

Supporting Information: Designable exciton mixing through layer alignment in WS₂-graphene heterostructures

Amir Kleiner,¹ Daniel Hernangómez-Pérez,¹ and Sivan Refaely-Abramson¹

¹*Department of Molecular Chemistry and Materials Science,
Weizmann Institute of Science, Rehovot 7610001, Israel*

Keywords: transition-metal dichalcogenides, twist, graphene, excitons, heterostructures, GW, BSE

CONTENTS

I. Supplementary methods - Computational Details	2
A. Geometries	2
B. Density functional theory	2
C. GW calculations	2
D. Bethe-Salpeter equation	3
E. Observables	3
II. Supplementary methods - Convergence tests	4
A. GW convergence tests	4
B. BSE convergence tests	4
III. Supplementary Discussion - Folding and symmetry effects at the mini-BZ	5
A. Atomic reconstruction and BZ \mathbf{k} -point assignment	5
B. Effect of the chalcogen-chalcogen distance on the level alignment	6
IV. Supplementary methods - Assignment of layer contributions	7
A. Projected density of states and layer decomposition	7
B. Layer contributions to the excitons	8
C. Indirect transitions	10
V. Supplementary discussion - Exciton integrated contribution in an optical energy range	10
VI. Supplementary discussion - Analysis of the exciton nature	12
VII. Supplementary methods - Peak broadenings	13
VIII. Supplementary references	14

I. SUPPLEMENTARY METHODS - COMPUTATIONAL DETAILS

A. Geometries

Geometry of the 0°-twist heterobilayer. The 0°-twist heterobilayer is composed of 4×4 TMD (WS_2) unit cells and 5×5 graphene unit cells in a rhombohedral structure with supercell in-plane lattice parameters of length $a = 12.6\text{\AA}$ (see Fig. 1 (a) in the main text). The heterobilayer is capped by a vacuum layer of width $d_{vac} \simeq 13.5\text{\AA}$ in the perpendicular direction in order to prevent residual interactions between the heterostructure (HS) and its periodic replicas. The in-plane area of the heterobilayer is fixed by the TMD unit cell parameter, which is essentially described by the average nearest-neighbor W-W distance, $\bar{d}_{W-W} = 3.15\text{\AA}$. This value is less than 0.1% smaller than the experimental value of 3.153\AA [1]. The average chalcogen-chalcogen distance that results from the geometry optimization in the system is found to be $\bar{d}_{S-S} = 3.10\text{\AA}$. The average interlayer distance between the TMD and graphene layer within the supercell, defined as the minimal distance between the layers, is $\bar{d}_{inter} = 3.43\text{\AA}$.

Geometry of the 30°-twist heterobilayer The 30°-twist heterobilayer is composed of 4×4 TMD unit cells and $3\sqrt{3} \times 3\sqrt{3}$ graphene unit cells in a rhombohedral structure with supercell in-plane lattice parameters of length $a = 12.6\text{\AA}$ (see Fig. 1 (b) in the main text). The initial supercell geometry for this twist angle was created with CELLMATCH [2]. As in the 0°-twist case, the supercell contains a vacuum layer of width $d_{vac} \simeq 13.5\text{\AA}$ to prevent residual spurious interactions between the periodic replicas of the heterobilayer. The in-plane area of the heterobilayer is constrained by the TMD unit cell parameter. As for the 0°-twist case, this value is given by the average nearest-neighbor W-W distance of $\bar{d}_{W-W} = 3.15\text{\AA}$, resulting in the same small amount of residual strain. The average chalcogen-chalcogen distance that results from the geometry relaxation in the system is $\bar{d}_{S-S} = 3.16\text{\AA}$, which is around 2% larger than for the 0°-twist HS due to the differences in the long-range (Moiré) potentials. The average interlayer distance between the TMD and the graphene layer within the supercell, defined as the minimal distance between the layers, is $\bar{d}_{inter} = 3.40\text{\AA}$.

B. Density functional theory

We performed the density functional theory (DFT) calculations using the QUANTUM ESPRESSO (QE) package [3–5]. We employed a generalized gradient approximation for the exchange–correlation functional (PBE, [6]). The DFT wavefunctions are expanded in a plane-wave basis set with a basis cut-off of 50 Ry. We included non-collinear spin–orbit interaction within the fully relativistic norm-conserving pseudopotentials from the PSEUDO DOJO project [7]. A $6 \times 6 \times 1$ Monkhorst-Pack [8] \mathbf{k} -grid was employed for the convergence of the charge-density with a Fermi-Dirac smearing of 10^{-4} Ry. The calculation is considered to be converged once the total energy varies less than the threshold value of 10^{-9} Ry between consecutive self-consistent field iterations.

The geometries of both heterostructure supercells were optimized using QE with the PBE functional, a basis set energy cut-off of 75 Ry, and van der Waals interactions incorporated within the Tkatchenko-Scheffler method [9] to properly account for the interlayer separation. The positions of the atoms within the supercell were relaxed (while keeping the supercell lattice vectors fixed) using convergence criteria for the force of 10^{-4} Ry/ a_0 and considering the total energy converged within 10^{-6} Ry.

C. GW calculations

Using the BerkeleyGW package [10–13], we computed the quasi-particle energy spectrum including spin-orbit interaction within many-body perturbation theory by performing a one-shot non-self-consistent GW calculation (G_0W_0), employing the DFT wavefunctions and energies as a starting point. The dielectric function was obtained using the generalized plasmon-pole approximation [10]. We considered as well a non-uniform neck \mathbf{k} -space subsampling scheme [14] for faster convergence of the dielectric function in 2D materials with respect to the Brillouin zone (BZ) sampling. A uniform \mathbf{q} -grid of $6 \times 6 \times 1$ points was considered within this method, with 10 additional points in the surroundings of and replacing the $\mathbf{q} = \mathbf{0}$ point. We employed a dielectric screening cut-off 10 Ry and a total of ~ 2500 bands (~ 1900 unoccupied bands) in the summation over occupied and empty states. The Coulomb interaction was truncated in the out-of-plane direction to prevent interactions between the periodic images in this direction [15]. Overall, these parameters led to converged quasiparticle band gaps within 50 meV.

D. Bethe-Salpeter equation

The excitonic properties of the heterobilayers were obtained by solving the Bethe-Salpeter equation (BSE) [16]

$$(E_{c\mathbf{k}} - E_{v\mathbf{k}})A_{vc\mathbf{k}}^X + \sum_{v'c'\mathbf{k}'} K_{vc\mathbf{k};v'c'\mathbf{k}'}^{\text{eh}} A_{v'c'\mathbf{k}'}^X = \Omega_X A_{vc\mathbf{k}}^X, \quad (1)$$

where $E_{c(v)\mathbf{k}}$ are the quasiparticle energies of the conduction (valence) bands, $K_{vc\mathbf{k};v'c'\mathbf{k}'}^{\text{eh}} = \langle v\mathbf{c}\mathbf{k} | \hat{K}^{\text{eh}} | v'\mathbf{c}'\mathbf{k}' \rangle$ are the matrix elements of the electron-hole interaction kernel, the eigenvalues Ω_X correspond to the energy of the excitons $\{S\}$ and $A_{vc\mathbf{k}}^X$ represents the amplitude of the exciton state $|\Psi^X\rangle$ expressed in the electron-hole basis. In other words, the BSE (1) assumes that the excitonic wavefunction can be written as a coherent superposition of electron-hole states with crystal momentum \mathbf{k}

$$|\Psi^X\rangle = \sum_{vc\mathbf{k}} A_{vc\mathbf{k}}^X |\psi_{v\mathbf{k}}^*\rangle |\psi_{c\mathbf{k}}\rangle. \quad (2)$$

with $|\psi_{c\mathbf{k}}\rangle$ the spinor state describing an electron with conduction band quantum number c and crystal momentum \mathbf{k} , and $|\psi_{v\mathbf{k}}^*\rangle$ the spinor state describing a hole characterized by the valence band quantum number v and same crystal momentum \mathbf{k} .

