Observation of ultrafast interfacial Meitner-Auger energy transfer in a van der Waals heterostructure

Shuo Dong[1*], Samuel Beaulieu[1,2], Malte Selig[3], Philipp Rosenzweig[4], Dominik Christiansen[3], Tommaso Pincelli[1], Maciej Dendzik[1,5], Jonas D. Ziegler[6], Julian Maklar[1], R. Patrick Xian[1,7], Alexander Neef[1], Avaise Mohammed[4], Armin Schulz[4], Mona Stadler[8], Michael Jetter[8], Peter Michler[8], Takashi Taniguchi[9], Kenji Watanabe[10], Hidenori Takagi[4,11,12], Ulrich Starke[4], Alexey Chernikov[6,13], Martin Wolf[1], Hiro Nakamura[4,14], Andreas Knorr[3], Laurenz Rettig[1*] & Ralph Ernstorfer[1,15*]

[1]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany*

[2]*Université de Bordeaux - CNRS - CEA, CELIA, UMR5107, F33405, Talence, France*

[3]*Nichtlineare Optik und Quantenelektronik, Institut für Theoretische Physik, Technische Universität Berlin, 10623 Berlin, Germany*

[4]*Max Planck Institute for Solid State Research, 70569 Stuttgart, Germany*

[5]*Department of Applied Physics, KTH Royal Institute of Technology, Hannes Alfvéns väg 12, 114 19 Stockholm, Sweden*

[6]*Department of Physics, University of Regensburg, Regensburg D-93053, Germany*

[7]*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*

[8]*Institute of Semiconductor Optics and Functional Interfaces, Research Center SCoPE and IQST, University of Stuttgart, 70569 Stuttgart, Germany*

[9]*International Center for Materials Nanoarchitectonics, National Institute for Materials Science,*

1-1 Namiki, Tsukuba 305-0044, Japan

[10]*Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan*

[11]*Department of Physics, University of Tokyo, 113-0033 Tokyo, Japan*

[12]*Institute for Functional Matter and Quantum Technologies, University of Stuttgart, 70569 Stuttgart, Germany*

[13]*Institute for Applied Physics, Dresden University of Technology, Dresden, 01187, Germany*

[14]*Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA*

[15]*Institut für Optik und Atomare Physik, Technische Universität Berlin, 10623 Berlin, Germany*

*To whom correspondence should be addressed;

E-mail: dong@fhi-berlin.mpg.de, rettig@fhi-berlin.mpg.de, ernstorfer@fhi-berlin.mpg.de.

Atomically thin layered van der Waals heterostructures feature exotic and emergent optoelectronic properties. With growing interest in these novel quantum materials, the microscopic understanding of fundamental interfacial coupling mechanisms is of capital importance. Here, using multidimensional photoemission spectroscopy, we provide a layer- and momentum-resolved view on ultrafast interlayer electron and energy transfer in a monolayer-WSe_2/graphene heterostructure. Depending on the nature of the optically prepared state, we find the different dominating transfer mechanisms: while electron injection from graphene to WSe_2 is observed after photoexcitation of quasi-free hot carriers in the graphene layer, we establish an interfacial Meitner-Auger energy transfer process following the excitation of excitons in WSe_2. By analysing the time-energy-momentum distributions of excited-state carriers with a rate-equation model, we distinguish these two types of interfacial dynamics and identify the ultrafast conversion of excitons in WSe_2 to valence band transitions in graphene. Microscopic calculations find interfacial dipole-monopole coupling underlying the Meitner-Auger energy transfer to dominate over conventional Förster- and Dexter-type interactions, in agreement with the experimental observations. The energy transfer mechanism revealed here might enable new hot-carrier-based device concepts with van der Waals heterostructures.

The unique physical properties of atomically thin two-dimensional (2D) materials[1,2] and constantly improving fabrication methods[3,4] have lead to a great interest in novel quantum materials based on van der Waals (vdW) heterostructures[5]. By stacking 2D materials, vdW heterostructures inherit the properties from individual constituents, and exotic physical phenomena may emerge

due to the interfacial interaction[5–7]. An emblematic example is the emergence of superconductivity in twisted bilayer graphene when stacked at the so-called 'magic angle'[8]. As another example, interlayer excitons, which are spatially separated yet Coulomb-bound electron-hole pairs in semiconducting transition metal dichalcogenide (TMDC) heterostructures allow exceptional control of optoelectronic properties[9,10]. Out of the vdW heterostructure library, a basic optoelectronic building block is a monolayer (ML) semiconducting TMDC in contact with graphene[11]. This hybrid structure represents a model system as it combines the strong light-matter coupling of TMDCs and the high mobility of massless Dirac carriers of graphene[12]. The gapless electronic structure of graphene allows for harvesting low-energy photons, extending the spectral range covered by conventional photodetectors to the near-infrared wavelength, which is highly beneficial for photovoltaic applications[13].

Optoelectronic functionality in vdW heterostructures arises from careful design and control of optical transitions and interfacial transfer processes. Particularly, interfacial charge (ICT) and energy transfer (IET) are key processes which have triggered extensive experimental and theoretical efforts[14–17]. Using time-resolved optical spectroscopies, a strong reduction of the exciton lifetime and optically active charge-transfer excitations of TMDC/graphene heterostructures have been observed[18,19], suggesting strong interlayer coupling. Moreover, the efficiency of IET processes like Förster-type coupling (based on electronic dipole-dipole interaction) has recently been investigated theoretically, pointing out the importance of energy-momentum conservation between participating quasiparticles[14]. Therefore, a momentum resolved probe is required to directly monitor the dynamics and reveal the mechanism of interfacial transfer process in vdW heterostructures,

including those involving momentum-forbidden dark states.

Here, we use time- and angle-resolved photoemission spectroscopy (trARPES) to investigate ultrafast interlayer carrier interactions in an epitaxially grown ML-WSe_2/graphene heterostructure. Our trARPES setup combines a high-repetition-rate (500 kHz) femtosecond extreme ultraviolet (XUV) source [20] coupled to a time-of-flight momentum microscope[21] (see Methods). It allows the measurement of the four-dimensional (4D) photoemission intensity $I(E_{kin}, k_x, k_y, \Delta t)$, where E_{kin} is the outgoing photoelectron kinetic energy, k_x,k_y are the in-plane momenta and Δt is the pump-probe delay, as shown in Fig.1**a,b**. The probe photon energy of 21.7 eV allows accessing the entire Brillouin zone of the heterostructure and the variable pump wavelength allows us to photoexcite the heterostructure in a state-resolved manner. In the following, we present a time-, energy-, and momentum-resolved study on the excited-state dynamics in the heterostructure with two different pump photon energies: below the optical bandgap of WSe_2 (1.2 eV) and in resonance with its first excitonic transition (1.55 eV).

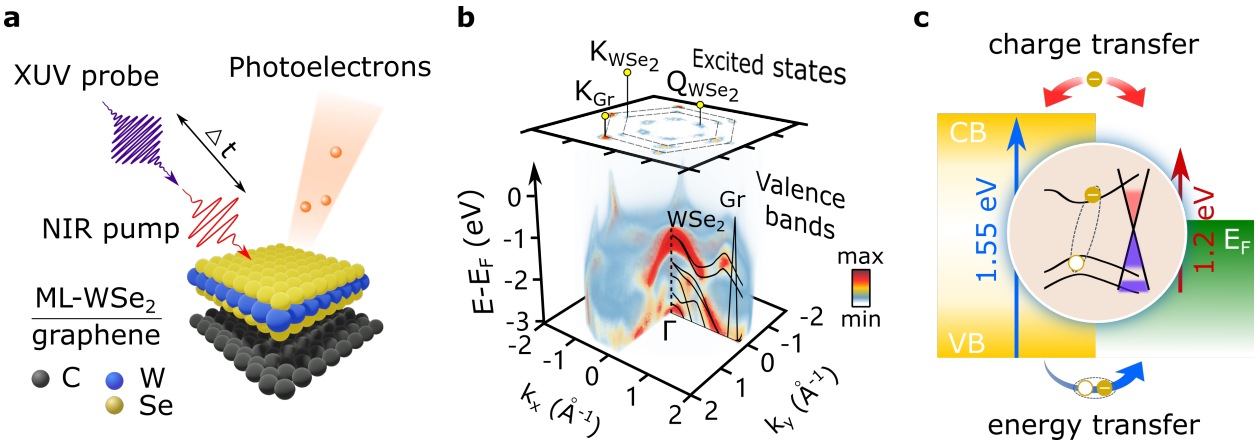


Fig. 1: Time- and angle-resolved photoemission measurement of interlayer charge and energy transfer in a ML-WSe_2/graphene heterostructure. **a**, Following the near-infrared pump, electrons are photoionized by the delayed XUV probe pulses and collected by a three-dimensional (3D) (E_{kin}, k_x, k_y) detector as a function of pump-probe delay Δt. **b**, The 3D snapshot of the 4D data, $I(E_{kin}, k_x, k_y, \Delta t = 0\ \mathrm{fs})$ presents the valence band structures from the Γ point to the Brillouin zone boundary of WSe_2, as well as the linearly dispersing graphene bands. The excited state population can be clearly mapped at the $\mathrm{K_{WSe_2}}$ and $\mathrm{Q_{WSe_2}}$ valleys, and the π^* band of graphene ($\mathrm{K_{Gr}}$). **c**, By changing the pump wavelength, we can selectively prepare different initial excited states: quasi-free carriers in graphene with the below-bandgap excitation (red arrow) or excitons in WSe_2 using excitation on the excitonic resonance (blue arrow).

Interlayer quasi-free carrier transfer

First, we photoexcite the heterostructure with the pump photon energy centred at $\hbar\omega_{pump}$=1.2 eV (pump pulse duration 200 fs FWHM), well below the optical bandgap of WSe_2[22]. The NIR-

pump/XUV-probe experiments were performed with a pump fluence of $F = 5.3\ \mathrm{mJ/cm^2}$ and at room temperature. Fig.2**a-d** show energy-resolved photoemission signals along the $\mathrm{K'-K}$ cut of the Brillouin zone, at selected time delays. The momentum distributions above E_F within the first 400 fs reveal that the excited states are localized in three different types of valleys: the Dirac cones of graphene at its K points ($\mathrm{K_{Gr}}$) and the K and Q valleys of WSe_2 ($\mathrm{K_{WSe_2}}$, $\mathrm{Q_{WSe_2}}$), as shown in Fig.2**e**. The $\mathrm{Q_{WSe_2}}$ valley localizes between the $\mathrm{K_{WSe_2}}$ valley and the Γ point. By performing energy-momentum integration in selected regions of interest (ROIs), we extracted excited-state dynamics within these three valleys (Fig.2**f**). Upon arrival of the pump pulses, the excited-state population rapidly builds up at $\mathrm{K_{Gr}}$ (black curve) and decays with a time scale of $\sim$ 200 fs. Strikingly, the conduction band minima (CBMs) at $\mathrm{K_{WSe_2}}$ (red curve) and $\mathrm{Q_{WSe_2}}$ valleys (green curve) are also being populated, however, with a delay of $\Delta t = 51 \pm 9$ fs (see SI) compared to the rise of hot-carrier population in graphene. Since the below-bandgap pump photon energy does not allow the direct photoexcitation of WSe_2, the delayed electron populations in the conduction bands arise through charge transfer from graphene to WSe_2.

These observations support the following picture of the underlying processes with a below-bandgap excitation: light is absorbed by graphene and populates unoccupied states at $E^{el}_{Gr} = E_D + \hbar\omega_{pump}/2$, leaving holes at $E^{h}_{Gr} = E_D - \hbar\omega_{pump}/2$ (Dirac energy $E_D > 0$ for a p-doped system or $E_D < 0$ for an n-doped system). The energy position of the Dirac point in our heterostructure is estimated to be ~ -0.1 eV below the Fermi level, obtained from the conical crossing[23,24] (see SI). The photoexcited carriers quickly reach a quasi-thermalized states in $\sim$ 10 fs[25] and could further increase their energy via intraband electron-electron scattering and interband Auger recom-

bination in few tens of femtoseconds[26,27]. Once electrons gained a sufficient amount of energy to overcome the energy barrier, they scatter to WSe_2 via a phonon-assisted tunneling process, filling the single-particle CBMs at $\mathrm{K}_{\mathrm{WSe_2}}$ and $\mathrm{Q}_{\mathrm{WSe_2}}$. This ICT mechanism is called *interlayer hot carrier injection*, and is schematically illustrated in Fig.2**g**. The excited electrons in WSe_2 may subsequently scatter back to graphene and relax down towards the Fermi energy (E_F). Based on the observed carrier dynamics, we performed microscopic calculations of the phonon-assisted interlayer tunneling process, allowing us to estimate the electronic wavefunction overlap between the involved conduction bands of WSe_2 and graphene to be around 4% (see SI for details).

