Microscopic understanding of ultrafast charge transfer in van-der-Waals heterostructures: Supplemental Material

R. Krause,[1,2,*] S. Aeschlimann,[1,2] M. Chávez-Cervantes,[2] R. Perea-Causin,[3] S. Brem,[4] E. Malic,[4,3] S. Forti,[5] F. Fabbri,[5,6,7] C. Coletti,[5,7] and I. Gierz[1,†]

[1] *University of Regensburg, Institute for Experimental and Applied Physics, Regensburg, Germany*

[2] *Max Planck Institute for the Structure and Dynamics of Matter, Center for Free Electron Laser Science, Hamburg, Germany*

[3] *Chalmers University of Technology, Department of Physics, Gothenburg, Sweden*

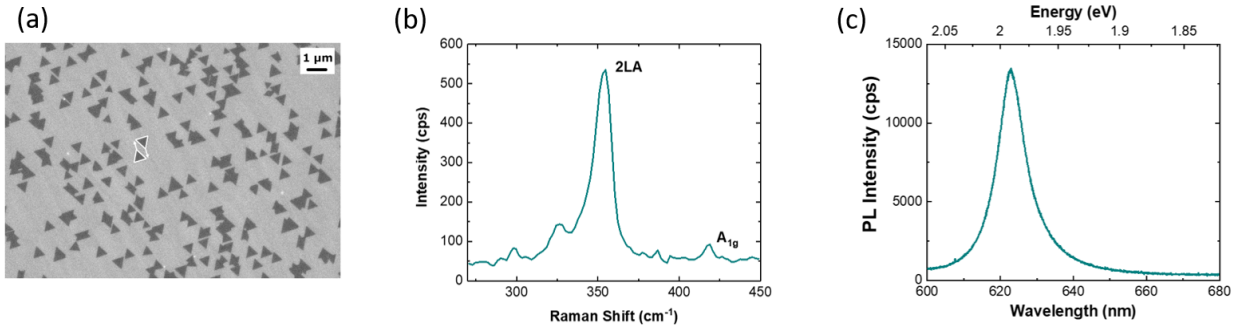
[4] *Philipps-Universität Marburg, Department of Physics, Marburg, Germany*

[5] *Center for Nanotechnology Innovation at NEST, Istituto Italiano di Tecnologia, Pisa, Italy*

[6] *NEST, Istituto Nanoscienze, CNR and Scuola Normale Superiore, Pisa, Italy*

[7] *Graphene Labs, Istituto Italiano di Tecnologia, Genova, Italy*

(Dated: September 29, 2021)



SFig. 1. **Sample characterization:** (a) Secondary electron microscopy image. (b) Raman spectrum. (c) Photoluminescence spectrum.

I SAMPLE GROWTH AND CHARACTERIZATION

Graphene samples were grown on N-doped 6H-SiC(0001) with a miscut below 0.5° bought from SiCrystal GmbH. The substrate was first cleaned and smoothed via etching in hydrogen atmosphere and then graphitized by annealing in argon atmosphere [1] until a carbon buffer layer with a $(6\sqrt{3} \times 6\sqrt{3})$R30° surface reconstruction was formed. The carbon buffer layer was then decoupled from the substrate via hydrogen intercalation at 800°C [2] in order to obtain a quasi-free-standing graphene monolayer. The whole process was carried out in a commercial Black Magic™ reactor from Aixtron. The WS_2 growth was carried out in a standard hot-wall reactor by chemical vapor deposition (CVD) [3, 4]. WO_3 and S powders with a weight ratio of 1:50 were used as precursors. The WO_3 and S powders were kept at 900°C and 120°C, respectively. The WO_3 powder was placed close to the substrate. Argon was used as carrier gas with a flow of 80 sccm. The pressure in the reactor was kept at 1 mbar. The synthesis process took 30 min. The samples were characterized with secondary electron microscopy, Raman, and photoluminescence spectroscopy. The results are shown in SFig. 1.

II TIME- AND ANGLE-RESOLVED PHOTOEMISSION SPECTROSCOPY DATA ANALYSIS

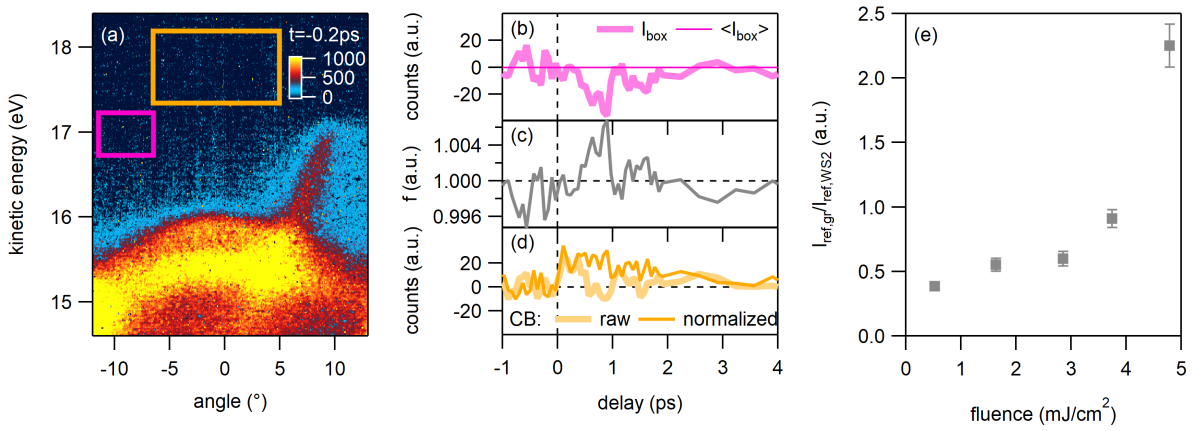
experimental setup

We performed the tr-ARPES experiments in an 2 eV-pump/XUV-tr-ARPES-probe setup based on a 1 kHz Titanium:Sapphire amplifier (Coherent Legend Elite Duo). Extreme ul-

traviolet (XUV) pulses were obtained from high harmonics generation (HHG) in an argon gas jet. Probe pulses at 26 eV photon energy with a nominal pulse duration of 100 fs were selected with a grating monmochromator. Pump pulses were obtained by frequency doubling of the signal output of an optical parametric amplifier (HE-TOPAS from Light Conversion). The pump fluence range employed in the present study was limited by noise on the lower side and pump-induced space charge on the upper side. ARPES snapshots were measured with a hemispherical analyzer (SPECS Phoibos 100). The photocurrent is detected on a two-dimensional detector as a function of emission angle θ and kinetic energy E_{kin} of the photoelectrons. The emission angle θ is related to the in-plane momentum of the electrons inside the solid via $k_{||} = \frac{\sqrt{2m_e}}{\hbar} \sin\theta \sqrt{E_{\mathrm{kin}}}$, where m_e is the electron mass. The binding energy E_B of the electrons inside the solid can be obtained from $E_B = \hbar\omega - E_{\mathrm{kin}} - \phi$, where $\hbar\omega$ is the photon energy and ϕ is the work function of the analyzer, that — in the absence of space charge effects — determines the kinetic energy where the Fermi level appears in the measurement. The energy resolution of the tr-ARPES measurements was determined from Fermi-Dirac fits of the carrier distribution inside the Dirac cone at negative pump-probe delay. The fitting function consisted of the Fermi-Dirac distribution convolved with a Gaussian to account for the finite energy resolution. These fits revealed that the kinetic energy at which the Fermi level appeared in the measurement (SFig. 4a) as well as the energy resolution (SFig. 4b) depend on the applied pump fluence. The temporal resolution of the tr-ARPES measurements of 200 fs was deduced from the full width at half maximum (FWHM) of the width of the derivative of the rising edge of the photocurrent integrated over the area marked by the red box in Fig. 1f. The conversion of the measured photocurrent from $(E_{\mathrm{kin}}, \theta)$ to $(E_B, k_{||})$ relies on an accurate value for the kinetic energy of the Fermi level that is commonly used as energy reference in ARPES data. To solve this issue, we extrapolated the kinetic energy of the Fermi level in SFig. 4a to zero pump fluence, yielding $E_{F,0} = 17.48\,\mathrm{eV}$.