We solved Eq. (1) using the BerkeleyGW package [11, 12, 16]. The electron-hole interaction kernel matrix elements were computed within a coarse uniform $9 \times 9 \times 1$ \mathbf{k} -grid and the results were interpolated to a finer uniform $36 \times 36 \times 1$ \mathbf{k} -grid. This finer grid is then used for the decomposition of the absorption spectrum, excitonic energies as well as any further analysis. The interaction kernel was computed as an *extended* kernel which takes into account all possible transitions between pairs of bands $(n, m) \rightarrow (n'm')$. The BSE was then solved within the Tamm-Dancoff approximation (TDA) [16]. The kernel was computed with 16 valence and 16 conduction bands, absorption calculations consider 14 valence and 14 conduction bands. These bands include all the relevant energy levels of the graphene and TMD high-symmetry points, in particular the folded \tilde{K}_{TMD} and Λ_{TMD} states, as well as additional bands for convergence reasons. These parameters converge the excitonic spectra of the heterobilayers up to 10 meV. Finally, the absorption calculations were performed by evaluating the momentum operator $\hat{\mathbf{p}} = i\nabla$ instead of the velocity operator $\hat{\mathbf{v}}$, unless otherwise stated explicitly. We thus neglected for the most part, non-local terms [11, 16] at the BSE level.

E. Observables

Absorbance. We computed the absorbance from the absorption obtained using BerkeleyGW as [17],

$$A(\omega) = \frac{\omega L_z}{c} \epsilon_2(\omega), \quad (3)$$

where ω is the optical frequency, c is the speed of light, L_z is the supercell dimension in the out-of-plane direction and $\epsilon_2(\omega)$ is the absorption (imaginary part of the dielectric function).

Exciton binding energy. The exciton binding energy is defined from the difference between the “exciton energies” without electron–hole interactions and the full exciton energies obtained from the solution of the BSE. In practice, we compute the expectation value of the diagonal part of the BSE Hamiltonian and subtract from this value the exciton energy, *i.e.*,

$$\begin{aligned} E_{\text{bind}}^X &= \langle \Psi^X | \hat{H}^{\text{BSE}} - \hat{K}^{\text{eh}} | \Psi^X \rangle - \langle \Psi^X | \hat{H}^{\text{BSE}} | \Psi^X \rangle, \\ &= \sum_{vc\mathbf{k}} |A_{vc\mathbf{k}}^X|^2 (E_{c\mathbf{k}} - E_{v\mathbf{k}}) - \Omega_X, \end{aligned} \quad (4)$$

where \hat{H}^{BSE} is the BSE Hamiltonian, \hat{K}^{eh} the electron-hole interaction kernel; $E_{c(v)\mathbf{k}}$ are the quasi-particle energies of the conduction (c) or valence (v) bands at point \mathbf{k} .

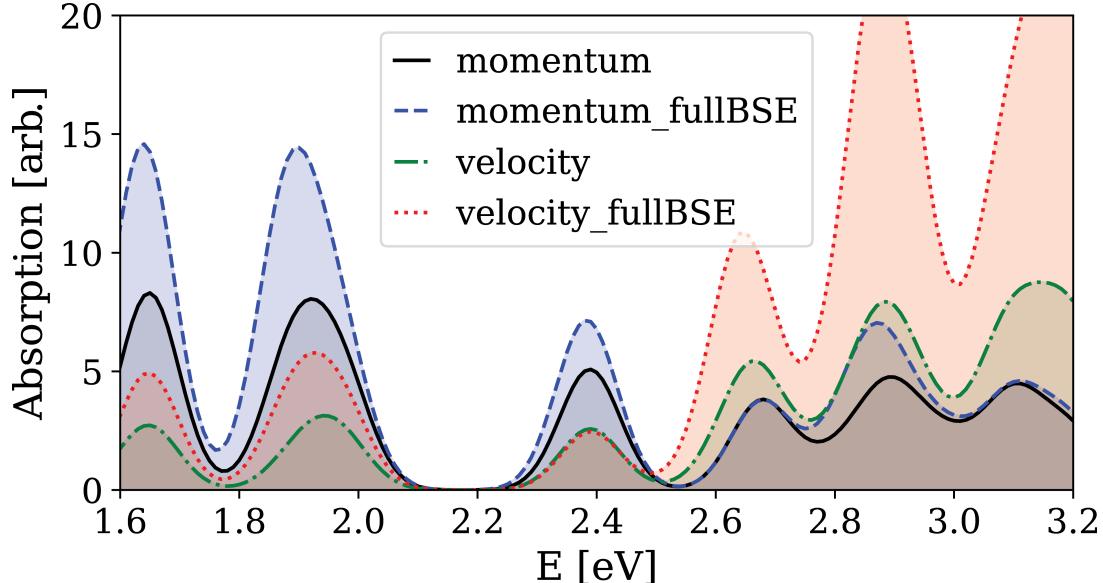


FIG. 1. Comparison of absorption spectra for the 0° -twist HS, computed with the momentum operator within the TDA (dot-dashed red line), the momentum operator solving the f-BSE (solid black line), the velocity operator solving within the TDA (dashed green line), and the velocity operator solving the f-BSE (dotted blue line). The optical activity of the resonances depends on approximation, but the qualitative form of the absorption spectrum or the positions of the resonant peaks remain unaffected.

II. SUPPLEMENTARY METHODS - CONVERGENCE TESTS

A. GW convergence tests

Energy cutoff for the dielectric screening The large dimensions of the system cause GW calculations, and in particular the dielectric function, to be difficult to calculate both by means of the CPU time and the memory requirements. The main factor contributing to the computational effort required is the dielectric cutoff chosen for the calculations. For this reason, we chose to perform our calculations with an under-converged energy cut-off value of 10 Ry. In order to assess the error in these calculations, less converged calculations with an energy cut-off value of 5 Ry were performed as well. The resulting calculations exhibit the same band structure, with a slight renormalization of the band gaps, which increase by ~ 45 meV for the 0° -twist HS and by ~ 65 meV for the 30° -twist HS with increasing energy cut-off. This slight renormalization was found to be within a reasonable enough error bar.

B. BSE convergence tests

Approximations in the BSE. While solving the BSE several approximations can be applied, namely solve the full BSE (f-BSE) or employ the Tamm-Danoff approximation (TDA). Moreover, in the absorption calculation, we can evaluate the velocity operator (\hat{v}) or approximate it by the momentum operator (\hat{p}). Semiconducting periodic solid systems are usually studied by BSE within the TDA and evaluating the momentum operator. Due to the semimetallic nature of our systems, we check the validity of the approximations by means of the difference in the absorption spectra computed for the 0° -twist HS. The results of these calculations, computed with a $6 \times 6 \times 1$ coarse \mathbf{k} -grid for the kernel and interpolated into a finer $9 \times 9 \times 1$ \mathbf{k} -grid for the absorption are shown in Fig. 1. We conclude that though there can be substantial differences in the optical activity at the absorption resonances, there are no significant differences in the shape of the spectra as well as in the position of the resonances between the different approximations.

Grid convergence. The absorption that results from the BSE is highly sensitive to the \mathbf{k} -space sampling since the optical transitions can be strongly dependent on the selected points in the \mathbf{k} -space. In the main paper, we employ a $36 \times 36 \times 1$ \mathbf{k} -grid. Here, we assess the error associated with the \mathbf{k} -grid sampling by considering coarser \mathbf{k} -grids for the 0° -twist HS and comparing the resulting absorption spectra in Fig. 2. We observe that for samplings of $18 \times 18 \times 1$ or finer, the position of the high-energy absorption resonances associated with the TMD \bar{A} and \bar{B} transitions are already