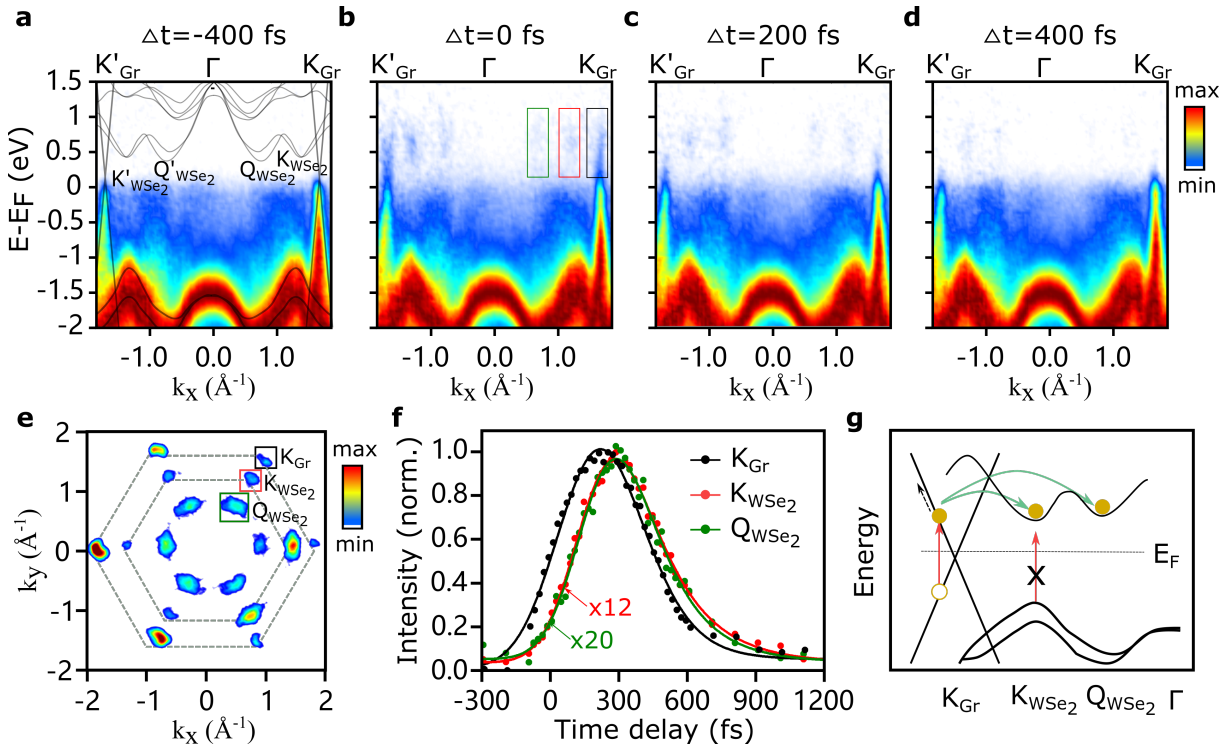


Fig. 2: Layer- and valley-resolved ultrafast dynamics upon below-bandgap pumping. **a-d**, Energy-momentum cuts of the photoemission signal along the $\mathrm{K'_{Gr}}$- $\mathrm{K'_{WSe_2}}$-Γ-$\mathrm{K_{WSe_2}}$-$\mathrm{K_{Gr}}$ high symmetry direction, at selected pump-probe time delays. **a**, The 2D spectrum at negative time delay reveals the equilibrium band structure of ML-WSe_2 as well as the linearly dispersing π band of graphene. The grey lines represent the DFT-calculated band structures (details in methods). Snapshots of the energy-momentum cuts at time delays of **b** $\Delta t = 0$ fs, **c** $\Delta t = 200$ fs and **d** $\Delta t = 400$ fs, respectively. **e**, Momentum map of the excited states (energy integrated for $E > E_\mathrm{F}$ and time integrated for the first 400 fs), showing the $\mathrm{K_{Gr}}$ points of graphene (black box) as well as the $\mathrm{K_{WSe_2}}$ and $\mathrm{Q_{WSe_2}}$ valleys (red and green boxes, respectively). The dashed grey lines represent the hexagonal Brillouin zones of both layers. **f**, Normalized population dynamics within the three

ROIs defined in **e**: $\mathrm{K_{WSe_2}}$ (red) and $\mathrm{Q_{WSe_2}}$ (green) are populated with a ∼50 fs delay with respect to $\mathrm{K_{Gr}}$ (black). **g**, Schematic of the ICT upon the below-bandgap excitation: the hot carriers within the graphene layer are transferred to the conduction bands of WSe_2 via hot electron injection.

Interlayer energy transfer: from excitons in WSe_2 to intraband transitions in graphene

Next, we select a pump photon energy of $\hbar\omega_{pump}$=1.55 eV (pump pulse duration: 35 fs FWHM, pump fluence: $F = 1.7\ \mathrm{mJ/cm^2}$), near-resonant to the A-excitonic transition of WSe_2. In this case, the pump photon energy allows both the WSe_2 and the graphene layer to be simultaneously photoexcited. One striking observation is that the energy distribution of excited carriers at the $\mathrm{K_{WSe_2}}$ valleys is centred at 0.63 eV (Fig.3**a**), ∼100 meV lower than with below-bandgap excitation (Fig. 3**b**), as apparent from the energy distribution curves (EDCs) (first 100 fs). As discussed above, with 1.2 eV excitation, the $\mathrm{K_{WSe_2}}$ valleys are filled with quasi-free electrons that have tunneled from the graphene layer. Therefore, this ∼ 100 meV energy difference is a direct photoemission signature of exciton formation, when resonantly pumping using 1.55 eV photons[28]: the bound electron-hole (el-h) pair reduces the quasi-free particle bandgap by the exciton binding energy. In addition to this excitonic feature, we also observe a transient shift of WSe_2 valence bands. In Fig. 3**d**, EDCs at $\mathrm{K_{WSe_2}}$ are shown at $\Delta t = 0$ fs (red) and $\Delta t = -200$ fs (black), in which the top two valence bands, VB1 and VB2, are fitted using Gaussian lineshape functions (see SI). The peak position of VB1 shifts towards the conduction band within the first 100 fs, transiently shrinking the electronic bandgap. This is due to the arrival of ICT-induced charge carriers from the graphene

layer. With resonantly pumping the A-exciton, the occurrence of ICT and injection of quasi-free carriers from graphene to WSe_2 is expected, similar to the case of below-bandgap excitation. This could lead to dynamical screening effect and the observed bandgap renormalization, as reported in highly-excited or doped ML TMDC materials[29–32]. As the magnitude of such a transient band gap renormalization has been shown to scale with the excited charge carrier density[31,33], we utilize the VB shift in the following as a measure of the ICT transferred carriers dynamics from graphene layer.

In addition to the excited-state dynamics in WSe_2, important insight can be drawn from the energy-momentum distribution of hot carriers in graphene. As shown in the early-time 2D differential spectrum $\Delta I(E, k, \Delta t = 0 \text{ fs})$ (Fig. 3**c**), obtained by subtracting the spectrum at negative time, hot carriers distribute in a broad energy range. The momentum-integrated spectrum along the linearly dispersing band in Fig. 3**e** clearly features the energy distribution of net electron gain (positive; red area) and loss (negative; blue area) following resonant photoexcitation. Remarkably, besides the modification of the distribution function near the Fermi level, we notice a strong negative peak at $E - E_F = -1.8$ eV. As noted earlier, for direct photoexcitation in graphene the photoexcited carriers are expected to be spread ± 0.77 eV ($\hbar\omega_{pump}/2$) around the Dirac point and quickly relax back to the Fermi level. Thus, this simple excitation mechanism cannot explain this peculiar feature in the valence band spectrum. The electron-electron scattering and Auger recombination could lead to a transient broadening of the momentum-space carrier distribution, but without any preferential energy localization[27,34,35]. Interlayer hot *hole* transfer can also be ruled out, as the top valence band of WSe_2 lies at $E - E_F = -1.0$ eV. It would require a multi-phonon absorption to

populate the hole-states localized deeply in the valence band, taking the typical phonon energy of $\sim$0.17 eV in graphene[36], a process of very low probability. However, the energy difference of deep-lying valence holes ($E-E_F = -1.8\,\mathrm{eV}$) and states near E_F ($E-E_F = -0.2\,\mathrm{eV}$) in graphene well matches the energy of the A-exciton in WSe_2 ($E_{ex} \sim 1.6\,\mathrm{eV}$). Combined with the fast depletion of exciton population shown in Fig.4**a** (black curve) extracted from the excited state of WSe_2 (ROI_1 in Fig.3**c**), this brings about the following scenario for the excitation of these carriers: annihilation of excitons in WSe_2 drives the intraband excitation of deep-lying valence electrons in graphene into empty hole states below the Dirac point. In more detail, this exciton energy transfer process, which we term Meitner-Auger energy transfer, considers recombination of excitons in WSe_2 with center-of-mass (COM) momentum $\mathbf{Q}$ and exciton energy E_{ex}. The photoexcitation prepares the required hot hole vacancy below E_F in graphene, thus enabling the intraband excitation. The momentum of the valence electron-hole pair k_{Gr} is determined by the Fermi velocity of the graphene bands and the transition energy E_{Gr}. This required momentum is provided by the optically pumped excitons which gain finite COM momenta during the population formation process via phonon-mediated dephasing and intravalley thermalization[37–40] (see the discussion in SI). The highly efficient IET of the excitons and intraband electron-hole pairs is thus possible under the conservation of energy and momentum, *i.e.*, $E_{ex} = E_{Gr}$ and $\mathbf{Q} = k_{Gr}$. In a similar trARPES study of a ML WS_2/graphene heterostructure, dominating interfacial charge transfer has been observed[16]. While the additional exciton energy transfer was not excluded, its relative efficiency might be reduced due to the larger COM momentum required at the larger A-exciton energy of WS_2 and the energy level alignment of these specific samples.

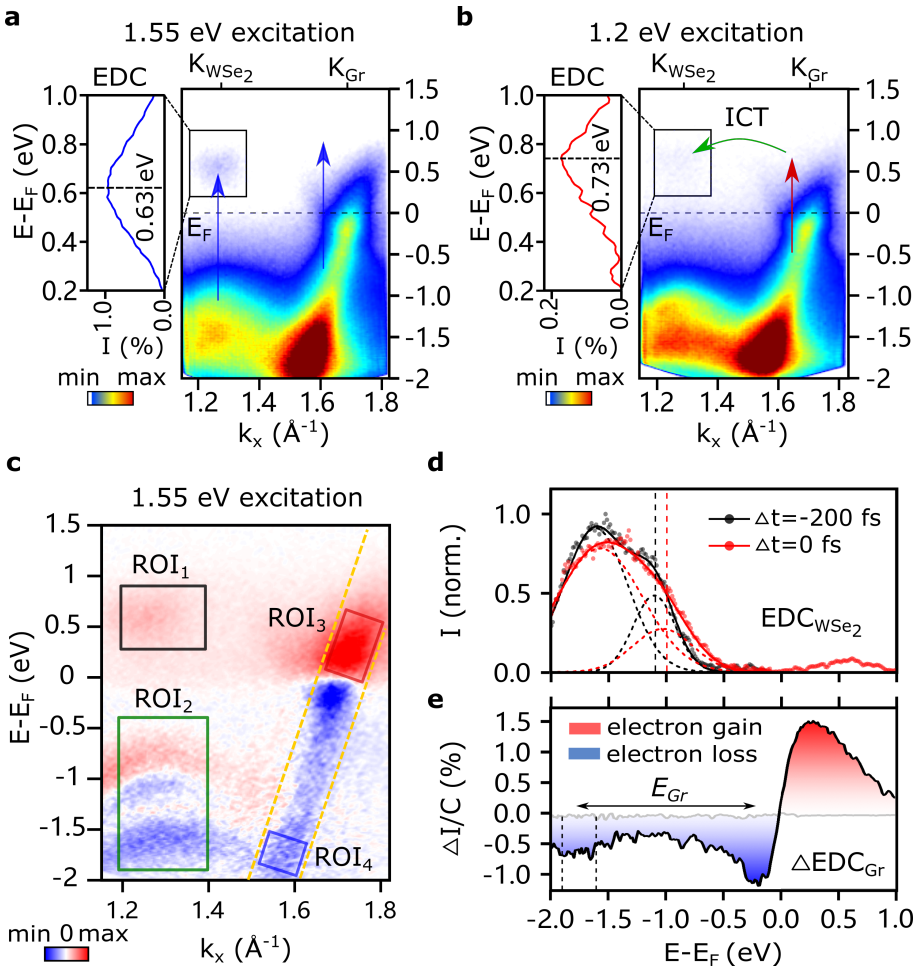


Fig. 3: Photoemission signatures of exciton formation and interfacial interactions. **a**, With the resonant A-exciton pump (1.55 eV), carriers within both the WSe_2 and the graphene layer are photoexcited (time integration of 100 fs). The energy of the excited-states carriers at K_{WSe_2} is 0.63 eV, shown in the EDC (left panel figure). **b**, With the below-bandgap excitation (1.2 eV), the local CBM of K_{WSe_2} is filled with ICT-induced electrons and centered at 0.73 eV. **c**, Differential energy-momentum cut with 1.55 eV pump at time zero, obtained by subtracting the negative time delay spectrum. **d**, The normalized EDC of K_{WSe_2} (momentum integration of 0.2 Å^{-1}) at $\Delta t = -200$ fs (black) and $\Delta t = 0$ fs

(red). The VBs are fitted with two Gaussian functions (dashed curves) and the positions of VB1 are indicated by the dash lines. **e**, The momentum-integrated spectrum of graphene Dirac bands (between the dashed yellow lines in **c**) shows the electron gain (positive, red area) and loss (negative, blue area) following photoexcitation. The intensity is normalized by the total electron count C obtained from negative time delay spectrum. Apart from the carriers accumulation near the E_F, the hole population forms another prominent peak around $E - E_F = -1.8$ eV, indicated between the dash lines.

In order to gain information on the time scales of the energy and charge transfer processes, next we analyze the dynamics of excited-state populations extracted from the ROIs shown in Fig. 3**c**, including the excited-state carriers in WSe_2 (ROI_1), VB1 shifting (ROI_2), hot electrons in graphene (ROI_3) and IET-driven deep valence band holes (ROI_4). The time trace of hot carriers in the CBM of WSe_2 (black curve in Fig.4**a**) contains two types of quasi-particles dynamics: the photo-generated excitons N_T^{ex} and the ICT-induced quasi-free electrons N_T^{el}. The decay of excitons excite the valence band electrons in graphene via IET with a transfer time of τ_{IET} (Fig. 4**f**). On the other hand, the arrival of ICT-induced electrons transiently shift VB1 of WSe_2 (green curve in Fig.4**a**) which therefore represents the dynamics of N_T^{el} as discussed before. The VB1 shifting shows a time delay of ~ 65 fs compared to the CB signal, evidencing the occurrence of interlayer hot electron injection after photoexcitation. The population of N_T^{el} subsequently relaxes back to $\mathrm{K_{Gr}}$, refilling the excited-states of graphene (Fig. 4**h**). From the graphene side, the photoexcited hot electrons N_{Gr}^{el} (red curve in Fig.4**b**) could either scatter to conduction bands of WSe_2 or relax by interband decay channels in graphene. Therefore, the relaxation of N_{Gr}^{el} could be characterized

with the charge transfer time of τ_{ICT} and a decay time of τ_{Gr}^{el}. The deep valence band holes N_{Gr}^{h} (blue curve in Fig.4**b**) are populated by exciton energy transfer on a time scale of τ_{IET}, which would relax back to the Fermi level with a lifetime of τ_{Gr}^{h}.

By numerically solving a rate equation model based on a multi-state scheme (see SI), we disentangle the dynamics of IET and ICT. Our global fit describes the data well and yields the transfer times of $\tau_{IET} = 67 \pm 7$ fs and $\tau_{ICT} = 118 \pm 18$ fs. The lifetimes of electrons and IET-populated hot holes in graphene are simultaneously extracted as $\tau_{Gr}^{el} = 84 \pm 7$ fs and $\tau_{Gr}^{h} = 7 \pm 4$ fs. Combining all our observations and analysis of the energy-momentum dynamics in WSe_2 and graphene, we summarize the interfacial phenomena governing the non-equilibrium behaviour of our heterostructure: first, the optical pump generates excitons in WSe_2 and quasi-free carriers in graphene (Fig. 4**e**). Following photoexcitation, the exciton annihilation excites deep valence electrons in graphene via an IET process (Fig. 4**f-g**). Simultaneously, hot electrons in graphene are injected to the conduction bands of WSe_2 via ICT which transiently shift the valence bands of WSe_2 (Fig. 4**h**).