treatment of raw data

The qualitiy of the tr-ARPES raw data in the present study was impaired by three main factors: (1) a featureless background due to dark counts on the detector as well as photoelectrons emitted from the pump rather than the probe pulse, (2) fluctuating ARPES counts



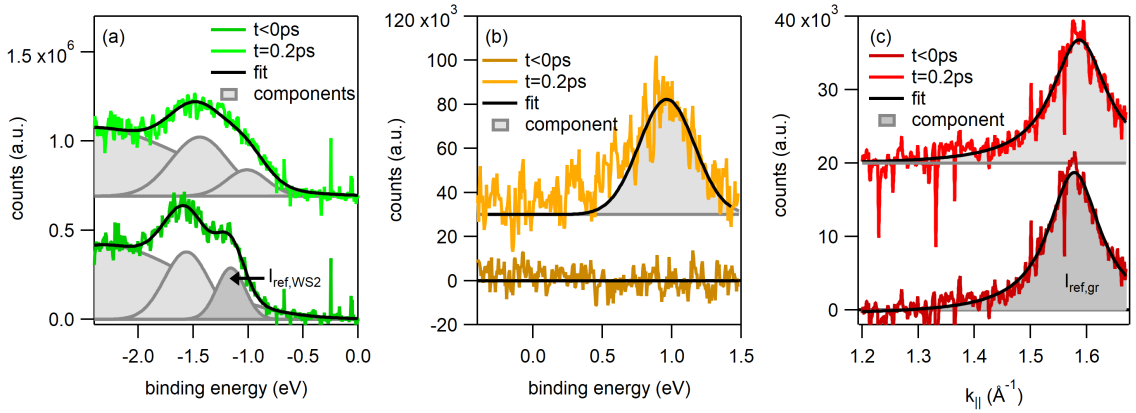
SFig. 2. Treatment of raw data: (a) Tr-ARPES snapshot at negative pump-probe delay taken with a pump fluence of 1.63 mJ/cm^2. The pink box is used to extract $I_{box}(t)$ in panel (b). The orange box captures the carriers at the bottom of the conduction band shown in panel (d). (b) Photocurrent $I_{box}(t)$ integrated over the area marked by the pink box in (a) together with its average for negative pump-probe delays $< I_{box} >$. (c) Normalization factor $f =< I_{box} > /I_{box}(t)$ as a function of pump-probe delay. (d) Effect of the raw data correction on the time dependence of the photocurrent integrated over the area marked by the orange box in panel (a). (e) Ratio of the reference intensities $I_{ref,gr}$ and $I_{ref,WS2}$ as defined in SFig. 3c and a, respectively, for the datasets obtained with different pump fluences.

due to shot-to-shot intensity variations of the XUV pulses, and (3) variations of the relative intensity of WS_2 and graphene bands due to sample inhomogeneity. These problems were solved as follows:

(1) We measured the photocurrent with pump but without probe pulse and subtracted this background from the tr-ARPES snapshots.

(2) We integrated the counts over the area marked by the pink box in SFig. 2a where the photocurrent is expected to be zero after background subtraction. This yields a photocurrent $I_{box}(t)$ that fluctuates with time t (SFig. 2b). Next, we averaged $I_{box}(t)$ at negative delays yielding $< I_{box} >$. We then multiplied each tr-ARPES snapshot with a factor $f =< I_{box} > /I_{box}(t)$ (SFig. 2c). The effect of this procedure on the transient photocurrent at the bottom of the conduction band obtained by integrating the tr-ARPES data over the area marked by the orange box in SFig. 2a is illustrated in SFig. 2d.

(3) As the WS_2 coverage in our WS_2/graphene heterostructure is not uniform, the relative intensity of the WS_2 and the graphene bands varies from spot to spot and, hence, between different tr-ARPES runs. In order to turn our results independent of the precise value of the WS_2 coverage we divided the pump-probe traces shown in Fig. 2a by the integrated

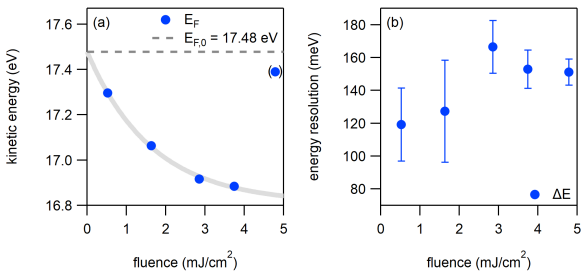


SFig. 3. Determination of transient band positions: (a) EDCs extracted along the dashed green line in Fig. 1a at $t < 0$ fs (light green) and at $t = 200$ fs (dark green) together with Gaussian fit (black) in the WS_2 VB region. The shaded areas indicate the three individual Gaussians used to fit the data. The area of the dark-shaded Gaussian $I_{ref,WS2}$ was used to turn the measured photocurrent independent of the WS_2 coverage. (b) EDCs extracted along the dashed green line in Fig. 1a from the differential tr-ARPES data shown in the lower panel of Fig. 1 in the WS_2 CB region for $t < 0$ fs (dark orange) and $t = 200$ fs (light orange) together with Gaussian fit (black). (c) MDCs extracted along the dashed red line in Fig. 1a for $t < 0$ fs (dark red) and $t = 200$ fs (light red) together with Lorentzian fit. The area of the dark-shaded Lorentzian $I_{ref,gr}$ was used to turn the measured photocurrent independent of the WS_2 coverage. The data was obtained for a pump fluence of 2.85 mJ/cm^2.

intensity of the upper WS_2 valence band I_{ref,WS_2} from SFig. 3a and the pump-probe trace shown in Fig. 2b by the integrated intensity of the Dirac cone $I_{ref,gr}$ from SFig. 3c. The ratio $I_{ref,gr}/I_{ref,WS_2}$ for the different tr-ARPES runs with different pump fluences is shown in SFig. 2c. The measurement with the highest pump fluence was done on a sample area where the WS_2 coverage was particularly low. This reduced the absolute number of photoexcited electron-hole pairs and, hence, the absolute number of holes that were transferred to the graphene layer as well as the measured shift of the Dirac cone. For this reason, the high fluence data point in SFig. 10c and d was set in brackets and not included for the determination of the guide to the eye.

k-space conversion and experimental resolution

The photocurrent is detected on a two-dimensional detector as a function of emission angle θ and kinetic energy E_{kin} of the photoelectrons. The emission angle θ is related to the in-plane momentum of the electrons inside the solid via $k_{||} = \frac{\sqrt{2m_e}}{\hbar} \sin\theta \sqrt{E_{\text{kin}}}$, where m_e is

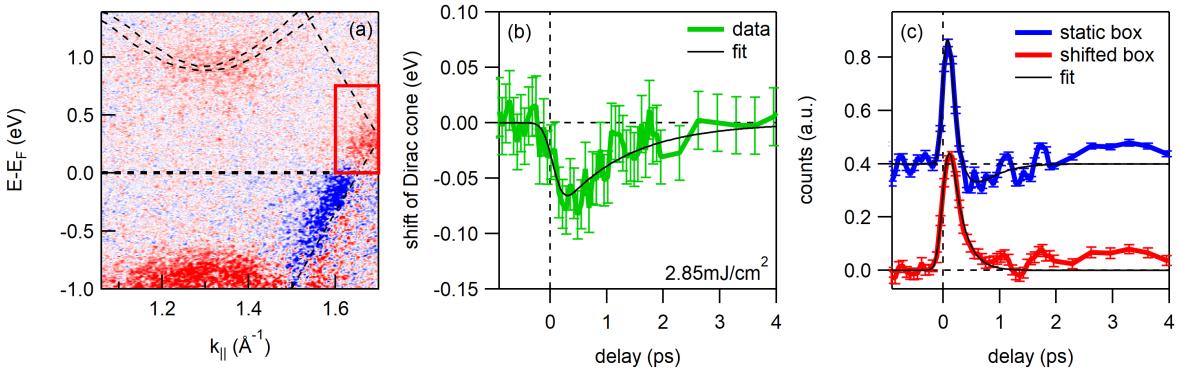


SFig. 4. Space charge effects: (a) Position of the Fermi level at negative delays as a function of pump fluence. The gray line is used to extrapolate the position of the Fermi level to zero pump fluence. The data point in brackets was obtained on an area of the sample where the WS_2 coverage was particularly low. The error bars are smaller than the symbol size. (b) Energy resolution for different pump fluences.