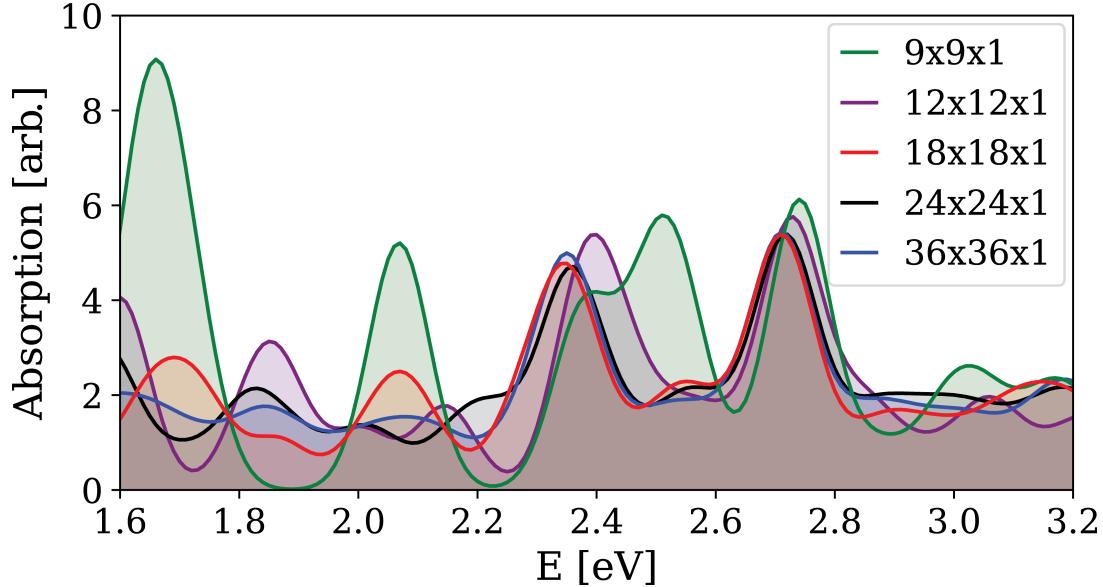


FIG. 2. Comparison of the full heterobilayer absorption spectra for five different Brillouin zone samplings. The position and width of the absorption resonances below the peak at 2.2 eV are highly sensitive to the \mathbf{k} -grid sampling. However, TMD-dominated absorption resonances at and above the 2.2 eV are robust to the choice of \mathbf{k} -space sampling, except for the coarse $9 \times 9 \times 1$ \mathbf{k} -grid, which is thus deemed under-converged.

converged, with a negligible change between the $24 \times 24 \times 1$ and the $36 \times 36 \times 1$ spectra. We further note that the oscillations in the graphene low-energy region become less pronounced with denser grids.

Band convergence. The absorption spectrum is highly sensitive to the amount of conduction and valence bands used in the calculation, since these affect which transitions are taken into account. A full description of the absorption spectrum requires a calculation with all valence and all conduction states of the system, but the spectrum in a limited energy region can be computed with a finite number of bands. A minimal calculation requires to include at least 4 valence and 10 conduction bands, in order to include the graphene low-energy bands and the TMD K and Λ -associated bands. We consider at least 10 valence bands in our calculations to account for accidental degeneracies. We checked the resulting absorption spectrum of the 0° -twist HS with 10, 12, and 14 valence and conduction bands in Fig. 3. We find that the convergence of the \bar{B} peak optical region requires a minimum of 12 conduction bands in the calculation.

GW energy cut-off convergence Since the quasiparticle band gap depends on the energy cutoff chosen for the dielectric matrix calculation, the absorption spectrum should depend on it as well. In Fig. 4 we show that the effect of the energy cut-off from our GW calculations is a blue shift of the relevant TMD features.

III. SUPPLEMENTARY DISCUSSION - FOLDING AND SYMMETRY EFFECTS AT THE MINI-BZ

In this section, we show how to identify the K- and Λ -originating energy levels for the WS_2 –Gr HS presented in the main manuscript, both folding on the same high-symmetry point (K) of the supercell mini Brillouin zone (BZ).

A. Atomic reconstruction and BZ \mathbf{k} -point assignment

In order to identify the origin of the energy levels in the HS band structure, we computed four related systems: a unit cell of TMD (WS_2), a supercell of 4×4 repeated unit cells of TMD, a TMD supercell in which the geometrical reconstruction (geometry optimization) caused by the presence of the graphene layer is taken into account, and finally the relaxed full HS. We focus on the lowest spin-orbit split conduction band and the two highest valence bands (note that the energy of the valence bands at Λ is substantially lower than the valence band energy at K and therefore, the role in the band folding is subdominant). The identification of the origin of the energy levels at K in the supercell

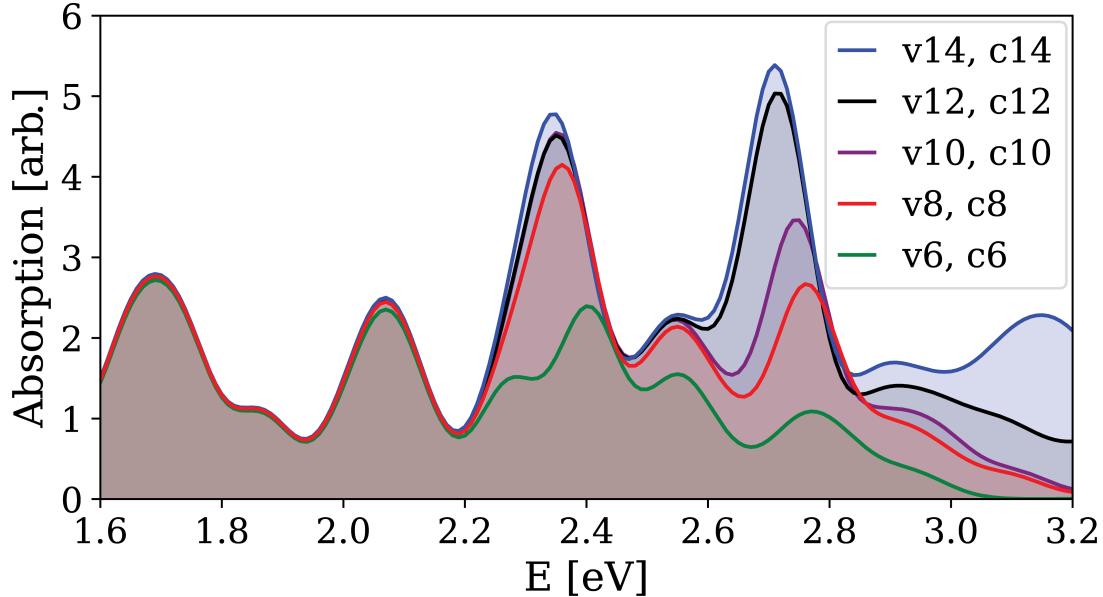


FIG. 3. Comparison of the 0°-HS twist absorption spectra computed with 6 (green), 8(red), 10 (purple line), 12 (black line) and 14 (blue line) conduction and valence bands.

arrangement can be performed straightforwardly due to the fact that the only difference between the supercell and the unit cell is the folding in \mathbf{k} -space, see Fig. 5 (a). In the valence band region, we find that the two topmost bands are split by ~ 0.4 eV, consistent with the large spin-orbit coupling in WS₂. In the conduction band region, the two states that come from the K point, which folds onto itself, are split by ~ 20 meV. The folded Λ bands appear to be triple-degenerate (six states in total) and they are split by ~ 0.3 eV, see red lines in Fig. 5 (b). After atomic reconstruction due to graphene adsorption, we find that the K-originating bands shift to higher energies, and the Λ -originating bands shift to lower energies. Furthermore, we find the degeneracy of the Λ states to be broken and levels split by ~ 1 meV. The relative shift in the case of the 0°-twist HS is such that the K-originating bands remain the lowest conduction bands that define the HS band gap. These energy shifts are shown in Fig. 5 (b) marked by black arrows. For the 30°-twist HS the atomic reconstruction leads to a larger shift of the K- and Λ -originating energy levels, therefore the K levels are no longer the lowest conduction bands and instead, we find that the band-gap is determined by the lowest triplet of states originally coming from Λ , followed in increasing energy far from the Fermi level by the K-originating energy levels, which are then followed by the higher spin-orbit split Λ -originating energy-level cluster.

As an illustration of the effect of the supercell, atomic reconstruction, and HS formation on the wavefunctions properties, we show the pseudo-charge density of the four systems for the 0°-twist, both for the K-originating (Fig. 6) and the Λ -originating states (Fig. 7). We can observe in panels (a) and (b) of both figures that the creation of the supercell does not affect the form of the wavefunction density. In the K-originating states, the states are strongly localized in the metallic atoms in the four cases. We estimate that the wavefunction overlap [18] between the supercell and the reconstructed supercell, as well as the reconstructed supercell and the HS for this type of states, is above 95%. Note that, for the Λ -originating wavefunctions, there is a strong impact of the atomic reconstruction that occurs after graphene adsorption. This is related to the previously discussed lift of level degeneracy in the band structure. We estimate that the overlap of the Λ -originating wavefunctions between the supercell and the reconstructed supercell (or HS) to be $\sim 30\%$.