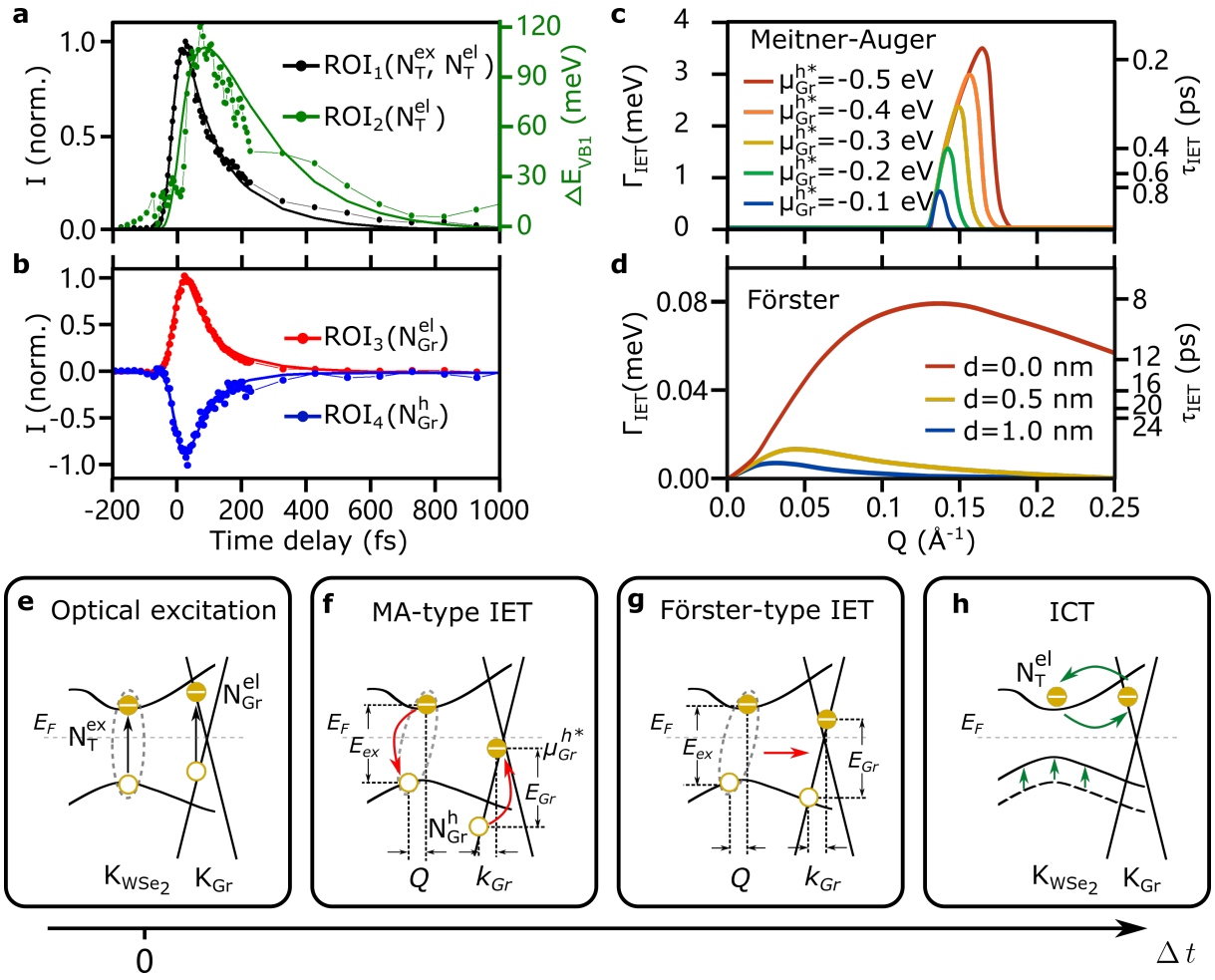


Fig. 4: Interlayer charge and energy transfer upon resonant A-exciton excitation. **a**, By integrating the ROI_1 in Fig.3**c**, the time trace of the normalized photoemission intensity of excited-state carriers at the CBM of WSe_2 (black) contains the dynamics of excitons (N_T^{ex}) and ICT-induced quasi-free carriers (N_T^{el}). The energy shift of VB1 (green) mainly reflects the dynamic of N_T^{el}, which are extracted from time-dependent EDCs in ROI_2. **b**, The time traces of hot electrons (red) and hot holes in the deep VB (blue) in graphene are extracted from the ROI_3 and ROI_4 in Fig.3**c**, respectively. The time traces in **a-b** are fitted based on a rate equation model (see text). **c**, Calculated Meitner-Auger mediated

IET transfer rate as a function of COM momentum $\mathbf{Q}$ with different photo-induced hole vacancy at $E = \mu_{Gr}^{h*}$. **d**, Calculated Förster coupling rate as a function of $\mathbf{Q}$ with varied interlayer distance of d. Sketch of the underlying carrier dynamics: **e**, Optical excitation of excitons in WSe_2 and quasi-free hot carriers in graphene. **f**, Meitner-Auger IET with creation of intraband electron-hole pairs in graphene by absorbing the exciton energy. **g**, Förster-type energy transfer with generation of interband electron-hole pairs in graphene. **h**, ICT-induced hot electron injection into WSe_2 and transient energy shift of its valence band.

To elucidate the interfacial coupling mechanism at play in our experiment, in particular the observed ultrafast energy transfer rate, we perform microscopic calculations of three types of IET mechanisms: Meitner-Auger, Förster and Dexter energy transfer. The interlayer MA process is described by the dipole-monopole energy transfer from excitons to valence band excitation, schematically shown in Fig. 4**f**. The photoexcited hot holes in graphene quickly relax and distribute below E_F near a transient chemical potential μ_{Gr}^{h*}. This allows an MA-type transition from the deep valence band to the hot hole vacancy by absorbing the exciton energy. The microscopically calculated transfer rate is plotted as a function of $\mathbf{Q}$ in Fig.4**c** with different transient chemical potentials for the hole distributions μ_{Gr}^{h*}. When the hole vacancy is located around $\mu_{Gr}^{h*} = -0.3\,\mathrm{eV}$, the maximum transfer rate reaches $\Gamma_{IET} = 2.4\,\mathrm{meV}$, corresponding to a $\tau_{IET} = 270\,\mathrm{fs}$ transfer time. The MA-type IET process could describe the observed energy-momentum distribution of intraband transition of valence electrons in a reasonable quantitative agreement with the extracted transfer rate.

Another IET mechanism is Förster energy transfer (Fig.4**g**). The energy of the exciton excites an *interband* transition from valence bands to above Dirac point via the dipole-dipole coupling.[41] The coupling strength is explicitly evaluated (for derivation, see SI) and determined by the momentum $\mathbf{Q}$ and interlayer distance d. The strong exciton oscillator strength and intrinsic in-plane exciton dipole moment in many 2D materials favor the Förster-type IET[42]. However, the calculated transfer rate is only $0.08\ \mathrm{meV}$ (a transfer time of ~ 8.1 ps), even assuming a tightly stacked heterostructure with interlayer distance of $d = 0\ \mathrm{nm}$ (Fig.4**d**). Our calculations reveal that the IET process preferably excites an *intraband* rather than an *interband* transition. The experimentally observed energy-momentum distribution of excited-state hot holes supports this conclusion. In addition, we also performed calculations of Dexter-type IET and found a very weak interlayer coupling strength, more than three orders of magnitude smaller compared to the other two mechanisms, due to the small wave function overlap and the finite momentum distance between $\mathrm{K}_{\mathrm{WSe_2}}$ and K_{Gr} (see SI). We can thus identity the MA-type conversion of excitons in WSe_2 to intraband excitations in graphene as the dominant IET mechanism.

In this work, we provide a detailed microscopic picture of interfacial charge and energy transfer processes in photoexcited ML-WSe_2/graphene heterostructures. Optical excitation of electrons in graphene leads to inter-layer charge transfer of quasi-free electrons from the graphene layer to the K and Q valleys of the semiconductor's conduction bands on a time scale of $\sim$50 fs. In contrast, excitons in WSe_2 decay through an interfacial Meitner-Auger energy transfer process with a time constant of $\sim$70 fs. This previously unidentified process is governed by inter-layer dipole-monopole interactions leading to annihilation of an exciton in WSe_2 and *non-vertical intraband*

excitations in graphene. The momentum of the electron-hole pair in graphene originates from the finite center of mass momentum of the hot excitons in WSe_2. The interfacial Meitner-Auger mechanism is found to dominate the energy transfer process over established mechanisms like Förster- and Dexter-type transfer. This mechanism results in transient hole distributions as low as 2 eV below the Dirac points. These observations enrich the physical toolbox for designing van der Waals heterostructures and might be utilized in hot-carrier photovoltaic device concepts to harness the ultrafast and efficient carrier transfer processes at interfaces[43].

Methods

Time- and angle-resolved photoemission spectroscopy We used a 500kHz tabletop femtosecond optical parametric chirped pulse amplification (OPCPA) laser system operated at a center wavelength of 800 nm and delivering average power up to 15 W. The high harmonic generation is produced in a vacuum chamber by tight focusing (10 μm) the second harmonic (400 nm) of the OPCPA fundamental on a thin and dense argon gas jet. We select the photons around 21.7 eV (110 meV FWHM bandwidth) as the probe arm for trARPES experiment[20]. Concerning the pump arm, we used two different beams for this study. One pump beam is directly obtained from the OPCPA (800 nm, FWHM=35 fs) and another one is the residual power of the compressed fiber amplifier (1030 nm, FWHM=200 fs). The pump and probe beams are coupled into an ultra-high-vacuum (UHV) chamber and spatially overlapped at the sample position which is controlled by a six-axis manipulator (Carving, SPECS GmbH). The main UHV chamber is equipped with an unique combination of a hemispherical electron energy analyzer (PHOIBOS150, SPECS GmbH)

and time-of-flight (ToF) momentum microscope (METIS1000, SPECS GmbH)[21]. On the one hand, the hemispherical analyzer, which can work in a multi-electrons per laser shot regime, provides high statistic energy/momentum cuts along a given momentum direction, as shown in Fig.3. On the other hand, the momentum microscope allows for efficient, parallel, momentum-resolved detection of the full photoemission horizon from the surface as shown in Fig.1**b** and Fig.2**a-e**. All the experiments are performed at room temperature.

ML-WSe_2/ML-Graphene vdW heterostructure fabrication Monolayer graphene on SiC was grown using the well-established recipe of sublimation growth at elevated temperatures in an argon atmosphere[4]. Note that, on SiC[44], the graphene monolayer resides on top of a $(6\sqrt{3} \times 6\sqrt{3})$R30° reconstructed carbon buffer layer that is covalently bound to the SiC substrate. WSe_2 films were grown on the thus prepared MLG/SiC substrates via hybrid-pulsed-laser deposition (hPLD) in ultra-high vacuum[24]. Pure tungsten (99.99%) was ablated using a pulsed KrF excimer laser (248 nm) with a repetition rate of 10 Hz, while pure selenium (99.999%) was evaporated from a Knudsen cell at a flux rate of around 1.5 $\text{Å}/\text{s}$ as monitored by a quartz crystal microbalance. The deposition was carried out at 450°C for 6 h, followed by two-step annealing at 640°C and 400°C for 1 h each.

DFT calculation of band structure We performed density functional theory (DFT) calculation of ML WSe_2 and graphene with the projector augmented wave code GPAW[45] using Perdew-Burke-Ernzerhof functional[46] including the spin-orbital coupling, separately. We sample the Brillouin Zone with a $(15 \times 15 \times 1)$ k-point mesh, and set the cutoff energy for the plane-wave expansion at 600 eV. The calculated band structures of both materials are superimposed on each other and

shown in Fig.2**a**.

Data availability

All data underlying this study are available from the Zenodo repository. Source data are provided with this paper.

References

1. Novoselov, K.S., Mishchenko, A., Carvalho, A. & Neto, A. H. 2D materials and van der Waals heterostructures. *Science* **353**, 6298 (2016).

2. Akinwande, D. *et al.* Graphene and two-dimensional materials for silicon technology. *Nature* **573**, 507–518 (2019).

3. Liu, F. *et al.* Disassembling 2D van der Waals crystals into macroscopic monolayers and reassembling into artificial lattices. *Science* **367**, 903–906 (2020).

4. Emtsev, K. V. *et al.* Towards wafer-size graphene layers by atmospheric pressure graphitization of silicon carbide. *Nat. Mater.* **8**, 203–207 (2009).

5. Jin, C. *et al.* Ultrafast dynamics in van der Waals heterostructures. *Nat. Nanotechnol.* **13**, 994–1003 (2018).

6. Lee, C.-H. *et al.* Atomically thin p–n junctions with van der Waals heterointerfaces. *Nat. Nanotechnol.* **9**, 676 (2014).

7. McGilly, L. J. *et al.* Visualization of moiré superlattices. *Nat. Nanotechnol.* **15**, 580–584 (2020).

8. Cao, Y. *et al.* Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43–50 (2018).

9. Kunstmann, J. *et al.* Momentum-space indirect interlayer excitons in transition-metal dichalcogenide van der Waals heterostructures. *Nat. Phys.* **14**, 801–805 (2018).

10. Rivera, P. *et al.* Observation of long-lived interlayer excitons in monolayer $MoSe_2$–WSe_2 heterostructures. *Nat. Commun.* **6**, 6242 (2015).

11. Avsar, A. *et al.* Spin–orbit proximity effect in graphene. *Nat. Commun.* **5**, 4875 (2014).

12. Garcia, J. H., Vila, M., Cummings, A. W. & Roche, S. Spin transport in graphene/transition metal dichalcogenide heterostructures. *Chem. Soc. Rev.* **47**, 3359–3379 (2018).

13. Massicotte, M. *et al.* Photo-thermionic effect in vertical graphene heterostructures. *Nat. Commun.* **7**, 12174 (2016).

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

15. Froehlicher, G., Lorchat, E. & Berciaud, S. Charge versus energy transfer in atomically thin graphene-transition metal dichalcogenide van der Waals heterostructures. *Phys. Rev. X* **8**, 011007 (2018).

16. Aeschlimann, S. *et al.* Direct evidence for efficient ultrafast charge separation in epitaxial WS_2/graphene heterostructures. *Sci. Adv.* **6**, 0761 (2020).

17. Hill, H. M. *et al.* Exciton broadening in WS_2/graphene heterostructures. *Phys. Rev. B* **96**, 205401 (2017).

18. Yuan, L. *et al.* Photocarrier generation from interlayer charge-transfer transitions in WS_2-graphene heterostructures. *Sci. Adv.* **4**, 1700324 (2018).