the electron mass. The binding energy E_B of the electrons inside the solid can be obtained from $E_B = \hbar\omega - E_{\text{kin}} - \phi$, where $\hbar\omega$ is the photon energy and ϕ is the work function of the analyzer, that — in the absence of space charge effects — determines the kinetic energy where the Fermi level appears in the measurement. The energy resolution of the tr-ARPES measurements was determined from Fermi-Dirac fits of the carrier distribution inside the Dirac cone at negative pump-probe delay. The fitting function consisted of the Fermi-Dirac distribution convolved with a Gaussian to account for the finite energy resolution. These fits revealed that the kinetic energy at which the Fermi level appeared in the measurement (SFig. 4a) as well as the energy resolution (SFig. 4b) depend on the applied pump fluence. The temporal resolution of the tr-ARPES measurements of 200 fs was deduced from the full width at half maximum (FWHM) of the width of the derivative of the rising edge of the photocurrent integrated over the area marked by the red box in Fig. 1f. The conversion of the measured photocurrent from (E_{kin}, θ) to $(E_B, k_{||})$ relies on an accurate value for the kinetic energy of the Fermi level that is commonly used as energy reference in ARPES data. To solve this issue, we extrapolated the kinetic energy of the Fermi level in SFig. 4a to zero pump fluence, yielding $E_{F,0} = 17.48\,\text{eV}$.

determination of population dynamics in Fig. 2a

The red box in Fig. 1b follows the transient down-shift of the Dirac cone. For that reason the sign change of the pump-probe signal observed in Fig. 3(2) in our previous work [5] is not observed in Fig. 2a. This effect is illustrated in Fig. 5.



SFig. 5. Extraction of gain signal in graphene: (a) Pump-induced changes of the photocurrent at the peak of the pump-probe signal (+0.2ps). The red box indicates the area of integration for the pump-probe traces shown in (c). Thin dashed lines represent the calculated band structures of graphene [6] and WS_2 [7]. (b) Transient shift of the Dirac cone as a function of pump-probe delay. (c) Graphene gain signal obtained by integrating the counts over the area marked by the red box in (a) as a function of pump-probe delay for the static box (blue) and for the box following the transient shift of the Dirac cone (red).

determination of transient band positions

In SFig. 3 we show examples for the fits used to determine the transient peak positions in Fig. 2. The EDCs in SFig. 3a were obtained by integrating the tr-ARPES data in the first row of Fig. 1 over a momentum interval of 0.05 $Å^{-1}$ centered at the green dashed line in Fig. 1a. A total of three Gaussians (one for the upper and lower WS_2 VB, respectively, and one for the background) was used to fit the spectra. While the lower and upper VB are nicely resolved at negative pump-probe delays, the broadening of the spectrum after photoexcitation makes a correct determination of the transient peak positions difficult. This difficulty could be avoided by constraining the parameters of the fit as follows. (1) No peak was allowed to gain spectral weight with respect to negative pump-probe delay. (2) Only the upper valence band (VB_A) was allowed to loose spectral weight as a result of photoexcitation. (3) No peak was allowed to narrow with respect to negative pump-probe delay. (4) The separation of the upper and lower WS_2 was fixed at 430 meV. (5) The position of the third Gaussian for the background was fixed at $E = -2.5$ eV. The EDCs in SFig. 3b were obtained by integrating the differential tr-ARPES data in the second row of Fig. 1 over a momentum interval of 0.1 $Å^{-1}$ centered at the green dashed line in Fig. 1a. These spectra were fitted with a single Gaussian for all pump-probe delays where the CB was found to be occupied. The integration

range for the MDCs in SFig. 3c was $\pm 0.06\,\mathrm{eV}$ around $E = -0.5\,\mathrm{eV}$. The MDCs were fitted with a single Lorentzian peak.

Fermi-Dirac fits of electron distribution inside Dirac cone

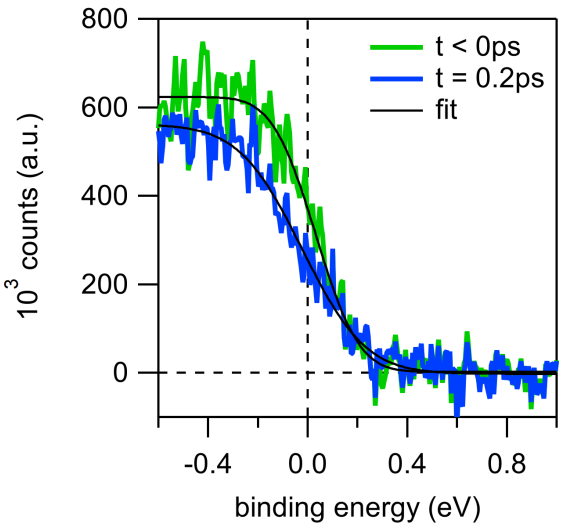
In SFig. 6 we provide an example for Fermi-Dirac fits of the transient electron distribution inside Dirac cone used to determine the electronic temperature and the chemical potential in SFig. 9. The fitting function consisted of a Fermi-Dirac distribution convolved with a Gaussian to account for the finite energy resolution. From these fits the number holes transferred from the WS_2 into the graphene layer was determined as follows. The transient shift of the chemical potential referenced with respect to the graphene Dirac point $\mu_{e(ED)}$ was calculated by subtracting the band shift in Fig. 2c from the chemical potential in SFig. 9b that is referenced with respect to the vacuum level. From $\mu_{e(ED)}(t)$ and $T_e(t)$ (SFig. 9a) we can then directly calculate the change of the total number of electrons in the graphene layer via

$$\Delta n_e(t) = \int_{-\infty}^{\infty} dE\, \rho(E)\, [f_{FD}(E, \mu(t), T(t)) - f_{FD}(E, \mu_0, T_0)]$$

where $\rho(E) = \frac{2A_c}{\pi} \frac{|E-ED|}{\hbar^2 v_F^2}$ is the density of states with the unit cell are $A_c = \frac{3\sqrt{3}a^2}{2}$ and the lattice constant $a = 1.42\,\text{Å}$. The equilibrium chemical potential is $\mu_0 = -0.3\,\mathrm{eV}$. The transient chemical potential is given by $\mu(t) = \mu_0 + \Delta\mu_{e(ED)}(t)$. The equilibrium temperature is $T_0 = T(t < 0\,\mathrm{ps}) = 300\,\mathrm{K}$. The number of transferred holes shown in SFig. 9c is then given by $\Delta n_h(t) = -\Delta n_e(t)$.

possible impact of $T_e \neq T_h \neq T_{graphene}$

Assuming that T_e, T_h, and $T_{graphene}$ at least exhibit a similar exponential decay, $T_e = T_h \neq T_{graphene}$ would only modify the y-axis scaling in Fig. 4c and d. As the y-axis scaling is in arbitrary units, this would not be important. If $T_e \neq T_h$ our conclusion that the hole transfer is faster than the electron transfer because of the smaller barrier height for holes might be too simple. In particular, our conclusion would break down, if $T_e \geq 2.4 T_h$, i.e. if the temperature difference became big enough to compensate for the difference in barrier height. However, this situation should not occur: As the WS_2 band structure is roughly



SFig. 6. Fermi-Dirac fits: EDCs obtained by integrating the counts over the momentum width of the yellow box in Fig. 1a at $t < 0$ fs (color) and $t = 200$ fs (color) together with Fermi-Dirac fits (black).