B. Effect of the chalcogen-chalcogen distance on the level alignment

Monolayer TMDs are direct band gap materials with the conduction band minimum (CBM) and valence band maximum (VBM) located at the K point of the monolayer BZ [19]. However, the precise size of the band gap, as well as whether it is direct or indirect, can be very sensitive to the strain within the layer [20–23]. In-plane strain is an experimentally accessible property, and thus the effect of this type of strain on the TMD has been thoroughly studied [24–26]. Out-of-plane strain can appear spontaneously due to the long-range (van der Waals) interaction in multilayered homo- and heterostructures [25].

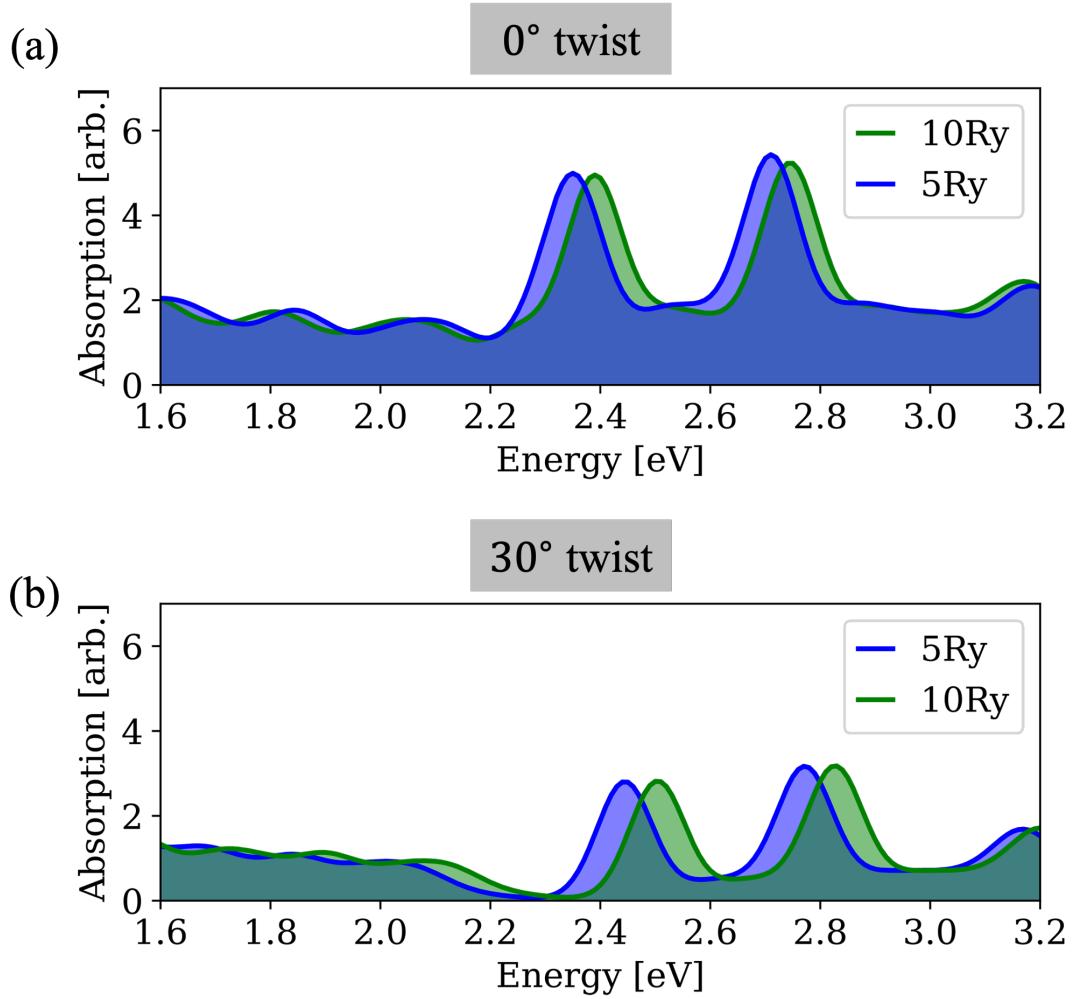


FIG. 4. (a) Comparison of the absorption spectra for the 0° -twist HS computed using a GW calculation with 5 Ry (blue) and 10 Ry (green) energy cutoff. (b) Same as (a) for the 30° -twist HS.

We find that for the unit cell parameter of length 3.150\AA changes in the out-of-plane sulfur-sulfur distance d_{S-S} as small as $1 - 2\%$ can have a large effect on the position of the conduction bands at K and Λ - enough to completely change their relative alignment, see Fig. 8. Increasing the out-of-plane strain causes the K point energy to shift to higher energies, while the Λ point energy shifts to lower energies. This causes a direct-indirect transition in the TMD for certain d_{S-S} , with the exact value in which this transition occurs being dependent on the calculation details, with the DFT direct-to-indirect transition occurring for larger sulfur-sulfur distances compared to GW.

IV. SUPPLEMENTARY METHODS - ASSIGNMENT OF LAYER CONTRIBUTIONS

A. Projected density of states and layer decomposition

To determine the contribution of each layer to the bandstructure, we obtain the \mathbf{k} -resolved projected density of states (\mathbf{k} -PDOS), which is computed by projecting the DFT wavefunctions $\psi_{n\mathbf{k}}$ onto the atomic orbitals labeled by i_A , $\phi_{i_A}^A$, for the set of atoms, $\{A\}$,

$$\rho_{n\mathbf{k}}^{A,i_A}(E) = |\langle \phi_{i_A}^A | \psi_{n\mathbf{k}} \rangle|^2 \delta(E - E_{n\mathbf{k}}), \quad (5)$$

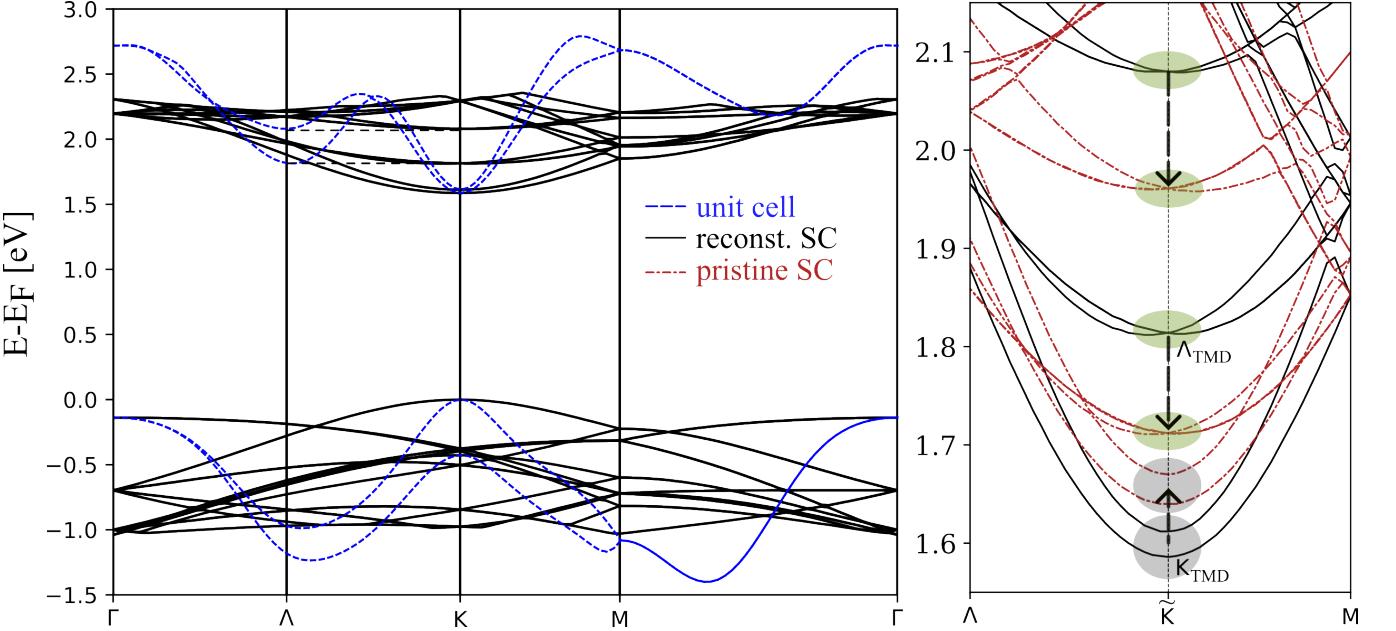


FIG. 5. (a) Band structure of the unit cell (dashed blue lines) and \$4 \times 4\$ supercell (SC) (solid black lines) (note that the high-symmetry points correspond to the crystal momentum of the BZ for the unit cell and the crystal momentum in the mini-BZ for the supercell). We also mark by dashed lines the projection of the unfolded \$\Lambda\$ point onto the folded supercell bands at the K point, allowing for easy identification of the band origin in terms of the energy spectrum. (b) Band structure of the conduction region near the K point of the \$4 \times 4\$ supercell (solid black lines) and the reconstructed monolayer corresponding to the \$0^\circ\$-twist HS (dot-dashed red lines). Black arrows show the renormalization of the K-originating energy levels, marked by grey shaded areas, and the \$\Lambda\$-originating energy levels, marked by green shaded areas.