19. He, J. *et al.* Electron transfer and coupling in graphene–tungsten disulfide van der Waals heterostructures. *Nat. Commun.* **5**, 5622 (2014).

20. Puppin, M. *et al.* Time-and angle-resolved photoemission spectroscopy of solids in the extreme ultraviolet at 500 kHz repetition rate. *Rev. Sci. Instrum.* **90**, 023104 (2019).

21. Maklar, J. *et al.* A quantitative comparison of time-of-flight momentum microscopes and hemispherical analyzers for time-resolved ARPES experiments *Rev. Sci. Instrum.* **91**, 123112 (2020).

22. Li, Y. *et al.* Measurement of the optical dielectric function of monolayer transition-metal dichalcogenides: MoS_2, $MoSe_2$, WS_2, and WSe_2. *Phys. Rev. B* **90**, 205422 (2014).

23. Bostwick, A., Ohta, T., Seyller, T., Horn, K. & Rotenberg, E. Quasiparticle dynamics in graphene. *Nat. Phys.* **3**, 36–40 (2007).

24. Nakamura, H. *et al.* Spin splitting and strain in epitaxial monolayer WSe_2 on graphene. *Phys. Rev. B* **101**, 165103 (2020).

25. Baudisch, M. *et al.* Ultrafast nonlinear optical response of Dirac fermions in graphene. *Nat. Commun.* **9**, 1018 (2018).

26. Breusing, M., Ropers, C. & Elsaesser, T. Ultrafast carrier dynamics in graphite. *Phys. Rev. Lett.* **102**, 086809 (2009).

27. Chen, Y., Li, Y., Zhao, Y., Zhou, H. & Zhu, H. Highly efficient hot electron harvesting from graphene before electron-hole thermalization. *Sci. Adv.* **5**, 9958 (2019).

28. Dong, S. *et al.* Direct measurement of key exciton properties: Energy, dynamics, and spatial distribution of the wave function. *Nat. Sci.* **1**,e10010 (2021).

29. Chernikov, A., Ruppert, C., Hill, H. M., Rigosi, A. F. & Heinz, T. F. Population inversion and giant bandgap renormalization in atomically thin WS_2 layers. *Nat. Photonics* **9**, 466–470 (2015).

30. Liu, F., Ziffer, M. E., Hansen, K. R., Wang, J. & Zhu, X. Direct determination of band-gap renormalization in the photoexcited monolayer MoS_2. *Phys. Rev. Lett.* **122**, 246803 (2019).

31. Gao, S., Liang, Y., Spataru, C. D. & Yang, L. Dynamical excitonic effects in doped two-dimensional semiconductors. *Nano Lett.* **16**, 5568–5573 (2016).

32. Dendzik, M. *et al.* Observation of an excitonic Mott transition through ultrafast core-*cum*-conduction photoemission spectroscopy. *Phys. Rev. Lett.* **125**, 096401 (2020).

33. Liang, Y. & Yang, L. Carrier plasmon induced nonlinear band gap renormalization in two-dimensional semiconductors. *Phys. Rev. Lett.* **114**, 063001 (2015).

34. Tomadin, A., Brida, D., Cerullo, G., Ferrari, A. C. & Polini, M. Nonequilibrium dynamics of photoexcited electrons in graphene: Collinear scattering, Auger processes, and the impact of screening. *Phys. Rev. B* **88**, 035430 (2013).

35. Gierz, I., Link, S., Starke, U. & Cavalleri, A. Non-equilibrium Dirac carrier dynamics in graphene investigated with time-and angle-resolved photoemission spectroscopy. *Faraday Discuss.* **171**, 311–321 (2014).

36. Na, M. *et al.* Direct determination of mode-projected electron-phonon coupling in the time domain. *Science* **366**, 1231–1236 (2019).

37. Selig, M. *et al.* Excitonic linewidth and coherence lifetime in monolayer transition metal dichalcogenides. *Nat. Commun.* **7**, 13279 (2016).

38. Christiansen, D., Selig, M., Malic, E., Ernstorfer, R. & Knorr, A. Theory of exciton dynamics in time-resolved ARPES: Intra-and intervalley scattering in two-dimensional semiconductors. *Phys. Rev. B* **100**, 205401 (2019).

39. Selig, M. *et al.* Ultrafast dynamics in monolayer transition metal dichalcogenides: Interplay of dark excitons, phonons, and intervalley exchange. *Phys. Rev. Res.* **1**, 022007 (2019).

40. Pöllmann, C. *et al.* Resonant internal quantum transitions and femtosecond radiative decay of excitons in monolayer WSe_2. *Nat. Mater.* **14**, 889–893 (2015).

41. F ö rster, T. Intermolecular energy migration and fluorescence. *Ann. Phys.* **437**, 55–75 (1948).

42. Kozawa, D. *et al.* Evidence for fast interlayer energy transfer in $MoSe_2$/WS_2 heterostructures. *Nano Lett.* **16**, 4087–4093 (2016).

43. Paul, K. K., Kim, J.-H. & Lee, Y. H. Hot carrier photovoltaics in van der Waals heterostructures. *Nat. Rev. Phys.* **3**, 178–192 (2021).

44. Riedl, C., Coletti, C. & Starke, U. Structural and electronic properties of epitaxial graphene on SiC (0 0 0 1): a review of growth, characterization, transfer doping and hydrogen intercalation. *J. Phys. D: Appl. Phys.* **43**, 374009 (2010).

45. Mortensen, J. J., Hansen, L. B. & Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B* **71**, 035109 (2005).

46. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).

Acknowledgements This work was funded by the Max Planck Society, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant No. ERC-2015-CoG-682843), the German Research Foundation (DFG) within the Emmy Noether program (Grant No. RE 3977/1), through Projektnummer 18208777-SFB 951 "Hybrid Inorganic/Organic Systems for Opto-Electronics (HIOS)" (CRC 951 project B12, M.S., D.C., A.K.), and the SFB/TRR 227 "Ultrafast 264 Spin Dynamics" (projects A09 and B07). S.B. acknowledges financial support from the NSERC-Banting Postdoctoral Fellowships Program. M.D. acknowledges financial support from the Göran Gustafssons Foundation. K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan (Grant Number JPMXP0112101001) and JSPS KAKENHI (Grant Numbers 19H05790 and JP20H00354).

Author contributions S.D., S.B., T.P., M.D., J.M., A.N. and L.R. performed the trARPES measurement. S.D. analyzed the data and wrote the first draft of the manuscript. R.E., L.R. and M.W. were responsible for developing all the experimental infrastructures. M.S. and D.C. performed the microscopic calculation with the guidance of A.K.. R.P.X. developed the 4D data processing code. P.R. and H.N. provided the epitaxially-grown heterostructure, with support from U.S. and H.T.. A.M., A.S., and M.S. conducted Raman and photoluminescence measurements, with guidance from M.J and P.M.. J.D.Z and A.C. prepared the exfoliated ML sample with the hBN substrate provided by K.W. and T.T.. All authors contributed to the

final version of the manuscript.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to dong@fhi-berlin.mpg.de, rettig@fhi-berlin.mpg.de and ernstorfer@fhi-berlin.mpg.de.

Supplementary Information for: Observation of ultrafast interfacial Meitner-Auger energy transfer in a van der Waals heterostructure

Shuo Dong[1*], Samuel Beaulieu[1,2], Malte Selig[3], Philipp Rosenzweig[4], Dominik Christiansen[3], Tommaso Pincelli[1], Maciej Dendzik[1,5], Jonas D. Ziegler[6], Julian Maklar[1], R. Patrick Xian[1,7], Alexander Neef[1], Avaise Mohammed[4], Armin Schulz[4], Mona Stadler[8], Michael Jetter[8], Peter Michler[8], Takashi Taniguchi[9], Kenji Watanabe[10], Hidenori Takagi[4,11,12], Ulrich Starke[4], Alexey Chernikov[6,13], Martin Wolf[1], Hiro Nakamura[4,14], Andreas Knorr[3], Laurenz Rettig[1*] & Ralph Ernstorfer[1,15*]

[1]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany*

[2]*Université de Bordeaux - CNRS - CEA, CELIA, UMR5107, F33405, Talence, France*

[3]*Nichtlineare Optik und Quantenelektronik, Institut für Theoretische Physik, Technische Universität Berlin, 10623 Berlin, Germany*

[4]*Max Planck Institute for Solid State Research, 70569 Stuttgart, Germany*

[5]*Department of Applied Physics, KTH Royal Institute of Technology, Hannes Alfvéns väg 12, 114 19 Stockholm, Sweden*

[6]*Department of Physics, University of Regensburg, Regensburg D-93053, Germany*

[7]*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*

[8]*Institute of Semiconductor Optics and Functional Interfaces, Research Center SCoPE and IQST, University of Stuttgart, 70569 Stuttgart, Germany*

[9]*International Center for Materials Nanoarchitectonics, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan*

[10]*Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan*

[11]*Department of Physics, University of Tokyo, 113-0033 Tokyo, Japan*

[12]*Institute for Functional Matter and Quantum Technologies, University of Stuttgart, 70569 Stuttgart, Germany*

[13]*Institute for Applied Physics, Dresden University of Technology, Dresden, 01187, Germany*

[14]*Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA*

[15]*Institut für Optik und Atomare Physik, Technische Universität Berlin, 10623 Berlin, Germany*

This file includes:

Supplementary Notes

Supplementary Methods

Supplementary Equation (1) to (64)

Supplementary Fig.1 to Supplementary Fig.9

Supplementary Table 1

References

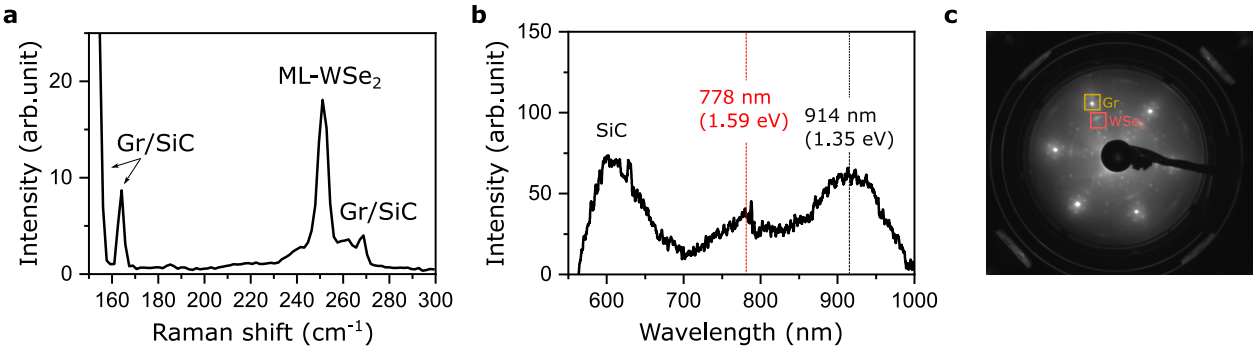
Supplementary Notes

Characterization of the ML-WSe_2/graphene heterostructure

Raman measurements were carried out at room temperature using a 532 nm laser with a power of 1 mW and a spot size of 5 to 10 μm. As shown in Supplementary Fig. 1**a**, ML-WSe_2 is confirmed by an intense peak at 250 cm^{-1} which comes from essentially degenerate A_{1g} and E_{2g} lattice vibration modes[1,2]. Photoluminescence (PL) measurements are performed using another system with a 532 nm excitation laser, power of 1 mW and spot size of 1 μm at room temperature. In Supplementary Fig. 1**b**, ML-WSe_2 on graphene presents two weak PL peaks (778 nm and 914 nm) only slightly above the background. The peak at 778 nm is close in energy to the A-exciton transition energy[3,4]. The origin of the peak at higher wavelength is unknown, and may come from the existence of in-gap defect states. The weakness of the PL signals is consistent with the quenching of PL known to occur for ML-TMDCs adjacent to graphene[5].

The sample is protected by the Se capping layer before sending to our lab. After introducing the sample into our ultrahigh vacuum (UHV) photoemission end-station, we have annealed the sample for 15 minutes at 400°C through direct current heating to remove the Se capping. After annealing, we recorded a low energy electron diffraction (LEED) pattern with the incident beam energy of 95 eV, to verify the surface cleanliness and ordering (Supplementary Fig. 1**c**). The six outer sharp LEED spots come from the bottom ML graphene layer (yellow box) and the inner six arc-shaped diffraction spots originate from the top ML-WSe_2 layer (red box). The occurrence of a well-oriented hexagonal pattern of WSe_2 spots aligned to the graphene pattern attests the epitaxial

nature of our heterostructure and single-domain lattice orientation. A certain level of strain-induced misalignment between nanoislands of WSe_2 with respect to the graphene layer is evident from the azimuthal widths of the diffraction spots.



Supplementary Fig. 1: Optical characteristics and surface analysis of the heterostructure sample. a, Raman measurement of ML-WSe_2/graphene (Gr) at room temperature. An intense peak at 250 cm^{-1} belongs to ML-WSe_2, whereas other peaks belong to the Gr/SiC substrate. **b**, Photoluminescence measurement of the ML-WSe_2/Gr heterostructure. The peak originating from the A-exciton of ML-WSe_2 is marked with the dashed red line. **c**, LEED pattern of the ML-WSe_2/Gr heterostructure at 95 eV after annealing.