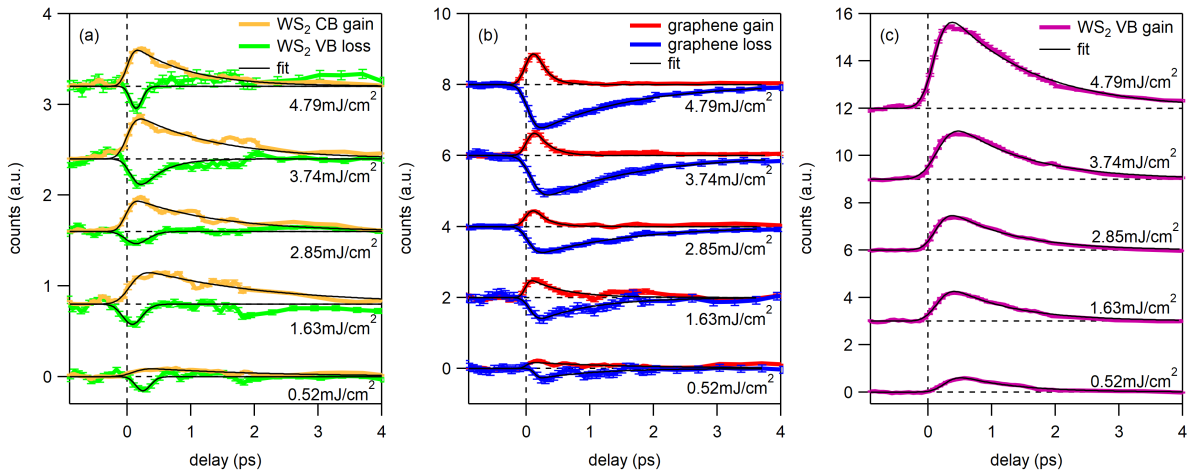
electron-hole symmetric at the K-point, we expect the initial electron and hole temperatures to be quite similar. This situation might change during charge separation. Because hole transfer is found to be much faster than electron transfer, we expect the remaining holes in WS_2 to cool much faster than the electrons. This means that $T_e = T_h$ at t = 0 will change to $T_e > T_h$ for later pump-probe delays. Therefore, $T_e \geq 2.4T_h$ is not expected to occur at any time. In conclusion, even if our assumptions $T_e = T_h$ and $T_e = T_{graphene}$ turned out not to be entirely applicable for all times, the main message of our study would not be affected.

additional tr-ARPES data for other pump fluences

The fluence dependent data corresponding to Fig. 2 is shown in SFigs. 7-8. SFig. 9 shows the fluence dependence of the hot carrier dynamics in graphene. The fluence dependence of various parameters at the peak of the pump-probe signal is shown in SFig. 10. A comparison of the fluence dependence of the gain above the equilibrium position of the upper WS_2 VB and the charging-induced shift of the VB is shown in SFig. 11.

fit function for exponential decay

The fitting function is given by



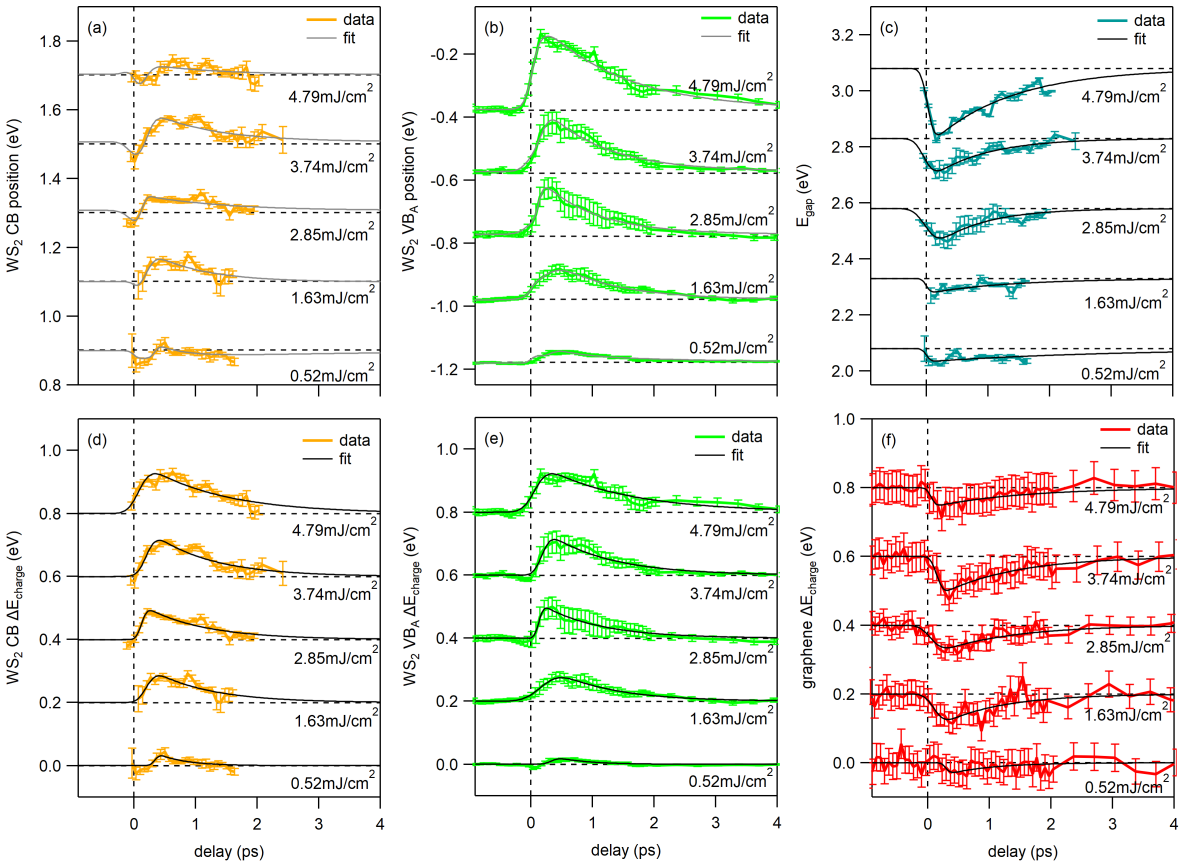
SFig. 7. Fluence dependence of the population dynamics: (a) Gain in the CB (orange) and loss in the VB (green) of WS_2. (b) Graphene dynamics: Gain above the Fermi level (red) and loss below the Fermi level (blue). (c) Gain above the equilibrium position of the upper WS_2 VB. Thin black lines are single exponential fits to the data.

$$f(t) = \frac{a}{2}\left(1 + \mathrm{erf}\left(\frac{(t-t_0)\tau - \sigma^2}{\sigma\,\tau}\right)\right)\exp\left(\frac{\sigma^2 - 2(t-t_0)\tau}{2\tau^2}\right) \tag{1}$$

where a is the amplitude of the pump-probe signal, σ is related to the full width at half maximum (FWHM) of the derivative of the rising edge via FWHM $= 2\sqrt{2\ln 2}\,\sigma$, t_0 is the middle of the rising edge, erf is the error function, and τ is the exponential lifetime. This fitting function is obtained by convolving the product of a step function and an exponential decay with a Gaussian to account for the finite rise time of the signal. In Fig 11 we show that the gain above the equilibrium position of the upper WS_2 VB from Fig. 2c and the WS_2 charging shift from Fig. 2b exhibit the same fluence dependence. Hence, the lifetime of the gain above the equilibrium position of the upper WS_2 VB is a measure for the lifetime of the charge separated state.

evidence for in-gap defect states

In SFig. 12 we show an energy distribution curve through the K-point of WS_2 in the energy region of the conduction band. At negative delay the WS_2 conduction band is unoccupied. At $t = 0.2\,\mathrm{ps}$ we observe a pronounced Gaussian peak at $E = 0.96\,\mathrm{eV}$ for the transiently populated WS_2 conduction band, as well as two other smaller Gaussians at $E = 0.29\,\mathrm{eV}$