where \$E_{n\mathbf{k}}\$ are the Kohn-Sham eigenvalues. Each layer contribution is computed from the summation of the \$\mathbf{k}\$-PDoS over all atoms \$\{A\}\$ and atomic orbitals \$\{i_A\}\$ that correspond a given layer \$L = \{\text{WS}_2, \text{Gr}\}

$$\rho_{n\mathbf{k}}^L(E) = \sum_{i_A, A \in L} \rho_{n\mathbf{k}}^{i_A, A}(E). \quad (6)$$

This quantity is normalized as \$\tilde{\rho}(E) = \rho(E) / \max[\rho(E)]\$, thus, the object

$$\tilde{\rho}_{n\mathbf{k}}(E) = \tilde{\rho}_{n\mathbf{k}}^{\text{WS}_2}(E) - \tilde{\rho}_{n\mathbf{k}}^{\text{Gr}}(E), \quad (7)$$

is defined in the range \$[-1, 1]\$, with the value of \$\tilde{\rho}_{n\mathbf{k}}(E) = 1\$ corresponding to pure contribution from the TMD layer and \$\tilde{\rho}_{n\mathbf{k}}(E) = -1\$ to pure contribution from the graphene layer. The quantity \$\tilde{\rho}_{n\mathbf{k}}(E)\$ is used for the color coding in this paper (e.g. Fig. 1 of the main paper).

B. Layer contributions to the excitons

The BSE outputs the excitonic amplitudes \$A_{cv\mathbf{k}}^X\$ which show for each exciton \$S\$ the amplitude of the transition from band \$v\$ to band \$c\$ at the \$\mathbf{k}\$-point \$\mathbf{k}\$ in the BZ. We use the absolute values of these amplitudes along with the quantity \$\tilde{\rho}_{n\mathbf{k}}(E)\$ from Eq. (7) to compute the contribution of each layer to each exciton. Since the BSE solutions have transitions to, in principle, multiple conduction levels from multiple valence levels, the basic step of the assignment of the layer contribution to the excitons consists of summing the layer contributions separately for each energy level and each \$\mathbf{k}\$-point beforehand. Thus, we use the following expression for the layer contributions of the valence bands for each exciton:

$$\chi_{v\mathbf{k}}^{LX} = \sum_c \rho_{v\mathbf{k}}^L |A_{cv\mathbf{k}}^X|^2 \quad (8)$$

Similarly, we get the overall layer composition of an exciton by summing over the layer contributions of all \$\mathbf{k}\$-points for each exciton as \$\chi_v^{LX} = \sum_{\mathbf{k}} \chi_{v\mathbf{k}}^{LX}\$, as well as the (\$\mathbf{k}\$-resolved) layer contributions for an absorption resonance \$\mathcal{R}\$ by

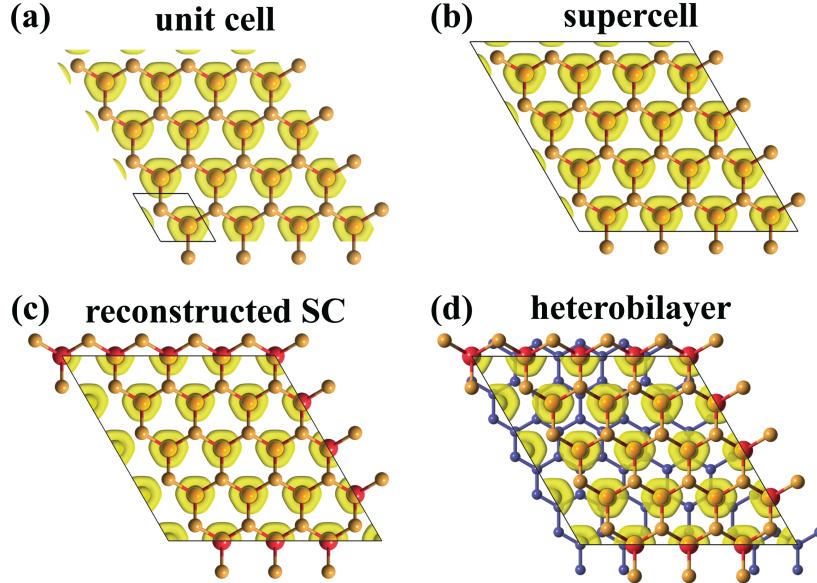


FIG. 6. Top view of the pseudo-charge density, $|\Psi_{n\kappa}(\mathbf{r})|^2$, for the wavefunctions corresponding to the energy levels with K_{TMD} origin. We display the pseudo-charge density for (a) TMD unit cell, (b) TMD 4×4 supercell, (c) 4×4 supercell with atomic reconstruction matching that of the HS and (d) full 0° -twist heterobilayer.

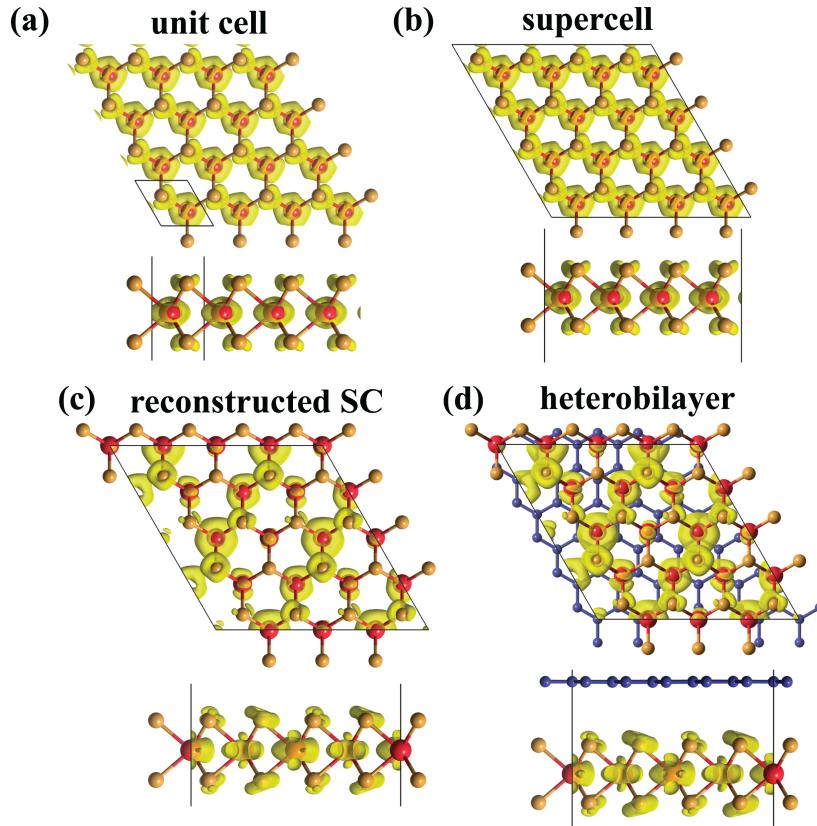


FIG. 7. Top and side view of the pseudo-charge density, $|\Psi_{n\kappa}(\mathbf{r})|^2$, for the wavefunctions corresponding to the energy levels with Λ_{TMD} origin. We display the pseudo-charge density for (a) TMD unit cell, (b) TMD 4×4 supercell, (c) 4×4 supercell with atomic reconstruction matching that of the HS and (d) full 0° -twist heterobilayer.