Delayed population rises at $\mathrm{K}_{\mathrm{WSe_2}}$ and $\mathrm{Q}_{\mathrm{WSe_2}}$ with the below-bandgap excitation

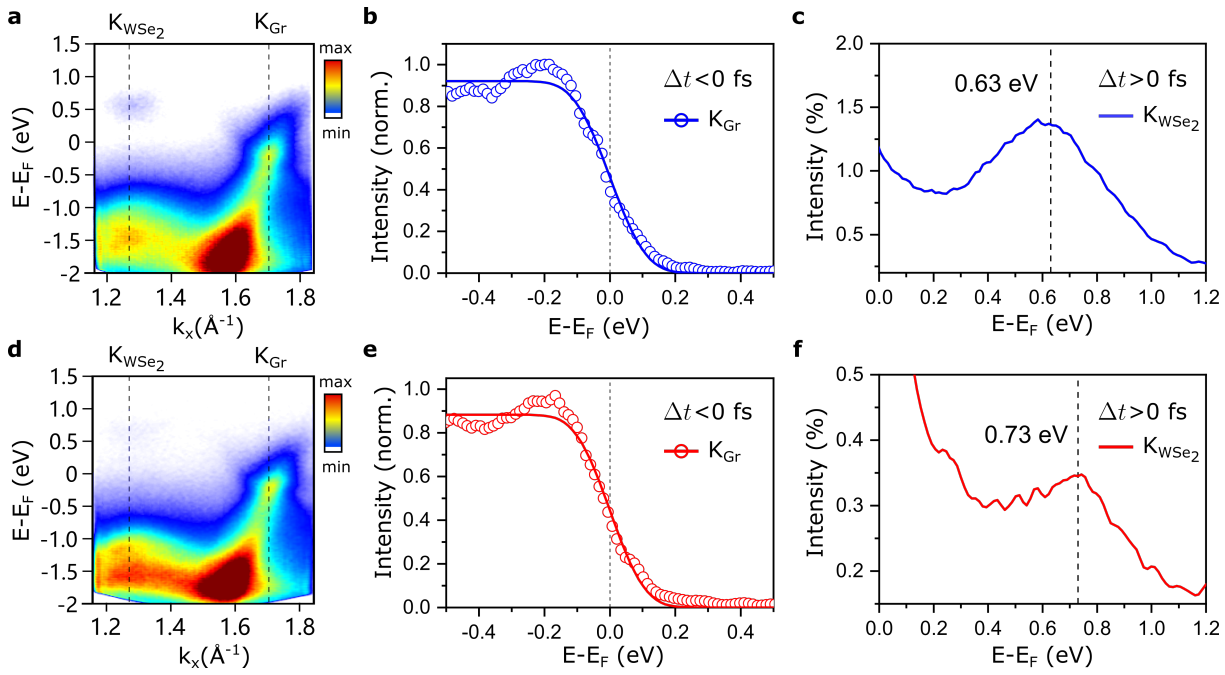
Upon 1.2 eV excitation, the time trace at the K_{Gr} points shown in Fig.2**f** is by fitting with a single exponential decay function convolved with the instrument response function (IRF), $I(t) = H(t-t_0) \times (A \cdot exp(-(t-t_0)/\tau) + C) \otimes \mathrm{IRF}$. Here, $H(t)$ is the Heaviside step function, IRF is Gaussian envelope function, A is the amplitude and C is the offset. In contrast, as the $\mathrm{K}_{\mathrm{WSe_2}}$ and

Q_{WSe_2} valleys are populated by interlayer charge transfer (ICT) processes, the corresponding time traces are fitted with an exponential growth function to describe the ICT process adding with a single exponential decay for relaxation. The delayed population rise between graphene and WSe_2, $\Delta t = 51 \pm 9$ fs, is obtained by taking the time difference between the centre of the IRF at K_{Gr} and the time delay when the population at the K_{WSe_2} valley reaches its maximum.

Identification of the Fermi level and distribution of excited-state carriers at K_{WSe_2}

In both experiments, the energy axis calibration has been performed using the position of the Fermi level of graphene, which is obtained from the energy distribution curve (EDC) at the Dirac point of graphene (Supplementary Fig. 2**a** and **d**). Before optical excitation, the EDCs at the K_{Gr} point (Supplementary Fig. 2**b,e**) are fitted with a Fermi-Dirac distribution function at 300 K convolved with the IRF ($\sim$150 meV FWHM) determined by the energy resolution of the spectrometer and the bandwidth of the probe pulses[6]. The chemical potentials are set to be zero for both experimental conditions to remove the XUV-probe-induced space charging effect in each measurement. The EDCs in Supplementary Fig. 2**b,e** are integrated over a momentum window, $\Delta k = 0.1$ Å^{-1} and selected at negative time delay. Based on the energy reference obtained from the Fermi level fits, the energy positions of the conduction band minima at the K_{WSe_2} point can be obtained from EDCs showing the excited state carrier distributions of WSe_2 upon resonant and off-resonant excitation as displayed in Supplementary Fig. 2**c** and **f**, respectively. The photoemission intensity has been normalized by the total electron count of the spectrum. The energy difference of the carrier distribution, more specifically, the smaller kinetic energy of excited carriers with 1.55 eV pump,

arises from the exciton formation upon the resonant excitation. By fitting with a single Gaussian lineshape on top of a empirical second-order polynomial background, the energies of excited-state carriers are extracted as 0.63 eV upon the resonant A-exciton excitation (Supplementary Fig. 2**c**) and 0.73 eV upon the below-bandgap excitation (Supplementary Fig. 2**f**). The background has been removed in the main text (side figures in Fig.3**a-b**).

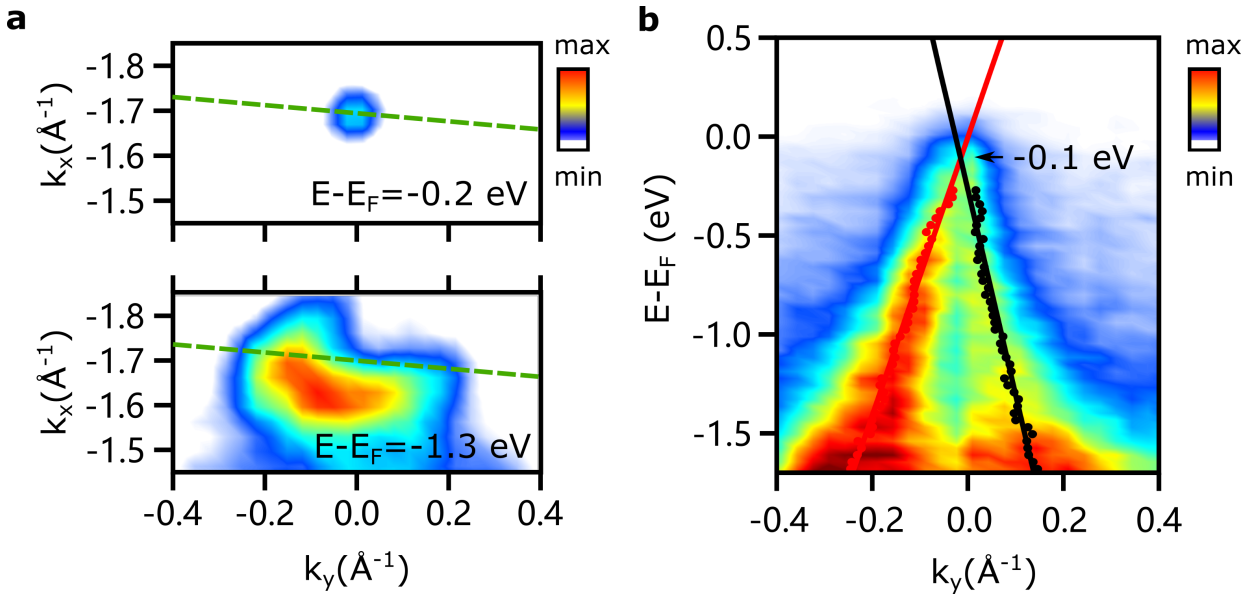


Supplementary Fig. 2: Fermi level calibration. a,d, The 2D photoemission intensity spectra as a function of energy and momentum at time zero with 1.55 eV and 1.2 eV pump, respectively. **b,e**, EDCs of $\mathrm{K_{Gr}}$ at negative time delay fitted with a Fermi-Dirac distribution convolved with the energy IRF for 1.55 eV and 1.2 eV pump, respectively. The chemical potentials are aligned to zero by rigidly shifting the energy axis in both cases. **c, f** EDCs at $\mathrm{K_{WSe_2}}$ integrated within the first 100 fs obtained with 1.55 eV and 1.2 eV pump, respectively. The dashed lines represent the center of the excited-state carrier

distributions extracted by a fitting procedure (see text).

Identification of the Dirac point energy

To identify the energy position of the Dirac point, we selected an energy-k_y cut ($\Delta t < 0$ fs) at Dirac point and along the green dashed line in Supplementary Fig. 3**a**. The small titling angle allows us to see both valence bands clearly in Supplementary Fig. 3**b**. We track the graphene valence band dispersion by fitting the momentum distribution curves of occupied bands with two Voigt lineshape functions. Then, each graphene valence band is fitted to a linear dispersion and the Dirac point is estimated at the intersection of two lines (red and black), $E - E_F = -0.10 \pm 0.05$ eV, in a reasonable agreement with the previous characterization of similar heterostructures[7]. This energy/momentum cut is different than the one presented in the main text (Fig.3) along the Γ-K direction, which is featured by the suppression of one side of the cone due to photoemission matrix element effects (sublattice interference)[8]. In the main text, we choose this cut direction because it allows us to clearly resolve the excited state dynamics from both layers.

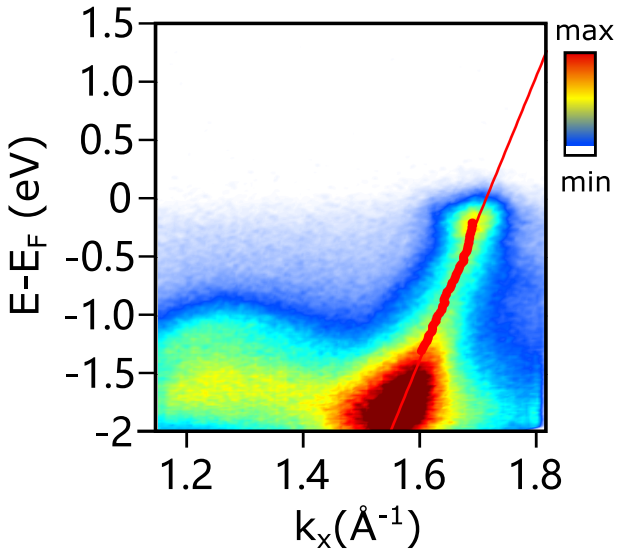


Supplementary Fig. 3: Experimental determination of the Dirac energy. **a**, 2D momentum distribution map $I(k_x, k_y)$ at $E - E_F = -0.2$ eV and $E - E_F = -1.3$ eV. At the boundary of the Brillouin zone, it shows the quasi-triangular-shaped π band of graphene. **b**, Energy/momentum cut $I(E, k)$ along the green dash line in **a** showing the conical band dispersion of graphene. Red and black markers indicate the band positions extracted from momentum distribution curve fits. Lines are linear fits of the band positions, yielding the energy position of the Dirac point of $E_D = -0.1 \pm 0.05$ eV.

Identification of the Fermi velocity

The Fermi velocity along Γ-K direction is extracted as $v_F = (1.8 \pm 0.1) \cdot 10^6\ m/s$ as shown in Supplementary Fig. 4 using the same band dispersion tracking method in the above paragraph. The energy/momentum spectrum is selected at negative time delay ($\Delta t < 0$ fs). The Fermi velocities

in the direction perpendicular to Γ-K with a slightly tilting angle (Supplementary Fig. 3**b**) are $v_F = (1.1 \pm 0.1) \cdot 10^6\ m/s$ (red) and $v_F = (1.5 \pm 0.1) \cdot 10^6\ m/s$ (black), respectively. The Fermi velocity of graphene has been found to to be in the range of $1 \cdot 10^6$ to $3 \cdot 10^6$ m/s, depending on the dielectric constant of the environment[9]. For epitaxially grown heterostructures, the dielectric constant of the embedding graphene layer between the bottom substrate and top TMDC layer could be modified by the coverage sizes of TMDC layer and substrate material, as the dielectric constant is determined by $\epsilon = (\epsilon_{top}+\epsilon_{substrate})/2$. At the same time, the Fermi velocity is also sensitive to the graphene's doping level[10]. We would like to note that because of the steep band dispersion of the graphene band, the momentum of transiently excited intraband electron-hole pairs is small. Therefore, it requires a relative small excitonic COM momentum **Q** to fulfill energy and momentum conservation, which favors the IET process.

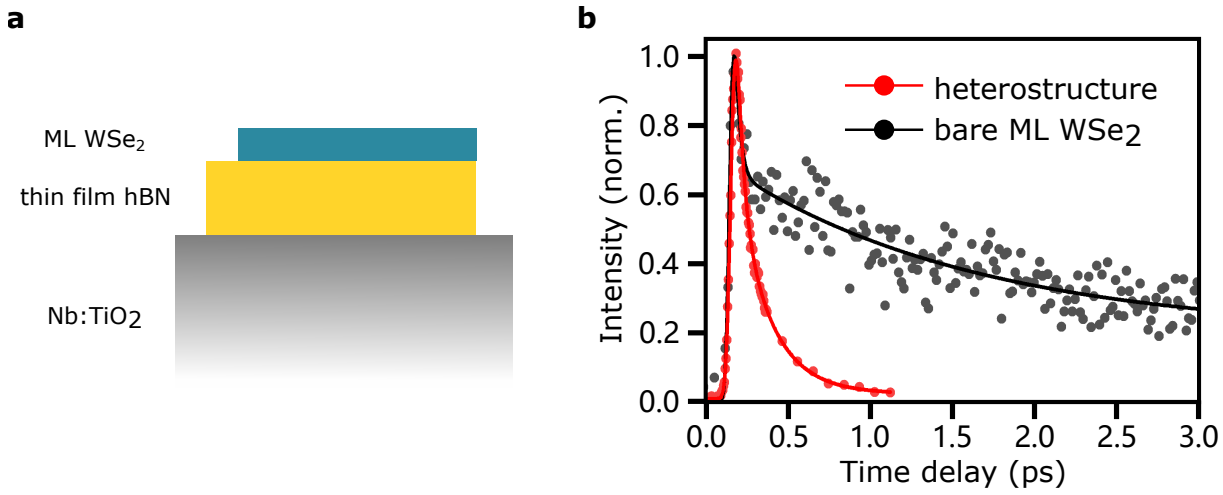


Supplementary Fig. 4: Experimental determination of the Fermi velocity. The energy/momentum cut $I(E,k)$ along Γ-K direction with band positions (red markers) and

the linear fit of the band dispersion (red line).

The near-unity efficiency of IET

To estimate the efficiency of energy transfer, we performed the same measurement (1.55 eV excitation) on bare ML-WSe_2, which is prepared by scotch-tape exfoliation and transferred on top of thin hexagonal boron nitrid (hBN) with conductive TiO_2 substrate (Supplementary Fig. 5**a**). The efficiency of the energy transfer process is commonly defined by the lifetime of the 'donor' material (here WSe_2) with and without the 'acceptor' material (here graphene) as: $\eta_{ET} = (\tau_{ML} - \tau_{hetero})/\tau_{ML}$, where τ_{ML} represents the exciton lifetime of the bare ML-WSe_2 and τ_{hetero} is the exciton lifetime in the WSe_2/graphene heterostructure. The excited-state population dynamics at the $\mathrm{K}_{\mathrm{WSe_2}}$ valley within each system are presented in Supplementary Fig. 5**b**. The lifetimes of the ML sample $\tau_{ML} = 1616 \pm 345$ fs is extracted by fitting with an exponential decay function convolved with the IRF. The exciton lifetime of the heterostructure $\tau_{hetero} = 67 \pm 7$ fs is obtained by solving the system of rate equations as described in the main text. Thus, we obtain for the interlayer energy transfer efficiency, $\eta_{ET} = 96 \pm 1\%$. This near-unity transfer efficiency is supported by the underlying conservation of energy and momentum. We note that the different sample fabrication methods of the bare ML and heterostructure may have influence on the exciton lifetime. However, picosecond to sub-nanosecond exciton lifetimes in ML samples, consistent with our observations, have been reported for samples fabricated with various methods[11,12]. Therefore, we believe that the comparison with the exciton lifetimes in the bare ML WSe_2 sample provides a reasonable estimate of the transfer efficiency.