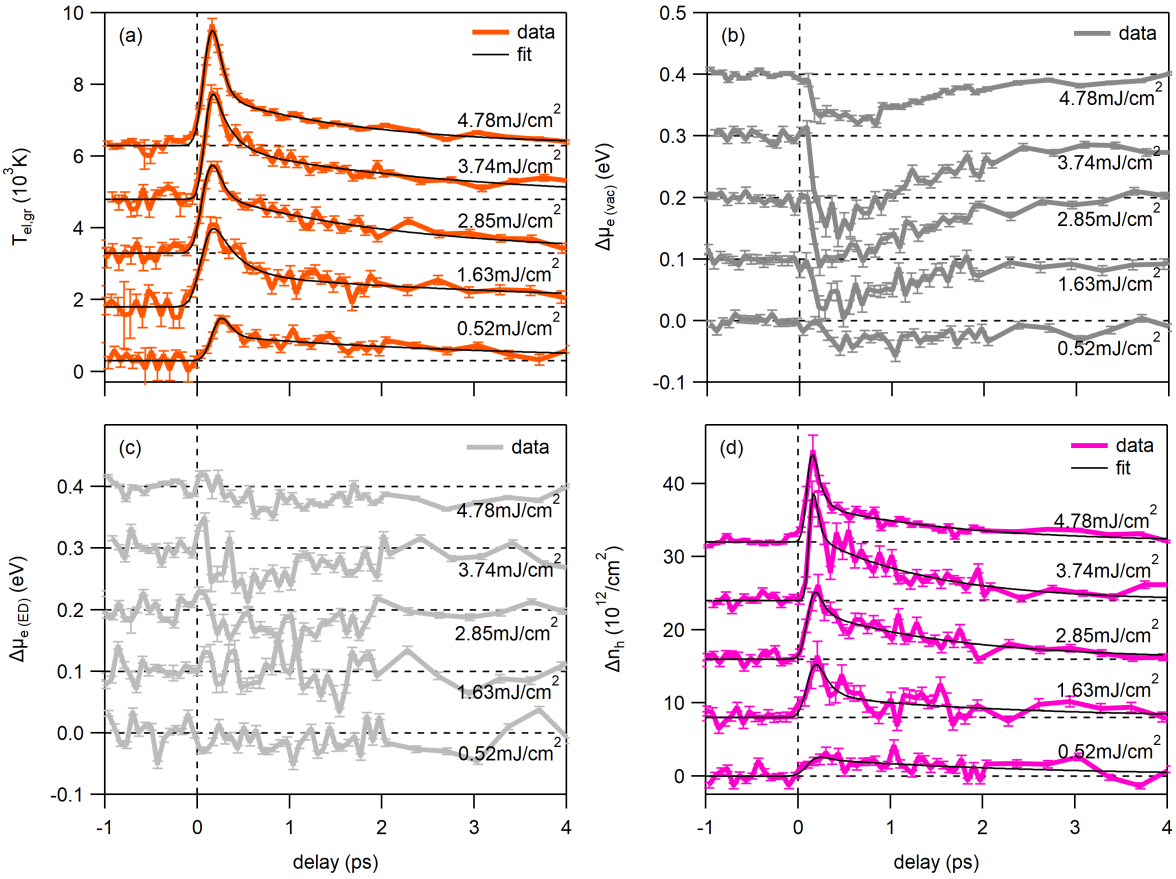
SFig. 8. Fluence dependence of the transient band positions: (a) Position of the WS_2 CB. (b) Position of the WS_2 VB. (c) Transient band gap of WS_2. (d) Shift of the WS_2 CB due to charging of the WS_2 layer. (e) Shift of the upper WS_2 VB due to charging of the WS_2 layer. (f) Shift of the Dirac cone.

and $E = 0.54\,\mathrm{eV}$ that we attribute to Sulfur vacancies in good agreement with scanning tunneling spectroscopy data in Ref. [8].

III COMPARISON WITH LITERATURE

In this section we compare our tr-ARPES data to other pump-fluence dependent experimental studies on similar vdW heterostructures [9, 10] and discuss whether the models proposed in literature [11–13] might also apply in our case.

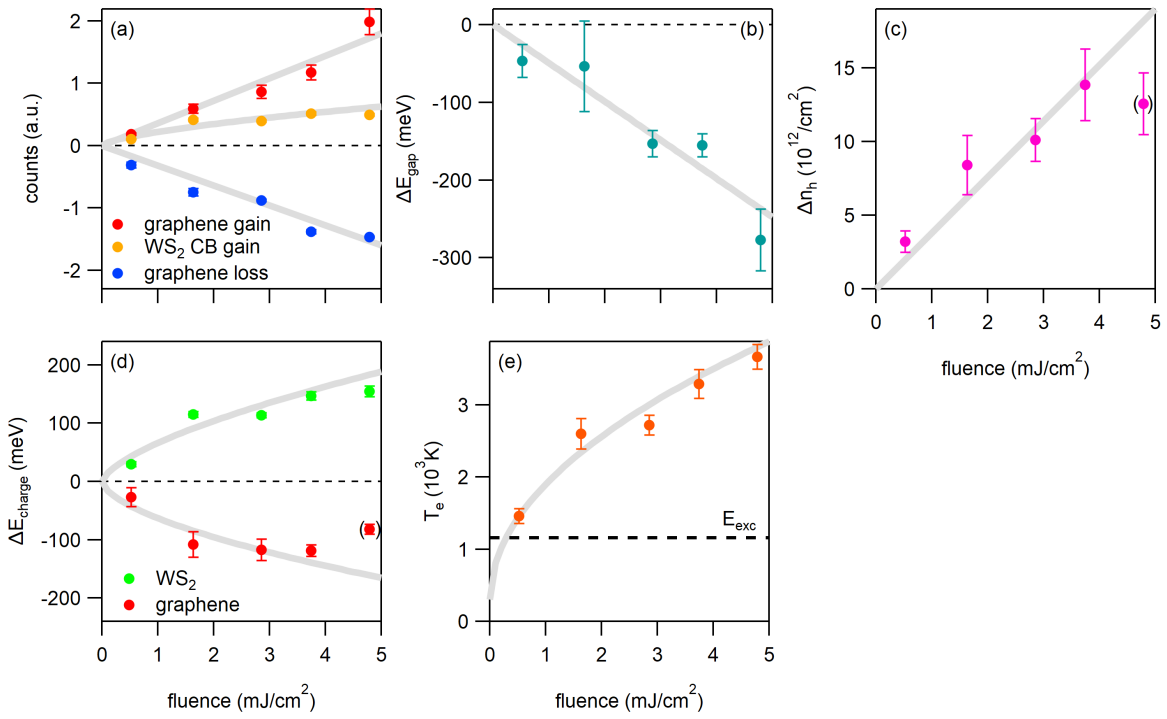
Reference [9] investigated an epitaxial MoS_2/graphene heterostructure on SiC(0001) with tr-ARPES. They observed a considerable band gap renormalization that increased with increasing hole density with a maximum value of $-450\,\mathrm{meV}$ at $n_h = 1.5 \times 10^{-12}\,\mathrm{cm}^{-1}$ with pump fluences in the $\mathrm{mJ/cm^2}$ regime. These values are of the same order of magnitude



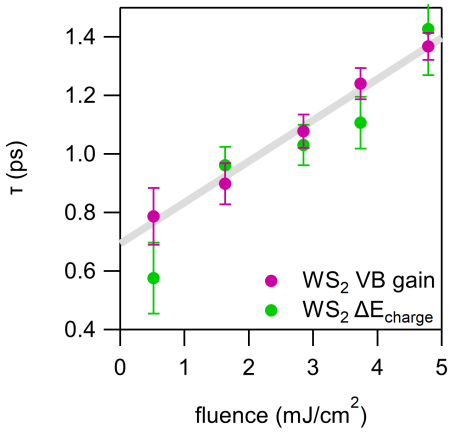
SFig. 9. Fluence dependence of hot carrier dynamics in graphene: (a) Electronic temperature, (b) chemical potential referenced with respect to the vacuum level, (c) chemical potential referenced with respect to the Dirac point, and (d) number of holes transferred from WS_2 to graphene as a function of pump-probe delay for different pump fluences.

as our data presented in SFig. 10b. Unlike the present study, Ref. [9] did not observe any indication for charge separation. We speculate that this might be related to the azimuthal alignment between the TMD and the graphene lattice which was 30° in [9] and which is 0° or 60° in the present study.