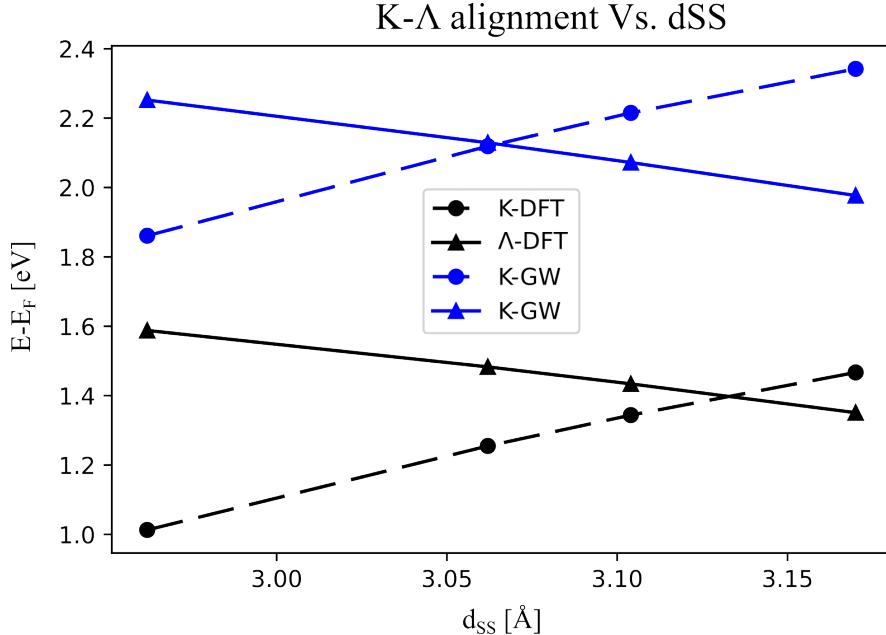


FIG. 8. Energies of the CBM at the K point (full circles) and the Λ point CBM (full triangles) as a function of the chalcogen-chalcogen distance (d_{S-S}) in a WS₂ unit cell. The lines are a guide to the eye. The unit cell lattice parameter has a length of 3.150 Å. We consider DFT (black) and GW (blue) calculations. For both DFT and GW, the K CBM increases in energy, while the Λ CBM decreases in energy with increasing d_{S-S} , but the direct-to-indirect transition of the material occurs at different d_{S-S} .

summing over all the excitons within an optical energy interval \mathcal{I} as $\chi_v^{LR} = \sum_{S \in \mathcal{I}_R} \mu_X \chi_v^{LX}$. For the conduction bands, the same expression applies with the replacement $v \rightarrow c$ in the sums and the layer compositions. Unless stated otherwise, summations over multiple excitons are performed by weighting each exciton with its optical activity as represented by the oscillator strength μ_X [11].

C. Indirect transitions

Due to the construction of the supercell and the BZ folding, momentum-indirect transitions in the unit cells become momentum-direct and thus optically allowed in the supercell arrangement. An example of these transitions are the transitions from the K valleys to the Dirac cone in the 0°-twist HS, which have different momentum in the unit-cell representation, but share the same crystal momentum in the supercell. We found another example of this phenomenon in the emergence of TMD intralayer transitions, between valence bands corresponding to the K point of the TMD and conduction bands corresponding to the Λ point of the TMD. Fig. 9 shows dots with size proportional to the transition coefficients at the K-point and a small vicinity of 0.1 Å⁻¹ from each valence band to each conduction band for all excitons within a 5 meV energy range under peak \bar{A} for both twist angles. This representation shows that though the "indirect" transitions are less favorable than the K-K transitions, they nevertheless occur, and with non-negligible probability.

V. SUPPLEMENTARY DISCUSSION - EXCITON INTEGRATED CONTRIBUTION IN AN OPTICAL ENERGY RANGE

In Fig. 2 of the main text, we display the integrated contribution of excitons within a 5 meV range around the \bar{A} , and \bar{B} peaks, and the largest absorption maximum in the low-energy region, X_l , for each of the HSs, with the excitons weighted by their corresponding oscillator strength (μ_X). We show below in Figs. 10 and 11 the robustness of these results with respect to different energy ranges in summation (*i.e.* different energy intervals, \mathcal{I}). We observe that for energy ranges from 1 meV to 10 meV, the excitonic picture stays consistent but with too-low energy ranges resulting in harder-to-interpret representations. The weighting of the excitons in the summation by their oscillator strength emphasizes the regions in reciprocal space with the largest optical activity for each energy range. In Figs.

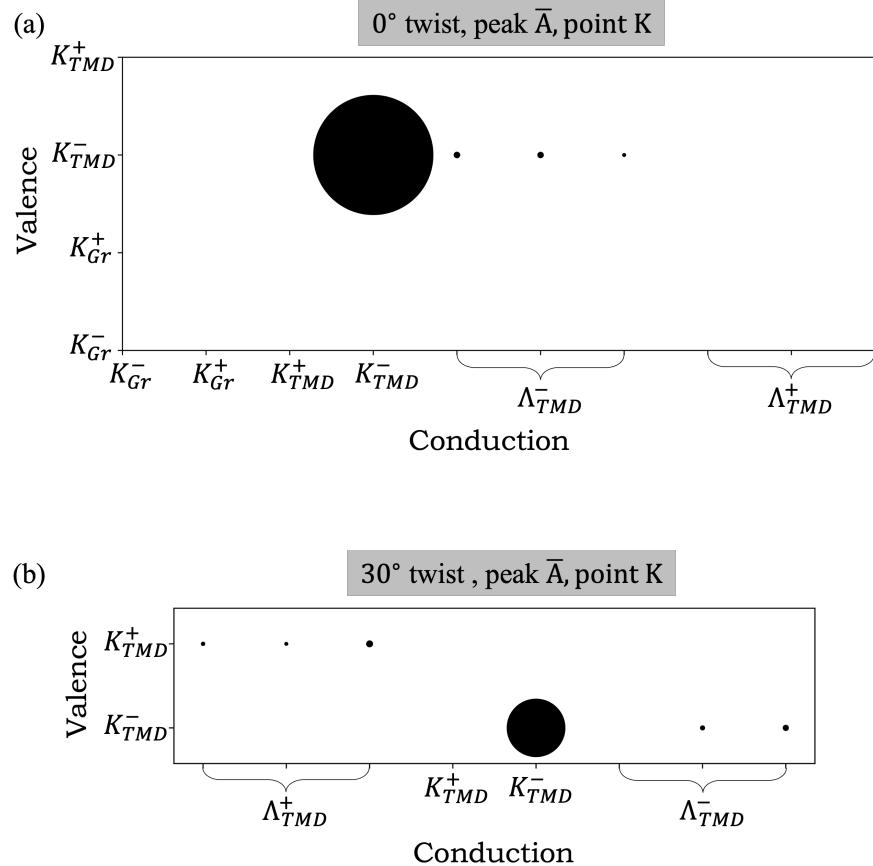


FIG. 9. Transitions contributing to the \bar{A} peak within a small energy range (5meV) at the K-point and it's surroundings (within 0.1\AA^{-1}) for the (a) 0° -twist HS and the (b) 30° -twist HS. Dot sizes are proportional to the transition coefficients.