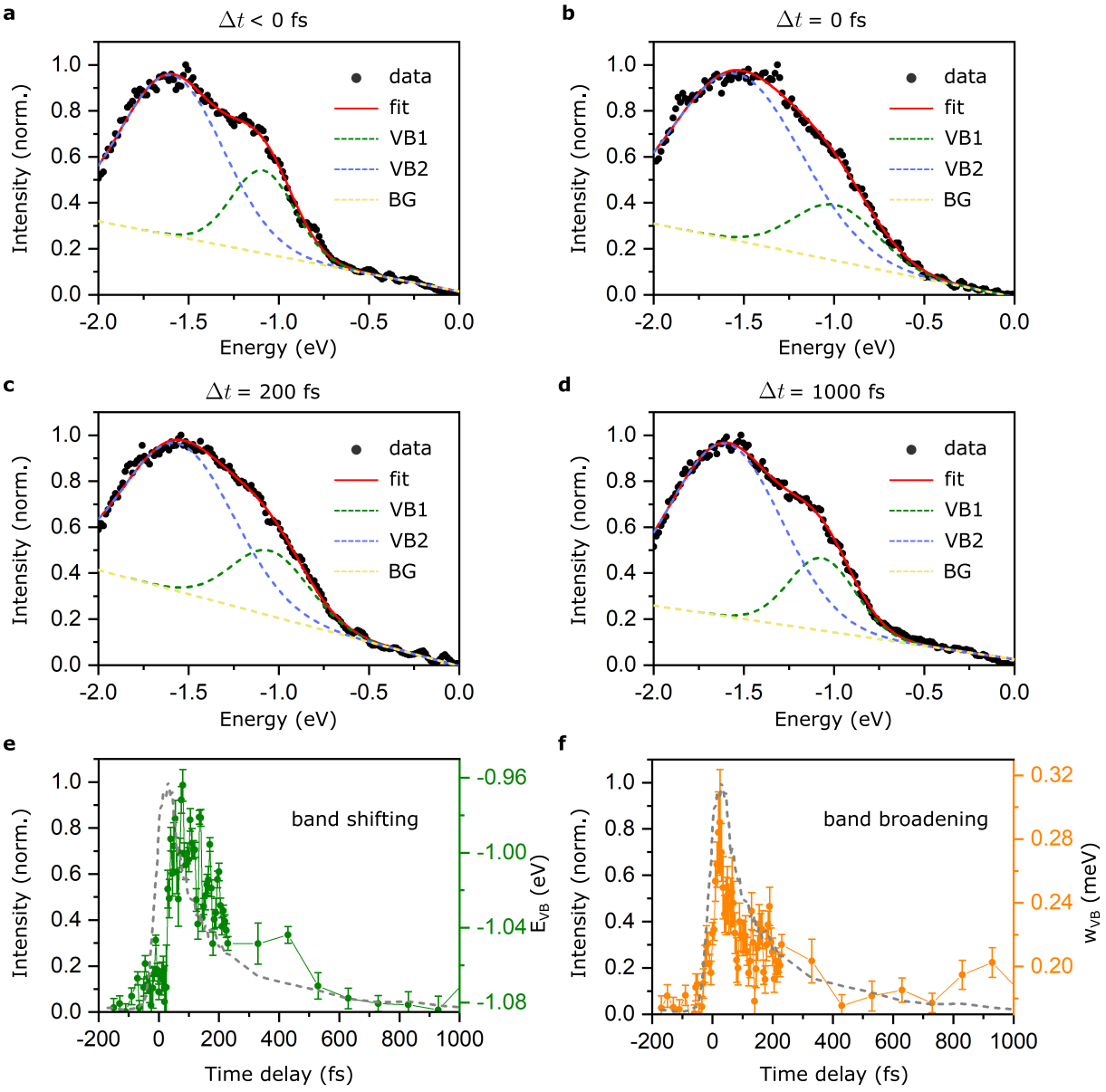


Supplementary Fig. 5: Estimation of the IET efficiency. **a**, Schematic of the bare ML-WSe_2 sample (blue slab) with the bottom hBN layer (yellow) mounted on a Nb:TiO_2 substrate (grey). **b**, Time traces of the excited-state carriers at the $\mathrm{K}_{\mathrm{WSe_2}}$ valleys of the bare ML sample (black) and heterostructure (red) sample, respectively.

Valence bands shifting and broadening effects

After photoexcitation, we observe shifting and broadening effects of the WSe_2 valence band as shown in the 2D difference spectrum (Fig. 3**c**) and EDCs at $\mathrm{K}_{\mathrm{WSe_2}}$ (Fig. 3**d**). To extract the transient lineshape, we fit the EDC of the top two VBs (VB1 and VB2) with two Gaussian functions on top of an empirical second-order polynomial background (BG), $I(E) = A_1 \cdot exp(-\frac{(E-E_1)^2}{2\omega_1^2}) + A_2 \cdot exp(-\frac{(E-E_2)^2}{2\omega_2^2}) + \mathrm{BG}$, where E_1, E_2 are peak positions and ω_1, ω_2 are the peak width. Supplementary Fig. 6**a-d** present representative fitting results at four time delays, $\Delta t < 0$ fs, $\Delta t = 0$ fs, $\Delta t = 200$ fs and $\Delta t = 1000$ fs. Because of the large spectral overlap between VB1 and VB2, the fitting is performed with the same shifting, $\Delta E_1(t) = \Delta E_2(t)$, and broadening parameter

, $\Delta\omega_1(t) = \Delta\omega_2(t)$, for the two peak functions, assuming the VBs respond to the interfacial coupling and the excitation-induced modification in the same way. The extracted time dependent peak shift and linewidth parameter are shown in Supplementary Fig. 6**e** and **f**, respectively. The band shifting reflects the electronic band gap renormalization due to the ICT-induced hot electrons[13,14]. Thus, a relative time delay can be observed compared with the excited-state population dynamics in the conduction bands. In contrast, the linewidth is a measure of the photohole self-energy, which depends on the many-body interactions with photoexcited carriers and phonons[15]. It follows more closely the transient of overall excited carriers in the system (grey dashed curve in Supplementary Fig. 6**f**).



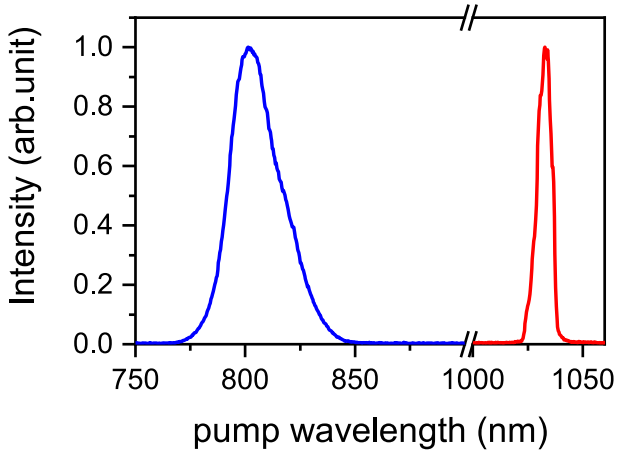
Supplementary Fig. 6: Transient lineshape of the valence bands of WSe_2. a-d, The EDCs at the $\mathrm{K}_{\mathrm{WSe_2}}$ valley present the spectral features of the first two valence bands at selected time delays, $\Delta t < 0$ fs, $\Delta t = 0$ fs, $\Delta t = 200$ fs and $\Delta t = 1000$ fs. The EDCs are fitted with two Gaussian functions describing VB1 (green dashed curve) VB2 (blue dashed curve). The second-order polynomial background is shown as yellow dashed

lines. **e**, Transient peak position of VB1 (green). **f** Peak linewidth both of VBs as function of time (yellow). The population dynamic of excited-states in the WSe_2 conduction band (grey) is shown in **e-f** as a reference.

Supplementary Methods

Characteristics of applied pump beams

In this work, we use two different pump beams which wavelengths are centred at 800 nm and 1030 nm, respectively (Supplementary Fig. 7). The pulse duration of 1030nm pump line is $\sim$ 200 fs FWHM, while the transform-limited pulse duration of 800 nm pump is $\sim$ 35 fs FWHM. In the measurement, the pump fluence of 800 nm is $F_{800} = 1.7\ \mathrm{mJ/cm^2}$ and that of 1030 nm is $F_{1030} = 5.3\ \mathrm{mJ/cm^2}$, with the consideration of effective pump-probe overlap profile based on the formula, $a = \frac{1}{\pi(\omega_{pump}^2+\omega_{probe}^2)}$, in the work of Harb *et al*[16]. Here, ω_{pump} and ω_{probe} are the respective beam widths of the pump and probe beams.



Supplementary Fig. 7: Pump spectra. Excitation spectra of the two light sources used for pumping as a function of wavelength.

Separating the interlayer charge and energy transfer by a rate equation model

As discussed in the main context, we observed the photoemission signatures of both ICT and IET upon resonant A-exciton excitation. To extract the corresponding transfer rates, Γ_{ICT} and Γ_{IET}, we develop a multi-state coupled rate equation model describing the interlayer charge and energy flow, as well as the hot carrier relaxations. In Fig.4**a**, the time trace of hot carriers in the CBM of WSe_2 (black curve) includes the dynamics of photo-generated excitons N_T^{ex} and ICT-induced quasi-free electrons N_T^{el}. The VB1 shifting (green curve in Fig.4**a**) mainly reflects the dynamic of N_T^{el}. Therefore, it provides the possibility to disentangle the dynamics of these two kinds of quasiparticles. Here, subscript ***T*** represents TMDC. Simultaneously, the deep valence band holes in graphene N_{Gr}^{h} are populated by the IET process and recombine with the rate Γ_h, as shown in Fig.4**b**. Finally, the dynamics of hot electrons in graphene N_{Gr}^{el} contains the ICT-induced charge flow (input and output towards WSe_2) and a decay process with the rate of Γ_{el}. With these considerations, the complete dynamics across the interface can be described with the following set of coupled rate equations:

$$\dot{N}_T^{ex} = -\Gamma_{IET} N_T^{ex} + S(t) \tag{1}$$

$$\dot{N}_T^{el} = -\Gamma_{ICT} N_T^{el} + \Gamma_{ICT} N_{Gr}^{el} \tag{2}$$

$$\dot{N}_{Gr}^{el} = +\Gamma_{ICT} N_T^{el} - \Gamma_{ICT} N_{Gr}^{el} - \Gamma_{el} N_{Gr}^{el} + S(t) \tag{3}$$

$$\dot{N}_{Gr}^{h} = -\Gamma_h N_{Gr}^{h} + \Gamma_{IET} N_T^{ex} \tag{4}$$

Here, $S(t)$ represents the optical excitation as a Gaussian-shaped pump envelope function. By numerically solving the system of coupled differential equations, and a global fit of the solution

to the data, we obtain the IET transfer time $\tau_{IET} = 67 \pm 7$ fs, and the ICT transfer time $\tau_{ICT} = 118 \pm 18$ fs ($\tau = \hbar/\Gamma$). At the same time, we get the relaxation times of electrons in graphene, $\tau_{Gr}^{el} = 84 \pm 7$ fs, and that of the deep valence holes, $\tau_{Gr}^{h} = 7 \pm 4$ fs. The fitting results are shown in Fig.4**a** and **b**.

Microscopic calculation of IET mechanisms

We perform microscopic calculations of the IET process mediated by the Meitner-Auger, Förster and Dexter type mechanisms.

Meitner-Auger-like energy transfer A schematic illustration of the Meitner-Auger type (MA) interlayer transfer is depicted in the main context in Fig.4**f**. Here, an exciton in the TMDC recombines non-radiatively; its energy excites an electron deep in the valence band of graphene to states close to the Dirac point but in the valence band.

Starting point for the calculation of the MA-type interlayer coupling is the Hamiltonian

$$H_F = \sum_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}',\lambda,\lambda',\nu,\nu'} V^{\lambda\nu\nu'\lambda'}_{\mathbf{k},\mathbf{q},\mathbf{q}',\mathbf{k}'} \lambda^{\dagger}_{\mathbf{k}} \nu^{\dagger}_{\mathbf{q}} \nu'_{\mathbf{q}'} \lambda'_{\mathbf{k}'}. \tag{5}$$

As a convention, we use $\lambda^{(\prime)}$ as band indices and $\mathbf{k}^{(\prime)}$ as momenta in WSe_2 layer and $\nu^{(\prime)}$ as band indices and $\mathbf{q}^{(\prime)}$ as momenta in graphene. The appearing matrix element is formally given as

$$V^{\lambda\nu\nu'\lambda'}_{\mathbf{k},\mathbf{q},\mathbf{q}',\mathbf{k}'} = \int_{\mathbb{R}^3} d^3r \int_{\mathbb{R}^3} d^3r' \Psi^{\lambda *}_{\mathbf{k}}(\mathbf{r}) \Psi^{\nu *}_{\mathbf{q}}(\mathbf{r}') V(\mathbf{r},\mathbf{r}') \Psi^{\nu'}_{\mathbf{q}'}(\mathbf{r}') \Psi^{\lambda'}_{\mathbf{k}'}(\mathbf{r}). \tag{6}$$

The band indices in the TMD are restricted to interband transitions $\lambda \neq \lambda'$ but the band indices in graphene are taken as the valence band $\nu = \nu' = v$. The remaining integrals can be evaluated

within a $\mathbf{k}\cdot\mathbf{p}$ expansion. Last we introduce exciton operators in WSe_2 $P^{\mu}_{\mathbf{Q}} = \sum_{\mathbf{q}} \varphi^{\mu}_{\mathbf{q}} c^{\dagger}_{\mathbf{q}+\frac{m_e}{m_h+m_e}\mathbf{Q}} v_{\mathbf{q}-\frac{m_h}{m_h+m_e}\mathbf{Q}}$ with quantum state μ and COM momentum $\mathbf{Q}$. The final Hamiltonian reads

$$H = \sum_{\mathbf{k},\mathbf{Q},\mu} W^{\mu}_{\mathbf{Q}} P^{\dagger\mu}_{\mathbf{Q}} v^{\dagger}_{\mathbf{k}-\mathbf{Q}} v_{\mathbf{k}} + h.c.\,, \tag{7}$$

with the coupling element

$$W_{\mathbf{Q}} = \frac{1}{e} V_{\mathbf{Q}} \mathbf{d}^{cv} \cdot \mathbf{Q} \varphi^{*\mu}(\mathbf{r}=0). \tag{8}$$

In the following we restrict ourselves to the lowest bound excitons $\mu = 1s$. From this Hamiltonian we calculate the equation of motion for the exciton occupation in the TMD $N_{\mathbf{Q}} = \langle P^{\dagger}_{\mathbf{Q}} P_{\mathbf{Q}} \rangle$ and the electron occupation in the valence band of graphene $f_{\mathbf{k}} = \langle v^{\dagger}_{\mathbf{k}} v_{\mathbf{k}} \rangle$ by exploiting Heisenberg equation of motion.