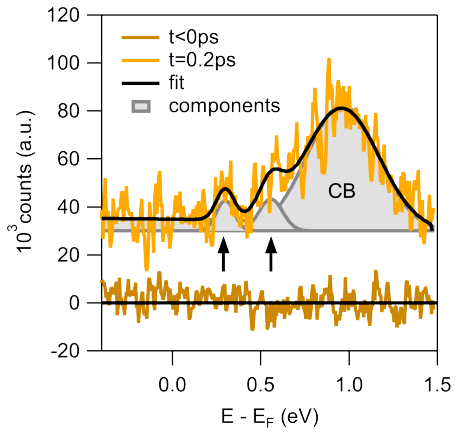
Reference [10] performed transient absorption measurements on CVD-grown WS_2/graphene heterostructures on SiO_2 but observed no systematic pump fluence dependence of the relaxation times with pump fluences in the $\mu J/cm^2$ regime. The authors proposed a model where the electric field across the interface that builds up due to charge separation caused by ultrafast hole transfer from WS_2 to graphene increases the transfer rate for the electrons which can be reconciled with our observation in Fig. 3b. However, this model cannot explain our observation that the hole transfer rate also increases with increasing fluence (see Fig. 3a).



SFig. 10. Fluence dependence of different parameters at the peak of the pump-probe signal ($t = 0.2$ ps): (a) Pump-induced change of photocurrent inside colored boxes from Fig. 1f. (b) Change of the WS_2 band gap from Fig. 2a. (c) Number of holes transferred into the graphene layer from SFig. 9c. (d) Charging-induced WS_2 and graphene band shifts from Figs. 2e and f. (e) Peak electronic temperature of Dirac carriers from SFig. 9a. Light gray lines are guides to the eye. The datapoints in brackets were obtained on areas of the sample with lower WS_2 coverage.



SFig. 11. Direct comparison of fluence dependence of gain above equilibrium position of upper WS_2 VB (purple) and WS_2 charging shift (green).



SFig. 12. Evidence for in-gap defect states. EDCs through the K-point of WS_2 in the conduction band region at negative delay and at the peak of the pump-probe signal at $t = 0.2$ ps. Gray lines indicate the Gaussian peaks used to fit the spectrum at $t = 0.2$ ps. The two Gaussians marked with black arrows are attributed to in-gap defect states due to Sulfur vacancies.

References [11–13] proposed a coherent phonon-driven charge transfer mechanism for ultrafast charge separation in vdW heterostructures with type II band alignment. The frequency of the associated coherent oscillations, however, is too high to be resolved given the limited temporal resolution of 200 fs in the present study.

IV MICROSCOPIC MODEL OF CHARGE TRANSFER

In the second quantization formalism, the Hamilton operator describing electron tunneling from one layer (l) to another ($\bar{l}$) is

$$H_T = \sum_{l\lambda \boldsymbol{k}\boldsymbol{q}} T^{\lambda l\bar{l}}_{\boldsymbol{k}\boldsymbol{q}} a^\dagger_{\lambda,\bar{l},\boldsymbol{k}+\boldsymbol{q}} a_{\lambda,l,\boldsymbol{k}}, \tag{2}$$

where $\lambda = v, c$ is the valence/conduction band, $\boldsymbol{k}$ the momentum of the initial state and $\boldsymbol{q}$ the momentum transfer of the process. The tunneling matrix element reads $T^{\lambda l\bar{l}}_{\boldsymbol{k}\boldsymbol{q}} = \langle \bar{l}, \lambda, \boldsymbol{k}+\boldsymbol{q}|V_T|l, \lambda, \boldsymbol{k}\rangle$ with the tunneling potential $V_T = V_l + V_{\bar{l}}$ being the sum of the potential of each layer. In order to find an expression for $T^{\lambda,\mathrm{WS_2}\to\mathrm{g}}_{\boldsymbol{k}\boldsymbol{q}}$, we use a tight-binding approach to describe the electronic wavefunctions in graphene and WS_2 [14]

$$\psi^{\mathrm{g}}_{\lambda\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N_{\mathrm{g}}}} \sum_{j=A,B} c^j_{\lambda\boldsymbol{k}} \sum_{\boldsymbol{R}_j} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{R}_j} \phi^{\mathrm{g}}_\lambda(\boldsymbol{r}-\boldsymbol{R}_j), \tag{3}$$

$$\psi^{\mathrm{WS_2}}_{\lambda\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N_{\mathrm{WS_2}}}} \sum_{\boldsymbol{R}_{\mathrm{W}}} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{R}_{\mathrm{W}}} \phi^{\mathrm{WS_2}}_{\lambda}(\boldsymbol{r}-\boldsymbol{R}_{\mathrm{W}}). \tag{4}$$

Here N_l is the number of unit cells in layer l, $c^j_{\lambda\boldsymbol{k}}$ are graphene's tight-binding coefficients, $\boldsymbol{R}_j$ are the atomic positions, and $\phi^l_\lambda(\boldsymbol{r})$ is the linear combination of atomic orbitals contributing to the band λ. Graphene's tight binding coefficients around the $\boldsymbol{K}$ point read $c^A_{\lambda\boldsymbol{k}} = \frac{1}{\sqrt{2}}$ and $c^B_{\lambda\boldsymbol{k}} = \frac{1}{\sqrt{2}}\sigma_\lambda \mathrm{e}^{i\theta_{\boldsymbol{k}-\boldsymbol{K}_{\mathrm{g}}}}$, where $\theta_{\boldsymbol{k}-\boldsymbol{K}_{\mathrm{g}}}$ is the angle of $\boldsymbol{k}$ with respect to graphene's $\boldsymbol{K}$ point. While for graphene the wavefunction has contributions from A and B lattice points, in semiconducting TMDs the main orbital contributions come from the metal atoms [15, 16]. Introducing the respective electronic wavefunctions, the tunneling matrix element reads

$$\begin{aligned} T^{\lambda,\mathrm{WS_2}\to\mathrm{g}}_{\boldsymbol{k}\boldsymbol{q}} &= \int d^3\boldsymbol{r} \left[\psi^{\mathrm{g}}_{\lambda\boldsymbol{k}+\boldsymbol{q}}(\boldsymbol{r})\right]^* V_T(\boldsymbol{r}) \psi^{\mathrm{WS_2}}_{\lambda\boldsymbol{k}}(\boldsymbol{r}) \\ &= \frac{1}{\sqrt{N_g N_{\mathrm{WS_2}}}} \sum_{j\boldsymbol{R}_j\boldsymbol{R}_{\mathrm{W}}} \left(c^j_{\lambda\boldsymbol{k}+\boldsymbol{q}}\right)^* \mathrm{e}^{-i(\boldsymbol{k}+\boldsymbol{q})\cdot\boldsymbol{R}_j} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{R}_{\mathrm{W}}} \\ &\quad \times \int d^3\boldsymbol{r} \left[\phi^{\mathrm{g}}_\lambda(\boldsymbol{r}-\boldsymbol{R}_j)\right]^* V_T(\boldsymbol{r}) \phi^{\mathrm{WS_2}}_\lambda(\boldsymbol{r}-\boldsymbol{R}_{\mathrm{W}}) \end{aligned} \tag{5}$$