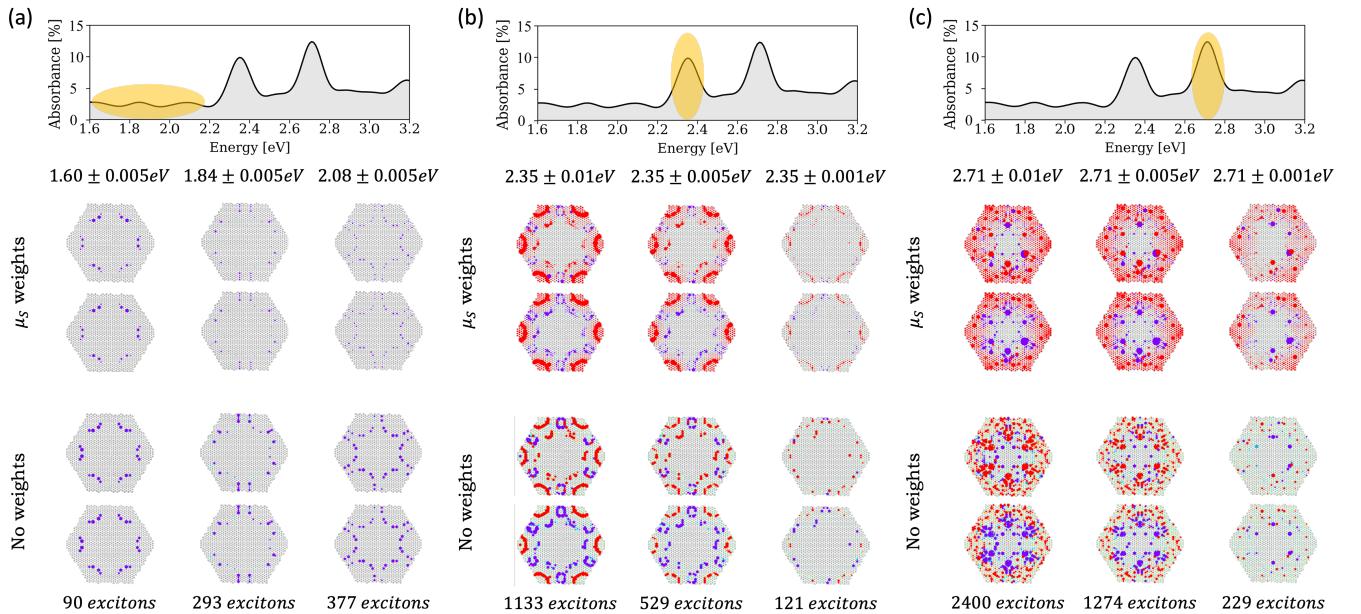


FIG. 10. Exciton integrated contribution within a given energy interval under the (a) low energy, (b) A, and (c) B peaks of the 0° -twist HS. The interval is marked with a yellow ellipse in the absorption spectrum. The upper rows show the summation with the excitons weighted by their corresponding oscillator strength (μ_x , and the lower rows show all non-dark excitons ($\mu_x \geq 1$) since all excitons have equal weights.

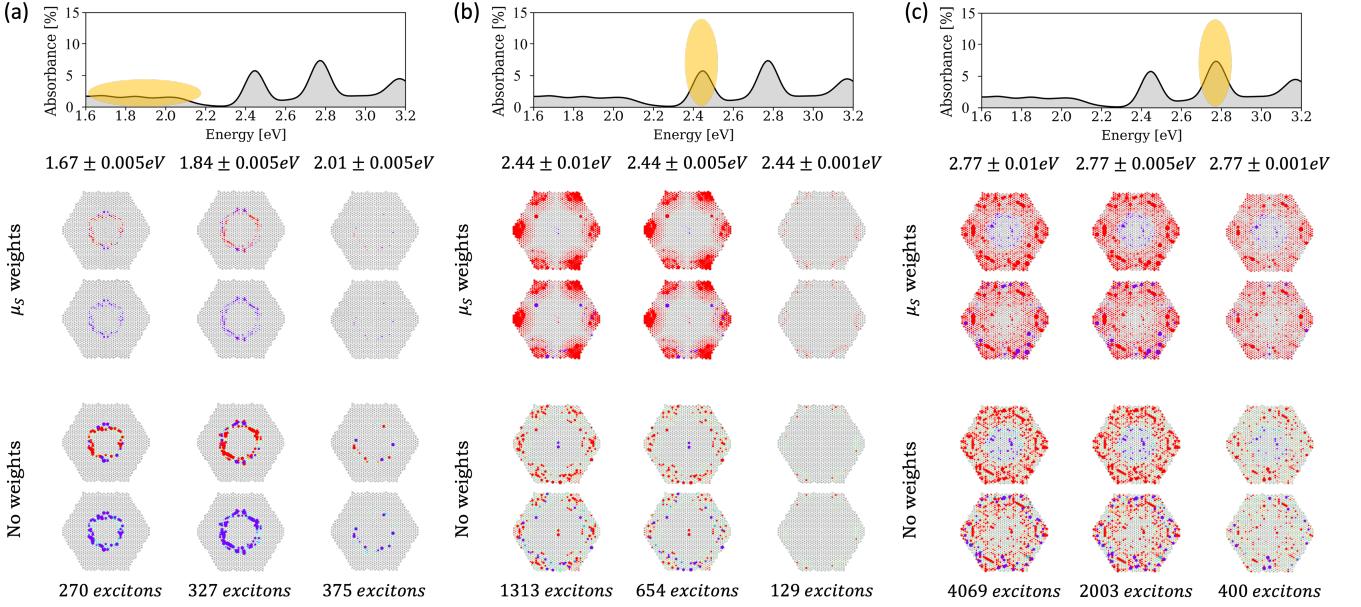


FIG. 11. Exciton integrated contribution within a given energy interval under the (a) low energy, (b) A, and (c) B peaks of the 30°-twist HS. The interval is marked with a yellow ellipse in the absorption spectrum. The upper rows show the summation with the excitons weighted by their corresponding oscillator strength (μ_X , and the lower rows show all non-dark excitons ($\mu_X \geq 1$) since all excitons have equal weights.

10 and 11, we compare the scenario resulting from weighting the excitons with the oscillator strength (upper rows) to those resulting from unweighted excitons (lower rows). We see that the qualitative picture, including the regions of highest optical activity, remains consistent between both representations.

Comparing now the situation between 0°-twist and 30°-twist HSs, we show in Figs. 10 (a) and 11 (a) the \mathbf{k} -resolved transitions for the low-energy region of each spectra. We observe that the excitonic picture exhibited is a K-centered ring for the 0°-twist HS, and a Γ -centered ring for the 30°-twist HS, with the diameter of the ring being larger with higher energies, consistent with the band structure calculations and the analysis of the main paper.

Finally, we show the excitonic picture of the dip in the absorption of the 30°-twist HS at 2.3eV, in Fig. 12. When compared with Fig. 11(a) we see that the dip in the absorption results from the lack of intralayer graphene transitions in this region, which leads to the sharp reduction in the oscillator strength of the excitons. Furthermore, the excitons in this region are Γ -centered, but fall in the K-regions of the mBZ, resulting in a transition from the Γ nature of the low energies, to the K nature of the A peak.

VI. SUPPLEMENTARY DISCUSSION - ANALYSIS OF THE EXCITON NATURE

In the \mathbf{k} -resolved representations of Figs.3 (c) and (d) in the main text, we can distinguish between intralayer TMD, intralayer graphene, and two types of interlayer transitions, which may coexist within a single exciton. For further analysis, we can choose to classify the excitons according to their degree of localization on each of the layers in the HSs (i.e. their *layer nature*). This results in a clear distinction between excitons that involve only TMD transition and only graphene transitions, but not between excitons that involve interlayer transitions or excitons that involve multiple intralayer transitions on both of the layers. Thus, we can categorize the excitons according to the degree of separation between the electron and the hole along the out-of-plane direction (i.e. referred to as *charge separation nature*). This results in a clear distinction between excitons involving interlayer transitions from TMD to graphene and excitons involving interlayer transitions from graphene to TMD but does not give a clear way of distinguishing between excitons that have the charge carriers spread evenly between the layers or located completely on a single layer. By combining these two ways of analyzing the excitons we can classify them according to intralayer TMD, intralayer graphene, interlayer TMD-graphene, interlayer graphene-TMD, and mixed non-charge separated.