The resulting equations of motion read

$$\partial_t N_{\mathbf{Q}} = \frac{2\pi}{\hbar} \sum_{\mathbf{k}} |W_{\mathbf{Q}}|^2 \left(f_{\mathbf{k}}(1 - f_{\mathbf{k}-\mathbf{Q}}) - N_{\mathbf{Q}}(f_{\mathbf{k}-\mathbf{Q}} - f_{\mathbf{k}}) \right) \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{Q}} - E_{\mathbf{Q}}) \tag{9}$$

$$\partial_t f_{\mathbf{k}} = \frac{2\pi}{\hbar} \sum_{\mathbf{Q}} |W_{\mathbf{Q}}|^2 \left(f_{\mathbf{k}-\mathbf{Q}}(1 - f_{\mathbf{k}}) - N_{\mathbf{Q}}(f_{\mathbf{k}} - f_{\mathbf{k}-\mathbf{Q}}) \right) \delta(\epsilon_{\mathbf{k}-\mathbf{Q}} - \epsilon_{\mathbf{k}} - E_{\mathbf{Q}}) \tag{10}$$

$$+ \frac{2\pi}{\hbar} \sum_{\mathbf{Q}} |W_{\mathbf{Q}}|^2 \left(N_{\mathbf{Q}}(f_{\mathbf{k}-\mathbf{Q}} - f_{\mathbf{k}}) - f_{\mathbf{k}}(1 - f_{\mathbf{k}-\mathbf{Q}}) \right) \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{Q}} - E_{\mathbf{Q}}) \tag{11}$$

Estimation of the decay rate of WSe_2 excitons From the Boltzmann equation we can identify the decay rate of WSe_2 excitons as

$$\Gamma_{\mathbf{Q}} = 8\pi \sum_{\mathbf{k}} (f_{\mathbf{k}-\mathbf{Q}} - f_{\mathbf{k}}) \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{Q}} - E_{\mathbf{Q}}), \tag{12}$$

where we have added a factor of 4 to account for the valley and spin degree of freedom in graphene. Analyzing the Dirac distribution, we find that $\mathbf{k}$ accounts for electrons close the Dirac point, and

$\mathbf{k} - \mathbf{Q}$ for electrons deep in the valence band. In order to get a simple expression for the decay rate, we assume that the electrons close to the Dirac point have much smaller momenta than the electrons deep in the valence band, i .e. $\mathbf{k} \ll \mathbf{Q}$ and $\mathbf{k} + \mathbf{Q} \approx \mathbf{Q}$.

This way, the Dirac distribution and the $\mathbf{k}$ can be evaluated analytically yielding

$$\Gamma_{\mathbf{Q}} = |W_{\mathbf{Q}}|^2 \frac{4}{\hbar v_F} \left(Q - \frac{E_{\mathbf{Q}}}{\hbar v_F} \right) \left(f_{\mathbf{Q}} - f_{Q - \frac{E_{\mathbf{Q}}}{\hbar v_F}} \right) \theta \left(Q - \frac{E_{\mathbf{Q}}}{\hbar v_F} \right) \tag{13}$$

The rate depends on the matrix element of the MA transfer, the density of states in graphene and on the occupation difference of the involved states in graphene which accounts for the Pauli blocking. The heavyside function $\theta \left(Q - \frac{E_{\mathbf{Q}}}{\hbar v_F} \right)$ accounts for the fact, that a minimal momentum is required to fulfil the energy and momentum conservation during the intervalley transfer.

Fig. 4**c** in the main context illustrates the MA rate of WSe_2 excitons to graphene for the photo-induced hole vacancies at different energy of μ_{Gr}^{h*}. We adjusted the graphene dispersion to the results from the ARPES measurement. From $\tau = \hbar/\Gamma$ we find scattering times of 270 fs ($\mu_{Gr}^{h*} = -0.3\,\mathrm{eV}$), 210 fs ($\mu_{Gr}^{h*} = -0.4\,\mathrm{eV}$), and 175 fs ($\mu_{Gr}^{h*} = -0.5\,\mathrm{eV}$).

Origin of the finite center-of-mass (COM) momentum The MA mediated IET requires nonzero COM momentum of the exciton. For example, the required COM momentum is $\sim 1.3\,\mathrm{nm}^{-1}$ as shown in Fig.4**c** (main context) with $\mu_{Gr}^{*} = -0.3$ eV, which corresponds to a kinetic energy of approximately 100 meV based on the effective mass of the exciton $m_{ex} = 0.65 m_e$.[17] Then, where does the energy (momentum) come from? In the following, we discuss the possible origins of COM momentum of the exciton which is quantified by the kinetic energy assuming the parabolic

excitonic band dispersion.

At room temperature, the mean kinetic energy of the excitons is 25.6 meV which is not enough to explain the required energy. Therefore, we calculate the temporal evolution of the exciton energy-momentum occupation during the optical pump for detuned excitation[18]. The equation of motion for the excitonic coherence in the rotating frame reads

$$\dot{P}_{\mathbf{0}}(t) = \frac{1}{i\hbar}\left(E_{\mathbf{0}} - \hbar\omega_L - i\gamma\right) P_{\mathbf{0}}(t) + \mathbf{d}\cdot\mathbf{E}(t), \tag{14}$$

where the first term accounts for the detuning of the excitonic transition energy $E_{\mathbf{0}}$ from the light pulse energy $\hbar\omega_L$. γ accounts for the dephasing of the excitonic coherence with contributions from radiative and exciton phonon coupling[19]. The last term accounts for the optical excitation with the dipole element $\mathbf{d}$ and the exciting electric field $\mathbf{E}(t)$. The equation of motion for the incoherent exciton occupation reads

$$\dot{N}_{\mathbf{Q}} = \Gamma^{Form}_{\mathbf{Q}}|P_{\mathbf{0}}|^2 + \sum_{\mathbf{K}} \Gamma^{in}_{\mathbf{Q},\mathbf{K}} N_{\mathbf{K}} - \sum_{\mathbf{K}} \Gamma^{out}_{\mathbf{Q},\mathbf{K}} N_{\mathbf{Q}}. \tag{15}$$

The first term accounts for the formation of incoherent exciton occupation from phonon induced dephasing from the excitonic coherence. The last two terms account for the thermalization of incoherent excitons[20]. The coupling element of the exciton formation reads

$$\Gamma^{Form}_{\mathbf{Q}} = \frac{2}{\hbar}\sum_{\pm,\alpha} |g_{\mathbf{Q}}|^2 \left(\frac{1}{2} \pm \frac{1}{2} + n^{\alpha}_{\mathbf{Q}}\right) \frac{\gamma}{(E_{\mathbf{Q}} - \hbar\omega_L \mp \hbar\Omega^{\alpha})^2 + \gamma^2}. \tag{16}$$

with the energy $\hbar\Omega^{\alpha}_{\mathbf{Q}}$ and the occupation $n^{\alpha}_{\mathbf{Q}}$ of phonons in the branch α with momentum $\mathbf{Q}$. The $\pm$ summation accounts for phonon emission/absorption processes.

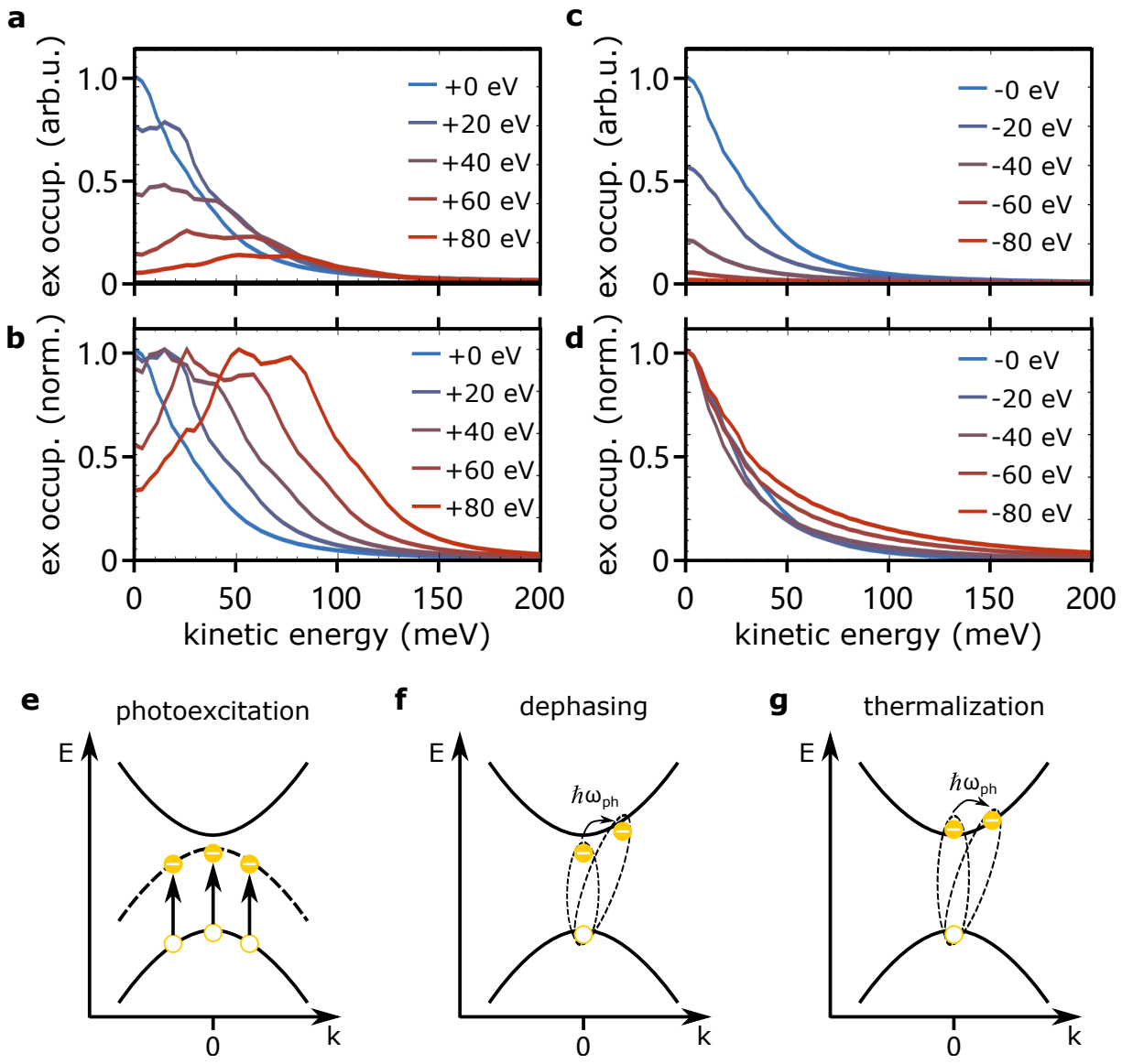
Supplementary Fig. 8**a** illustrates the snapshots of the exciton occupation directly at the

maximum of the pump pulse as a function of kinetic energy with selected detuning pump photon energy *above* the excitonic transition energy. The temperatures are set as room temperature for all the calculation. With increasing pump photon energy, the amount of injected excitons decreases due to the non-resonant excitation, eq. 14. However, at larger detunings, excitons occupy larger energy states due to the excess energy of the pump pulse, which is provided by acoustic and optical phonon transitions. The Supplementary Fig. 8**b** illustrates the exciton occupations but normalized to the maximum. Here it is even more obvious that the excitons obtain higher energies as the detuning increases. Interesting, for larger detunings two maxima can be observed, where the higher one originates from the formation of excitons via acoustic phonon scattering. The lower peak originates from the formation of excitons via optical phonon emission but also from relaxation of excitons from the higher peak via optical phonon emission.

Supplementary Fig. 8**c** illustrates snapshots of the exciton occupation directly at the maximum of the pump pulse as a function of kinetic energy for selected detunings *below* the excitonic transition energy. Similar to the previous scenario, the density of injected excitons decreases with increasing detuning due to the non-resonant excitation, eq. 14. In Supplementary Fig. 8**d**, the energy-dependent exciton occupations are normalized to their maximum. Interestingly, for pumping with larger negative detunings, the exciton distribution broadens such that the relative exciton occupation at large kinetic energies increases. The reason is, that for larger detunings the Lorentzian in equation 16 flattens which results in higher occupation of hot exciton at large energy range. As a consequence, for the near-resonant excitation below the excitonic transition, a substantial amount of excitons is formed at energies above 100 meV which contribute to the

Meitner-Auger scattering. To conclude, non-resonant excitation of the exciton, both above and below the resonance, introduce hot excitons with high kinetic energy, which are subjected to the Meitner-Auger IET.

In our experiment, the pump photon energy is $\hbar\nu_{pump} = 1.55\,\mathrm{eV}$ and the A-exciton transition energy is $E_{ex} = 1.63\,\mathrm{eV}$ determined by the energy difference of excited-state particles at CBM and VBM. After the photoexcition which prepares the *coherent* excitons with zero COM momentum (Supplementary Fig. 8**e**), the phonon-assisted dephasing process transfers the coherent excitons to *incoherent* exciton population which gain the finite COM momenta (Supplementary Fig. 8**f**). This dephasing process has been observed by our previous study[21]. The subsequent thermalization of excitons at the excitonic states (Supplementary Fig. 8**g**) also contribute to the nonzero COM momenta which is already included in our calculation.



Supplementary Fig. 8: Finite COM momentum of excitons at the K_{WSe_2} valley. **a**, The calculated energy-dependent hot exciton occupation with detuned pump photon energy above the resonant excitonic transition energy. **b**, Normalized hot exciton distributions in **a**. **c** The hot exciton occupation with detuned pump photon energy below the resonant excitonic transition energy and **d** is the corresponding normalized hot exciton distributions.

Schematic illustrations of involved ultrafast dynamics: **e**, photoexcitation creates the coherent exciton. The dash curve represents the coherent excitonic state. **f**, The phonon-assisted dephasing process transfers coherent excitons to incoherent excitons, at the same time, increases the COM momentum of excitons. **g**, The following thermalization also contributes to the finite COM momentum.