Using $V_T = V_l + V_{\bar{l}}$ and the periodicity of V_l in the lattice of l and writing the orbital wave functions in terms of their in-plane Fourier transform, $\phi(\boldsymbol{r}) = \frac{1}{A}\sum_{\boldsymbol{p}} \mathrm{e}^{i\boldsymbol{p}\cdot\boldsymbol{r}_\parallel}\phi_{\boldsymbol{p}}(z)$, we can define a tunneling parameter $h_{\lambda\boldsymbol{p}}$ that contains the overlap between the orbital wave functions and the tunneling potential. The tunneling matrix element now reads

$$T^{\lambda,\mathrm{WS_2}\to\mathrm{g}}_{\boldsymbol{k}\boldsymbol{q}} = \frac{1}{N_g N_{\mathrm{WS_2}}} \sum_{j\boldsymbol{R}_j\boldsymbol{R}_{\mathrm{W}}\boldsymbol{p}} \left(c^j_{\lambda\boldsymbol{k}+\boldsymbol{q}}\right)^* \mathrm{e}^{-i(\boldsymbol{k}+\boldsymbol{q}-\boldsymbol{p})\cdot\boldsymbol{R}_j} \mathrm{e}^{i(\boldsymbol{k}-\boldsymbol{p})\cdot\boldsymbol{R}_{\mathrm{W}}} h_{\lambda\boldsymbol{p}}, \tag{6}$$

with the tunneling parameter

$$h_{\lambda\boldsymbol{p}} = t^{\mathrm{WS_2,g}}_{\lambda\boldsymbol{p}} + \left(t^{\mathrm{g,WS_2}}_{\lambda\boldsymbol{p}}\right)^*, \qquad t^{l\bar{l}}_{\lambda\boldsymbol{p}} = \frac{1}{\sqrt{N_g N_{\mathrm{WS_2}}}} \int d^3\boldsymbol{r} \mathrm{e}^{i\boldsymbol{p}\cdot\boldsymbol{r}_\parallel} \left[\phi^{\bar{l}}_{\lambda\boldsymbol{p}}(z)\right]^* V_l(\boldsymbol{r}) \phi^l_\lambda(\boldsymbol{r}) \tag{7}$$

We perform the summation over the lattice points of each layer using $\frac{1}{N_l}\sum_{\boldsymbol{R}_{lj}} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{R}_{lj}} = \sum_{G_l} \mathrm{e}^{i\boldsymbol{G}_l\cdot\boldsymbol{\delta}_{jl}}\delta_{\boldsymbol{k},\boldsymbol{G}_l}$ and determine the allowed momentum exchange $\delta_{\boldsymbol{q},\boldsymbol{G}_{\mathrm{g}}-\boldsymbol{G}_{\mathrm{WS_2}}}$, where $\boldsymbol{G}_l$ is the reciprocal lattice vector of the layer l. The lattice offsets $\boldsymbol{\delta}_{jl}$ are defined as $\boldsymbol{\delta}_A = -\boldsymbol{\delta}_{\mathrm{W}} = \frac{1}{2}\boldsymbol{R}^0_A$ and $\boldsymbol{\delta}_A = \frac{1}{2}\boldsymbol{R}^0_A + \boldsymbol{\tau}_{AB} = \frac{1}{2}\boldsymbol{R}^0_B$, with $\boldsymbol{\tau}_{AB}$ being a vector from an A lattice point to a nearest

neighbor B. This yields

$$T_{\boldsymbol{kq}}^{\lambda,\mathrm{WS_2}\to\mathrm{g}} = \sum_{j\boldsymbol{G}_\mathrm{g}\boldsymbol{G}_{\mathrm{WS_2}}} \left(c_{\lambda,\boldsymbol{k}+\boldsymbol{q}}^{j}\right)^* \mathrm{e}^{i\frac{1}{2}(\boldsymbol{G}_\mathrm{g}\cdot\boldsymbol{R}_j^0+\boldsymbol{G}_{\mathrm{WS_2}}\cdot\boldsymbol{R}_A^0)} h_{\lambda,\boldsymbol{k}+\boldsymbol{G}_{\mathrm{WS_2}}} \delta_{\boldsymbol{q},\boldsymbol{G}_\mathrm{g}-\boldsymbol{G}_{\mathrm{WS_2}}}. \tag{8}$$

In a practical scenario, the initial momentum $\boldsymbol{k}$ will lie in the vicinity of $\boldsymbol{K}_{\mathrm{WS_2}}$. Assuming that the tunneling parameter decays quickly with momentum, the significant reciprocal lattice vectors will be those connecting two K points of the Brillouin zone, i.e. $\boldsymbol{K}_l + \boldsymbol{G}_l = C_3^n \boldsymbol{K}_l$. Hence we can reduce the allowed scattering transitions to $\delta_{\boldsymbol{q},C_3^n\Delta K-\Delta K}$. Finally, the rotational symmetry of the orbitals yields $h_{\lambda,C_3^n K_{\mathrm{WS_2}}} = h_{\lambda,K_{\mathrm{WS_2}}}\mathrm{e}^{i\varphi_n}$, with $\varphi_{\lambda,n} = \frac{2\pi}{3}n(M_\lambda^{\mathrm{WS_2}} - M_\lambda^\mathrm{g})$, where M_λ^l is the rotational quantum number. Defining the Moiré phase $\Phi_{j,n} = \frac{1}{2}(\boldsymbol{G}_\mathrm{g}\cdot\boldsymbol{R}_j^0 + \boldsymbol{G}_{\mathrm{WS_2}}\cdot\boldsymbol{R}_A^0)$, the tunneling matrix element now reads

$$T_{\boldsymbol{kq}}^{\lambda,\mathrm{WS_2}\to\mathrm{g}} = \sum_{j,n=0}^{2} \left(c_{\lambda,\boldsymbol{k}+\boldsymbol{q}}^{j}\right)^* \mathrm{e}^{i\Phi_{j,n}}\mathrm{e}^{i\varphi_{\lambda,n}} h_{\lambda,\boldsymbol{K}_{\mathrm{WS_2}}} \delta_{\boldsymbol{q},C_3^n\Delta K-\Delta K}. \tag{9}$$

Now we choose the offset $\boldsymbol{R}_A^0 = 0$, resulting in $\Phi_{A,n} = 0$, $\Phi_{B,n} = -\frac{2\pi}{3}n$. Introducing graphene's tight-binding coefficients, we obtain

$$T_{\boldsymbol{kq}}^{\lambda,\mathrm{WS_2}\to\mathrm{g}} = \sum_{n=0}^{2} \left(1 + \sigma_\lambda \mathrm{e}^{i\theta_{\boldsymbol{k}+\boldsymbol{q}-\boldsymbol{K}_\mathrm{g}}}\mathrm{e}^{-i\frac{2\pi}{3}n}\right)\mathrm{e}^{i\varphi_{\lambda,n}} h_{\lambda,\boldsymbol{K}_{\mathrm{WS_2}}} \delta_{\boldsymbol{q},C_3^n\Delta K-\Delta K}. \tag{10}$$

Finally, we are interested in the squared absolute value of the tunneling matrix element. Note that the condition $\boldsymbol{q} = C_3^n\Delta K - \Delta K$ cannot be fulfilled by more than one n-value simultaneously. Hence the phase $\varphi_{\lambda,n}$ will in fact be a global phase which cancels out when computing the absolute value. Thus, we can write our final expression for the tunneling matrix element as

$$|T_{\boldsymbol{kq}}^{\lambda,\mathrm{WS_2}\to\mathrm{g}}|^2 = \sum_{n=0}^{2} |h_{\lambda,\boldsymbol{K}_{\mathrm{WS_2}}}|^2 \left[1 + \sigma_\lambda \cos\left(\theta_{\boldsymbol{k}-\boldsymbol{K}_\mathrm{g}+\boldsymbol{q}} - \frac{2\pi}{3}n\right)\right] \delta_{\boldsymbol{q},C_3^n\Delta K-\Delta K}. \tag{11}$$

Here $\sigma_\lambda = \pm 1$ for the conduction (+) and the valence (−) bands, $\theta_{\boldsymbol{k}-\boldsymbol{K}_\mathrm{g}+\boldsymbol{q}}$ is the angle of $\boldsymbol{k}+\boldsymbol{q}$ with respect to the closest graphene $\boldsymbol{K}_\mathrm{g}$ point, $\Delta K = K_\mathrm{g} - K_{\mathrm{WS_2}}$ is the momentum difference between graphene's and $\mathrm{WS_2}$'s K points, C_3^n is a $2\pi/3$ rotation operator, and h_λ contains the overlap of the wavefunctions with the tunneling potential. From this expression we obtain

the behaviour shown in Fig. 5a-b, i.e. the tunneling is efficient for holes but suppressed for electrons. This effect, together with the cosine dependence, is a manifestation of graphene's pseudospin.