This distinction between interlayer excitons and those containing resonant intralayer excitations is performed by combining the layer contribution to each exciton (given by $\chi_c^{LX} + \chi_v^{LX}$), and another characteristic that encodes the degree of “interlayerness” of an exciton by checking whether the electron and the hole reside on the same layer or on

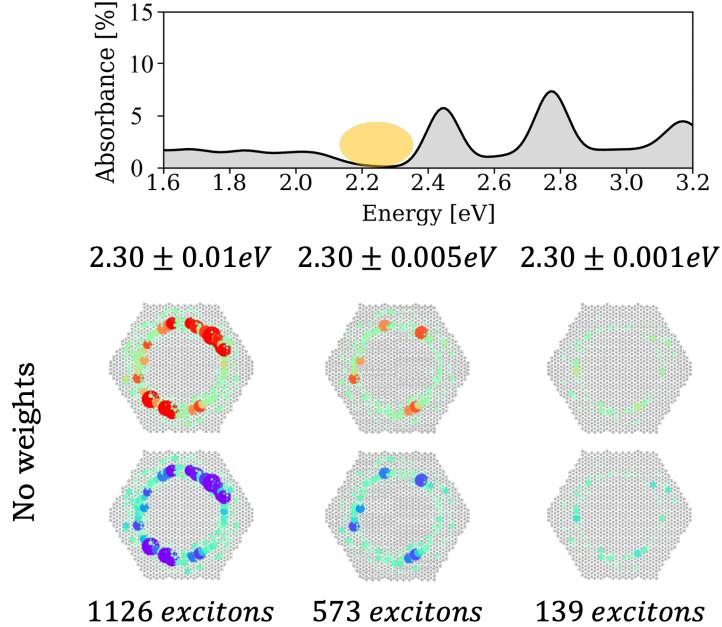


FIG. 12. Exciton integrated contribution within a given energy interval around the dip in the absorption at 2.3eV of the 30°-twist HS. Since all oscillator strengths in this region are negligible only non-weighted sums are performed.

different layers (given by $\chi_c^{LX} - \chi_v^{LX}$). Since this value essentially shows the direction of charge transition during the excitation: either from TMD to graphene, from graphene to TMD, or intralayer, we refer to it as *exciton transition*. The exciton transition for the 0°-twist and the 30°-twist HSs, colored according to the layer contributions to the excitons are shown in the main text in Fig. 3 (a) and (b), respectively. This representation allows the distinction between excitons made with the dot colors: purely TMD exciton marked in red, purely graphene excitons marked in purple, and mixed intralayer excitons marked colors in between according to the colorbar. The further from the center an exciton is located, the larger the interlayer nature of that exciton. Excitons above the center have holes in the graphene layer and electrons on the TMD, while excitons below the center have holes in the TMD layer and electrons on the TMD layer.

VII. SUPPLEMENTARY METHODS - PEAK BROADENINGS

The absorption peaks shown throughout this work are composed of a very large number of excitons, each broadened by a Lorentzian profile. By applying the central limit theorem, we can approximate the distribution of these excitons around the peak as a normal distribution. The basic shape of our computed peaks is then a convolution between the input Lorentzian shape of each exciton and the Gaussian distribution of the excitons, a.k.a. the Voigt profile. We then proceed to fit the peaks with a generalized form of the Voigt profile that takes into account a linear background C to deal with any remaining purely graphene excitons in the vicinity. The peaks are shifted and scaled to be centered at zero, and have a maximal intensity of unity. The resulting function is then:

$$V(x; \sigma, \gamma, C) = \int_{-\infty}^{\infty} G(x; \sigma) L(x - x'; \gamma) dx' + Cx \quad (9)$$

where $G(x; \sigma)$ and $L(x; \gamma)$ are a Gaussian function with variance σ and a Lorentzian function with scale parameter γ , respectively. The width of the Voigt peak (f_V) can then be estimated by using the width of the Gaussian ($f_V = 2\sqrt{2 \ln 2}\sigma$) and Lorentzian ($f_L = 2\gamma$) as $f_V = 0.5346f_L + \sqrt{0.2166f_L^2 + f_G^2}$ [27].

VIII. SUPPLEMENTARY REFERENCES

- [1] W. J. Schutte, J. L. De Boer, and F. Jellinek, Crystal structures of tungsten disulfide and diselenide, *Journal of Solid State Chemistry* **70**, 207 (1987).
- [2] P. Lazić, CellMatch: Combining two unit cells into a common supercell with minimal strain, *Computer Physics Communications* **197**, 324 (2015).
- [3] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougaussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [4] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, Advanced capabilities for materials modelling with Quantum ESPRESSO, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- [5] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. de Gironcoli, P. Delugas, F. Ferrari Ruffino, A. Ferretti, N. Marzari, I. Timrov, A. Urru, and S. Baroni, Quantum ESPRESSO toward the exascale, *J. Chem. Phys.* **152**, 154105 (2020).
- [6] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [7] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese, The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table, *Computer Physics Communications* **226**, 39 (2018).
- [8] H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B* **13**, 5188 (1976).
- [9] A. Tkatchenko and M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.* **102**, 073005 (2009).
- [10] M. S. Hybertsen and S. G. Louie, Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies, *Phys. Rev. B* **34**, 5390 (1986).
- [11] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures, *Computer Physics Communications* **183**, 1269 (2012).
- [12] M. Del Ben, F. H. da Jornada, A. Canning, N. Wichmann, K. Raman, R. Sasanka, C. Yang, S. G. Louie, and J. Deslippe, Large-scale GW calculations on pre-exascale HPC systems, *Computer Physics Communications* **235**, 187 (2019).
- [13] M. Wu, *Spin-Orbit Coupling, Broken Time-Reversal Symmetry, and Polarizability Self-Consistency in GW and GW-BSE Theory with Applications to Two-Dimensional Materials*, Ph.D. thesis, UC Berkeley (2020).
- [14] F. H. da Jornada, D. Y. Qiu, and S. G. Louie, Nonuniform sampling schemes of the Brillouin zone for many-electron perturbation-theory calculations in reduced dimensionality, *Phys. Rev. B* **95**, 035109 (2017).
- [15] S. Ismail-Beigi, Truncation of periodic image interactions for confined systems, *Phys. Rev. B* **73**, 233103 (2006).
- [16] M. Rohlfing and S. G. Louie, Electron-hole excitations and optical spectra from first principles, *Phys. Rev. B* **62**, 4927 (2000).
- [17] L. Yang, J. Deslippe, C.-H. Park, M. L. Cohen, and S. G. Louie, Excitonic Effects on the Optical Response of Graphene and Bilayer Graphene, *Phys. Rev. Lett.* **103**, 186802 (2009).
- [18] R. Steinitz-Eliyahu, D. Hernangómez-Pérez, F. S. Hegner, P. Nikačević, N. López, and S. Refaelly-Abramson, Mixed excitonic nature in water-oxidized BiVO_4 surfaces with defects, *Phys. Rev. Mater.* **6**, 065402 (2022).
- [19] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Atomically Thin MoS₂: A New Direct-Gap Semiconductor, *Phys. Rev. Lett.* **105**, 136805 (2010).
- [20] S. Cianci, E. Blundo, M. Felici, A. Polimeni, and G. Pettinari, Tailoring the optical properties of 2D transition metal dichalcogenides by strain, *Optical Materials* **125**, 112087 (2022).
- [21] H. Shi, H. Pan, Y.-W. Zhang, and B. I. Yakobson, Quasiparticle band structures and optical properties of strained monolayer MoS₂ and WS₂, *Phys. Rev. B* **87**, 155304 (2013).
- [22] L. Wang, A. Kutana, and B. I. Yakobson, Many-body and spin-orbit effects on direct-indirect band gap transition of strained monolayer MoS₂ and WS₂, *Ann. Phys.* **526**, L7 (2014).
- [23] J. Gusakova, X. Wang, L. L. Shiao, A. Krivosheeva, V. Shaposhnikov, V. Borisenko, V. Gusakov, and B. K. Tay, Electronic Properties of Bulk and Monolayer TMDs: Theoretical Study Within DFT Framework (GVJ-2e Method), *Phys. Status Solidi A* **214**, 1700218 (2017).
- [24] A. Kormányos, G. Burkard, M. Gmitra, J. Fabian, V. Zólyomi, N. D. Drummond, and V. Fal'ko, $\mathbf{K} \cdot \mathbf{p}$ theory for two-dimensional transition metal dichalcogenide semiconductors, *2D Mater.* **2**, 022001 (2015).
- [25] A. Ebnonnasir, B. Narayanan, S. Kodambaka, and C. V. Ciobanu, Tunable MoS₂ bandgap in MoS₂-graphene heterostructures, *Appl. Phys. Lett.* **105**, 031603 (2014).

- [26] E. Cappelluti, R. Roldán, J. A. Silva-Guillén, P. Ordejón, and F. Guinea, Tight-binding model and direct-gap/indirect-gap transition in single-layer and multilayer MoS₂, Phys. Rev. B **88**, 075409 (2013).
- [27] J. J. Olivero and R. L. Longbothum, Empirical fits to the Voigt line width: A brief review, Journal of Quantitative Spectroscopy and Radiative Transfer **17**, 233 (1977).