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, *Phys. Rev. Lett.* **108**, 196802 (2012).
- [12] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen, *Science* **344**, 1489 (2014).
- [13] J. Kim, X. Hong, C. Jin, S.-F. Shi, C.-Y. S. Chang, M.-H. Chiu, L.-J. Li, and F. Wang, *Science* **346**, 1205 (2014).
- [14] E. J. Sie, J. W. McIver, Y.-H. Lee, L. Fu, J. Kong, and N. Gedik, *Nat. Mat.* **14**, 290 (2015).
- [15] E. J. Sie, C. H. Lui, Y.-H. Lee, L. Fu, J. Kong, and N. Gedik, *Science* **355**, 1066 (2017).
- [16] X. Hong, J. Kim, S.-F. Shi, Y. Zhang, C. Jin, Y. Sun, S. Tongay, J. Wu, Y. Zhang, and F. Wang, *Nat. Nano* **9**, 682 (2014).
- [17] Y. Gong, J. Lin, X. Wang, G. Shi, S. Lei, Z. Lin, X. Zou, G. Ye, R. Vajtai, B. I. Yakobson, H. Terrones, M. Terrones, B. K. Tay, J. Lou, S. T. Pantelides, Z. Liu, W. Zhou, and P. M. Ajayan, *Nat. Mater.* **13**, 1135 (2014).
- [18] Y. Gong, S. Lei, G. Ye, B. Li, Y. He, K. Keyshar, X. Zhang, Q. Wang, J. Lou, Z. Liu, R. Vajtai, W. Zhou, and P. M. Ajayan, *Nano Lett.* **15**, 6135 (2015).
- [19] 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).
- [20] P. Rivera, K. L. Seyler, H. Yu, J. R. Schaibley, J. Yan, D. G. Mandrus, W. Yao, and X. Xu, *Science* **351**, 688 (2016).
- [21] 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).
- [22] M. Fogler, L. Butov, and K. Novoselov, *Nat. Commun.* **5**, 4555 (2014).
- [23] F.-C. Wu, F. Xue, and A. H. MacDonald, *Phys. Rev. B* **92**, 165121 (2015).
- [24] 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).
- [25] E. Wang, X. Lu, S. Ding, W. Yao, M. Yan, G. Wan, K. Deng, S. Wang, G. Chen, L. Ma, J. Jung, A. V. Fedorov, Y. Zhang, G. Zhang, and S. Zhou, *Nat. Phys.* **12**, 1111 (2016).
- [26] M. Yankowitz, J. Xue, D. Cormode, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, P. Jacquod, and B. J. LeRoy, *Nat. Phys.* **8**, 382 (2012).
- [27] 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).
- [28] 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).
- [29] K. Kim, A. DaSilva, S. Huang, B. Fallahzad, S. Larentis, T. Taniguchi, K. Watanabe, B. J. LeRoy, A. H. MacDonald, and E. Tutuc, *Proc. Natl. Acad. Sci. USA* **114**, 3364 (2017).
- [30] F. Wu and A. H. MacDonald, *Phys. Rev. B* **94**, 241108 (2016).
- [31] F. Wu, T. Lovorn, and A. H. MacDonald, *Phys. Rev. Lett.* **118**, 147401 (2017).
- [32] Q. Tong, H. Yu, Q. Zhu, Y. Wang, X. Xu, and W. Yao, *Nat. Phys.* **13**, 356 (2017).
- [33] H. Yu, Y. Wang, Q. Tong, X. Xu, and W. Yao, *Phys. Rev. Lett.* **115**, 187002 (2015).
- [34] C. Zhang, C. Gong, Y. Nie, K.-A. Min, C. Liang, Y. J. Oh, H. Zhang, W. Wang, S. Hong, L. Colombo, R. M. Wallace, and K. Cho, *2D Mater.* **4**, 015026 (2017).
- [35] N. R. Wilson, P. V. Nguyen, K. 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).
- [36] H. Yu, G.-B. Liu, P. Gong, X. Xu, and W. Yao, *Nat. Commun.* **5**, 3876 (2014).
- [37] M. M. Glazov, T. Amand, X. Marie, D. Lagarde, L. Bouet, and B. Urbaszek, *Phys. Rev. B* **89**, 201302 (2014).
- [38] T. Yu and M. W. Wu, *Phys. Rev. B* **89**, 205303 (2014).
- [39] F. Wu, F. Qu, and A. H. MacDonald, *Phys. Rev. B* **91**, 075310 (2015).
- [40] If we choose $\mathbf{Q} + \mathbf{K}_T - \mathbf{K}_B$ as the center-of-mass momentum, then optically active excitons are located at zero momentum (up to moiré reciprocal lattice vectors).
- [41] R. Bistritzer and A. H. MacDonald, *Proc. Natl. Acad. Sci. USA* **108**, 12233 (2011).
- [42] P. Giannozzi *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [43] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, *Comput. Phys. Commun.* **185**, 2309 (2014).
- [44] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.97.035306> for numerical results of WSe₂/MoSe₂ bilayer.
- [45] J. Jung, A. Raoux, Z. Qiao, and A. H. MacDonald, *Phys. Rev. B* **89**, 205414 (2014).
- [46] M. Wu, X. Qian, and J. Li, *Nano Lett.* **14**, 5350 (2014).
- [47] 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).
- [48] J. R. Schaibley, P. Rivera, H. Yu, K. L. Seyler, J. Yan, D. G. Mandrus, T. Taniguchi, K. Watanabe, W. Yao, and X. Xu, *Nat. Commun.* **7**, 13747 (2016).
- [49] J. Feng, H. Deschout, S. Caneva, S. Hofmann, I. Lončarić, P. Lazić, and A. Radenovic, [arXiv:1706.06313](https://arxiv.org/abs/1706.06313) (2017).
- [50] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, *Phys. Rev. B* **40**, 546 (1989).
- [51] H. Yu, G.-B. Liu, J. Tang, X. Xu, and W. Yao, *Sci. Adv.* **3**, e1701696 (2017).
- [52] 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).