In order to calculate tunneling rates, we insert the tunneling Hamilton operator in the Heisenberg's equation together with the carrier occupation $\rho^l_{\lambda\boldsymbol{k}}$ in the density matrix formalism [17]. Within a second-order Born-Markov approximation [18], we find the following Boltzmann-like equation:

$$\dot{\rho}^l_{\lambda\boldsymbol{k}} = \frac{2\pi}{\hbar}\sum_{\boldsymbol{q}} |T^{\lambda l\bar{l}}_{\boldsymbol{kq}}|^2 \left[\rho^{\bar{l}}_{\lambda,\boldsymbol{k}+\boldsymbol{q}}(1-\rho^l_{\lambda,\boldsymbol{k}}) - \rho^l_{\lambda,\boldsymbol{k}}(1-\rho^{\bar{l}}_{\lambda,\boldsymbol{k}+\boldsymbol{q}})\right]\delta(\varepsilon^l_{\lambda,\boldsymbol{k}} - \varepsilon^{\bar{l}}_{\lambda,\boldsymbol{k}+\boldsymbol{q}}). \tag{12}$$

We approximate the carrier distribution in the conduction band in WS_2 with a Boltzmann distribution. Integrating over momentum leads to a rate equation for the carrier density. We thus find that the carrier density that tunnels from WS_2 to graphene follows

$$\dot{n}^{\mathrm{WS_2}}_{\lambda}\big|_{\mathrm{WS_2}\to\mathrm{g}} = -\tau^{-1}_{\lambda} n^{\mathrm{WS_2}}_{\lambda}, \tag{13}$$

where

$$\tau^{-1}_{\lambda} = \frac{2\pi}{\hbar A n^{\mathrm{WS_2}}_{\lambda}}\sum_{\boldsymbol{kq}} |T^{\lambda l\bar{l}}_{\boldsymbol{kq}}|^2 \rho^{\mathrm{WS_2}}_{\lambda,\boldsymbol{k}}(1-\rho^{\mathrm{g}}_{\lambda,\boldsymbol{k}+\boldsymbol{q}})\delta(\varepsilon^{\mathrm{WS_2}}_{\lambda,\boldsymbol{k}} - \varepsilon^{\mathrm{g}}_{\lambda,\boldsymbol{k}+\boldsymbol{q}}) \tag{14}$$

* razvan.krause@ur.de

† isabella.gierz@ur.de

[1] K. V. Emtsev, A. Bostwick, K. Horn, J. Jobst, G. L. Kellogg, L. Ley, J. L. McChesney, T. Ohta, S. A. Reshanov, J. Röhrl, E. Rotenberg, A. K. Schmid, D. Waldmann, H. B. Weber, and T. Seyller, Towards wafer-size graphene layers by atmospheric pressure graphitization of silicon carbide, Nature Materials **8**, 203 (2009).

[2] C. Riedl, C. Coletti, T. Iwasaki, A. A. Zakharov, and U. Starke, Quasi-free-standing epitaxial graphene on SiC obtained by hydrogen intercalation, Physical Review Letters **103**, 246804 (2009).

[3] S. Forti, A. Rossi, H. Büch, T. Cavallucci, F. Bisio, A. Sala, T. O. Menteş, A. Locatelli, M. Magnozzi, M. Canepa, K. Müller, S. Link, U. Starke, V. Tozzini, and C. Coletti, Electronic

properties of single-layer tungsten disulfide on epitaxial graphene on silicon carbide, Nanoscale **9**, 16412 (2017).

[4] F. Fabbri, F. Dinelli, S. Forti, L. Sementa, S. Pace, G. Piccinini, A. Fortunelli, C. Coletti, and P. Pingue, Edge defects promoted oxidation of monolayer WS_2 synthesized on epitaxial graphene, The Journal of Physical Chemistry C **124**, 9035 (2020).

[5] S. Aeschlimann, A. Rossi, M. Chávez-Cervantes, R. Krause, B. Arnoldi, B. Stadtmüller, M. Aeschlimann, S. Forti, F. Fabbri, C. Coletti, and I. Gierz-Pehla, Direct evidence for efficient ultrafast charge separation in epitaxial WS_2/graphene heterostructures, Science Advances **6**, eaay0761 (2020).

[6] P. R. Wallace, The band theory of graphite, Physical review **71**, 622 (1947).

[7] H. Zeng, G.-B. Liu, J. Dai, Y. Yan, B. Zhu, R. He, L. Xie, S. Xu, X. Chen, W. Yao, and X. Cui, Optical signature of symmetry variations and spin-valley coupling in atomically thin tungsten dichalcogenides, Scientific reports **3**, 1608 (2013).

[8] B. Schuler, D. Y. Qiu, S. Refaely-Abramson, C. Kastl, C. T. Chen, S. Barja, R. J. Koch, D. F. Ogletree, S. Aloni, A. M. Schwartzberg, J. B. Neaton, S. G. Louie, and A. Weber-Bargioni, Large spin-orbit splitting of deep in-gap defect states of engineered sulfur vacancies in monolayer WS_2, Physical Review Letters **123**, 076801 (2019).

[9] S. Ulstrup, A. G. Cabo, J. A. Miwa, J. M. Riley, S. S. Grønborg, J. C. Johannsen, C. Cacho, O. Alexander, R. T. Chapman, E. Springate, M. Bianchi, M. Dendzik, J. V. Lauritsen, P. D. C. King, and P. Hofmann, Ultrafast band structure control of a two-dimensional heterostructure, ACS nano **10**, 6315 (2016).

[10] Z. Song, H. Zhu, W. Shi, D. Sun, and S. Ruan, Ultrafast charge transfer in graphene-WS_2 van der Waals heterostructures, Optik **174**, 62 (2018).

[11] R. Long and O. V. Prezhdo, Quantum coherence facilitates efficient charge separation at a MoS_2/$MoSe_2$ van der Waals junction, Nano letters **16**, 1996 (2016).

[12] H. Wang, J. Bang, Y. Sun, L. Liang, D. West, V. Meunier, and S. Zhang, The role of collective motion in the ultrafast charge transfer in van der Waals heterostructures, Nature communications **7**, 1 (2016).

[13] Q. Zheng, W. A. Saidi, Y. Xie, Z. Lan, O. V. Prezhdo, H. Petek, and J. Zhao, Phonon-assisted ultrafast charge transfer at van der Waals heterostructure interface, Nano letters **17**, 6435 (2017).

[14] M. Selig, E. Malic, K. J. Ahn, N. Koch, and A. Knorr, Theory of optically induced Förster coupling in van der Waals coupled heterostructures, Physical Review B **99**, 035420 (2019).

[15] A. Kormányos, G. Burkard, M. Gmitra, J. Fabian, V. Zólyomi, N. D. Drummond, and V. Fal'ko, k · p theory for two-dimensional transition metal dichalcogenide semiconductors, 2D Materials **2**, 022001 (2015).

[16] Y. Wang, Z. Wang, W. Yao, G.-B. Liu, and H. Yu, Interlayer coupling in commensurate and incommensurate bilayer structures of transition-metal dichalcogenides, Physical Review B **95**, 115429 (2017).

[17] E. Malic, T. Winzer, E. Bobkin, and A. Knorr, Microscopic theory of absorption and ultrafast many-particle kinetics in graphene, Physical Review B **84**, 205406 (2011).

[18] M. Kira and S. W. Koch, Many-body correlations and excitonic effects in semiconductor spectroscopy, Progress in Quantum Electronics **30**, 155 (2006).