Förster coupling To calculate the Förster rate from WSe_2 to graphene, we start with the Hamiltonian

$$H_F = \sum_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}',\lambda,\lambda',\nu,\nu'} V^{\lambda\nu\nu'\lambda'}_{\mathbf{k},\mathbf{q},\mathbf{q}',\mathbf{k}'} \lambda^\dagger_{\mathbf{k}} \nu^\dagger_{\mathbf{q}} \nu'_{\mathbf{q}'} \lambda'_{\mathbf{k}'}. \tag{17}$$

As a convention, we use $\lambda^{(\prime)}$ as band indices and $\mathbf{k}^{(\prime)}$ as momenta in WSe_2 layer and $\nu^{(\prime)}$ as band indices and $\mathbf{q}^{(\prime)}$ as momenta in graphene. The appearing matrix element reads

$$V^{\lambda\nu\nu'\lambda'}_{\mathbf{k},\mathbf{q},\mathbf{q}',\mathbf{k}'} = \int_{\mathbb{R}^3} d^3r \int_{\mathbb{R}^3} d^3r' \Psi^{\lambda *}_{\mathbf{k}}(\mathbf{r}) \Psi^{\nu *}_{\mathbf{q}}(\mathbf{r}') V(\mathbf{r},\mathbf{r}') \Psi^{\nu'}_{\mathbf{q}'}(\mathbf{r}') \Psi^{\lambda'}_{\mathbf{k}'}(\mathbf{r}). \tag{18}$$

Here, $\Psi^{\lambda/\nu}_{\mathbf{k}/\mathbf{q}}$ account for the electronic Bloch waves in WSe_2 and graphene. The appearing Coulomb potential shall take into account the dielectric environment of the heterostructure, including the WSe_2 and graphene layer which are distanced by a gap with dielectric constant ϵ_R and width z (closely stacked structures have $z = 0$)[22]. Additionally we take substrates below and above the structure into account.

We can evaluate the matrix element by Fourier transforming the Coulomb potential and calculating the real space integrals within a $\mathbf{k}\cdot\mathbf{p}$ expansion[23]. We introduce exciton operators in WSe_2 $P^\mu_{\mathbf{Q}} = \sum_{\mathbf{q}} \varphi^\mu_{\mathbf{q}} c^\dagger_{\mathbf{q}+\frac{m_e}{m_h+m_e}\mathbf{Q}} v_{\mathbf{q}-\frac{m_h}{m_h+m_e}\mathbf{Q}}$ with quantum state μ and COM momentum $\mathbf{Q}$ as

well as pair operators in graphene $R^{\mathbf{q}}_{\mathbf{Q}} = c^{\dagger}_{\mathbf{q}+\frac{1}{2}\mathbf{Q}} v_{\mathbf{q}-\frac{1}{2}\mathbf{Q}}$. The Hamiltonian then reads

$$H_F = \sum_{\mathbf{Q},\mathbf{q},\mu} F^{\mu}_{\mathbf{Q}}(z) P^{\dagger\mu}_{\mathbf{Q}} R^{\mathbf{q}}_{\mathbf{Q}} + h.c. \tag{19}$$

The appearing coupling element reads

$$F^{\mu}_{\mathbf{Q}}(z) = \frac{1}{e^2\sqrt{A}} V_{\mathbf{Q}}(z) \varphi^{\mu}(\mathbf{r}=0) \mathbf{Q} \cdot \mathbf{d_T}^{cv} \mathbf{Q} \cdot \mathbf{d}^{vc}_{G} \tag{20}$$

with $\mathbf{d}^{vc}_{T}$ the dipole element in WSe_2, $\mathbf{d}^{vc}_{G}$ the dipole element in graphene, $\varphi^{\mu}(\mathbf{r})$ the excitonic wave function in real space with quantum number μ in WSe_2. We restrict our analysis to the lowest lying excitons $\mu = 1s$.

The Förster induced transition rate is given as [24]

$$\Gamma_{\mathbf{Q}}(z) = 4\pi \sum_{\mathbf{q}} |F_{\mathbf{Q}}(z)|^2 \delta\left(E^{\mathbf{q}}_{\mathbf{Q}} - E^{1s}_{\mathbf{Q}}\right), \tag{21}$$

where we have already added a factor of 2 to account for the electron spin in graphene. We analytically treat the summation over the delta function, where the area which appears in equation (20) cancels. We arrive at

$$\Gamma_{\mathbf{Q}}(z) = \frac{|F_{\mathbf{Q}}(z)|^2 A E^{1s}_{\mathbf{Q}}}{2\hbar^2 v_F^2}. \tag{22}$$

v_F is the Fermi velocity in graphene. The area A cancels with the area in $|F_{\mathbf{Q}}(z)|$. In a last step we average over the angle dependence of $|F_{\mathbf{Q}}(z)|^2$, und sum the result over the K and K' point in graphene, which is already included in the $\mathbf{q}$ summation in eq. 21. This way we arive at the final expression

$$\Gamma_{\mathbf{Q}}(z) = \frac{|V_{\mathbf{Q}}(z)|^2 |\varphi^{1s}(\mathbf{r}=0)| d_T^2 d_G^2 E^{1s}_{\mathbf{Q}} Q^4}{8\hbar^2 v_F^2 e^2}. \tag{23}$$

The Coulomb potential $V_{\mathbf{Q}}(z)$ is given as

$$V_{\mathbf{Q}}(z) = \frac{e^2}{\epsilon_0 |\mathbf{Q}| \epsilon_{\mathbf{Q}}(z)}, \tag{24}$$

where the momentum dependent dielectric function $\epsilon_{\mathbf{Q}}(z)$ accounts for the dielectric screening from the surrounding[22]. As input parameters, we require the thickness of graphene and WSe_2 layers and their respective dielectric constants. Note, that in the limit of infinitely thin films and a uniform background, our results coincides with our previous one[20]. The required parameters are listed in table 1.

Fig. 4**d** in the main context illustrates the Förster transfer rate as a function of COM momentum and for different WSe_2 - graphene distances. For $\mathbf{Q} = 0$ we find a vanishing Förster rate followed by a monotonous increase. The large $\mathbf{Q}$ behavior is dictated by the interplay of the momentum dependence of the Coulomb potential and the factor Q^4. For the closest stacking, i.e. 0.0 nm we find a peak transition rate of about 0.08 meV.

Dexter Coupling The IET process could also mediated by Dexter-type two-particle exchange, whose transfer rate is determined by the wave function overlap[25]. Starting point for the calculation is the Hamiltonian

$$H_D = \sum_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}'} V^{cvvc}_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}'} c^\dagger_{\mathbf{k}} v^\dagger_{\mathbf{q}} v_{\mathbf{k}'} c_{\mathbf{q}'} + h.c., \tag{25}$$

with the same conventions for the notation as for the calculation of the Förster transfer. The appearing coupling element is defined as

$$V^{cvvc}_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}'} = \int_{\mathbb{R}^3} d^3r \int_{\mathbb{R}^3} d^3r' \Psi^{c*}_{\mathbf{k}}(\mathbf{r}) \Psi^{v*}_{\mathbf{q}}(\mathbf{r}') V(\mathbf{r},\mathbf{r}') \Psi^{v}_{\mathbf{k}'}(\mathbf{r}') \Psi^{c}_{\mathbf{q}'}(\mathbf{r}). \tag{26}$$

The Coulomb potential is translational invariant in the in-plane direction, i.e. $V(\mathbf{r},\mathbf{r}') = V(\mathbf{r}_\parallel - \mathbf{r}'_\parallel, z, z')$. Fourier transforming the Coulomb potential w.r.t. the in-plane components, writing the electronic wave functions as Bloch waves and decomposing the spatial coordinates into one component inside the unit cell and one which addresses the unit cells $\mathbf{r} \to \mathbf{r} + \mathbf{R}$, yields for the coupling element

$$V^{cvvc}_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}'} = \frac{1}{A}\sum_{\mathbf{K}} \int_{uc} dz \int_{uc} dz' \chi^c(z)\chi^v(z')V_{\mathbf{K}}(z,z')\delta_{\mathbf{k},\mathbf{q}'+\mathbf{K}}\delta_{\mathbf{q},\mathbf{k}'-\mathbf{K}}, \tag{27}$$

with

$$\chi^\lambda(z) = \frac{1}{V_{uc}} \int_{uc} d^2 r_\parallel u^{*\lambda}_{TMD}(\mathbf{r}_\parallel, z) u^{\lambda}_{Graphene}(\mathbf{r}_\parallel, z) \tag{28}$$

To evaluate the coupling element further, we restrict ourselves to the case with vanishing interlayer spacing. We decompose the z and z' integration into two integrals over WSe_2 and graphene

$$\begin{aligned} V^{cvvc}_{\mathbf{k},\mathbf{q},\mathbf{k}',\mathbf{q}'} = \frac{1}{A}\sum_{\mathbf{K}} \delta_{\mathbf{k},\mathbf{q}'+\mathbf{K}}\delta_{\mathbf{q},\mathbf{k}'-\mathbf{K}} \times \\ \Bigg(\int_{WSe_2} dz \int_{WSe_2} dz' \chi^c(z)\chi^v(z')V_{\mathbf{K}}(z,z') \\ + \int_{WSe_2} dz \int_{Graphene} dz' \chi^c(z)\chi^v(z')V_{\mathbf{K}}(z,z') \\ + \int_{Graphene} dz \int_{WSe_2} dz' \chi^c(z)\chi^v(z')V_{\mathbf{K}}(z,z') \\ + \int_{Graphene} dz \int_{Graphene} dz' \chi^c(z)\chi^v(z')V_{\mathbf{K}}(z,z') \Bigg) \end{aligned} \tag{29}$$

Given that the Coulomb potential varies only weakly with each layer, we can replace the z/z' dependence by the position of the layer $z = z_{WSe_2}, z_{Graphene}$ in the Coulomb potential. This way we arrive at

$$V^{cvvc}_{\mathbf{k},\mathbf{q},\mathbf{k'},\mathbf{q'}} = \frac{1}{A}\sum_{\mathbf{K}} \delta_{\mathbf{k},\mathbf{q'}+\mathbf{K}}\delta_{\mathbf{q},\mathbf{k'}-\mathbf{K}}\times$$

$$\left(\chi^c_{WSe_2}\chi^v_{WSe_2}V_{\mathbf{K}}(z = z_{WSe_2}, z' = z_{WSe_2})\right.$$

$$+\chi^c_{WSe_2}\chi^v_{Graphene}V_{\mathbf{K}}(z = z_{WSe_2}, z' = z_{Graphene})$$

$$+\chi^c_{Graphene}\chi^v_{WSe_2}V_{\mathbf{K}}(z = z_{Graphene}, z' = z_{WSe_2})$$

$$\left.+\chi^c_{Graphene}\chi^v_{Graphene}V_{\mathbf{K}}(z = z_{Graphene}, z' = z_{Graphene})\right), \tag{30}$$

with

$$\chi^{\lambda}_{WSe_2/Graphene} = \int_{WSe_2/Graphene} dz\chi^{\lambda}(z), \tag{31}$$

i.e. the contribution of the wave function overlap of the band λ in the individual layers. Assuming, that the integration in both layers contributes equally to the wave function overlap between WSe_2 and graphene in conduction and valence band, i.e. $\chi^{\lambda}_{WSe_2} = \chi^{\lambda}_{Graphene} = \frac{1}{2}\chi^{\lambda}$, we obtain the final expression for the matrix element

$$V^{cvvc}_{\mathbf{k},\mathbf{q},\mathbf{k'},\mathbf{q'}} = \frac{1}{4A}\chi^c\chi^v\sum_{\mathbf{K}} \delta_{\mathbf{k},\mathbf{q'}+\mathbf{K}}\delta_{\mathbf{q},\mathbf{k'}-\mathbf{K}}V^{Dex}_{\mathbf{K}} \tag{32}$$

with

$$V^{Dex}_{\mathbf{K}} = (V_{\mathbf{K}}(z = z_{WSe_2}, z' = z_{WSe_2}) + V_{\mathbf{K}}(z = z_{WSe_2}, z' = z_{Graphene})$$

$$+V_{\mathbf{K}}(z = z_{Graphene}, z' = z_{WSe_2}) + V_{\mathbf{K}}(z = z_{Graphene}, z' = z_{Graphene}))\,. \tag{33}$$

Reinserting the result back into the Hamiltonian (and indexing the operators according to the layer, since due to the momentum conservation our convention breaks down) yields

$$H_D = \sum_{\mathbf{K},\mathbf{k},\mathbf{q}} \frac{1}{4A} \chi^c \chi^v V^{Dex}_{\mathbf{K}+\mathbf{k}-\mathbf{q}} c^{\dagger WSe_2}_{\mathbf{k}+\mathbf{K}} v^{\dagger Gr}_{\mathbf{q}-\mathbf{K}} v^{WSe_2}_{\mathbf{k}} c^{Gr}_{\mathbf{q}} + h.c.. \tag{34}$$

So far, the momenta are defined w.r.t. the Γ point in graphene and WSe_2. Redefining the coordinates $\mathbf{k} \to K^W + \mathbf{k}$ and $\mathbf{q} \to K^G + \mathbf{q}$ expresses them w.r.t. the K point in WSe_2/graphene. A projection on excitonic wave functions in WSe_2 yields

$$H_D = -\sum_{\mathbf{K},\mathbf{q},\nu} \left(\frac{1}{4\sqrt{A}} \chi^c \chi^v \sum_{\mathbf{k}} \varphi^{*\nu}_{K^W+\mathbf{k}} V^{Dex}_{K^W-K^G+\mathbf{K}+\mathbf{k}-\mathbf{q}} \right) P^{\dagger\lambda}_{\mathbf{K}} R^{K^G+\mathbf{q}}_{\mathbf{K}} + h.c.. \tag{35}$$

In the Dexter coupling element, the momentum distance between the K points in graphene and the TMD directly enters. We have $|K^G - K^W| \approx 3.6\,\mathrm{nm}^{-1}$. As a first approximation we can ignore the COM and relative momenta inside the Coulomb potential $V^{Dex}_{K^W-K^G+\mathbf{K}+\mathbf{k}-\mathbf{q}} \approx V^{Dex}_{K^W-K^G}$ and get a first estimate for the Dexter coupling

$$H_D = -\sum_{\mathbf{K},\mathbf{q},\nu} D_{\mathbf{K}} P^{\dagger\nu}_{\mathbf{K}} R^{K^G+\mathbf{q}}_{\mathbf{K}} + h.c., \tag{36}$$

with

$$D_{\mathbf{K}} = \frac{1}{4\sqrt{A}} \chi^c \chi^v \varphi^{*\nu}(\mathbf{r}=\mathbf{0}) V^{Dex}_{K^W-K^G}. \tag{37}$$

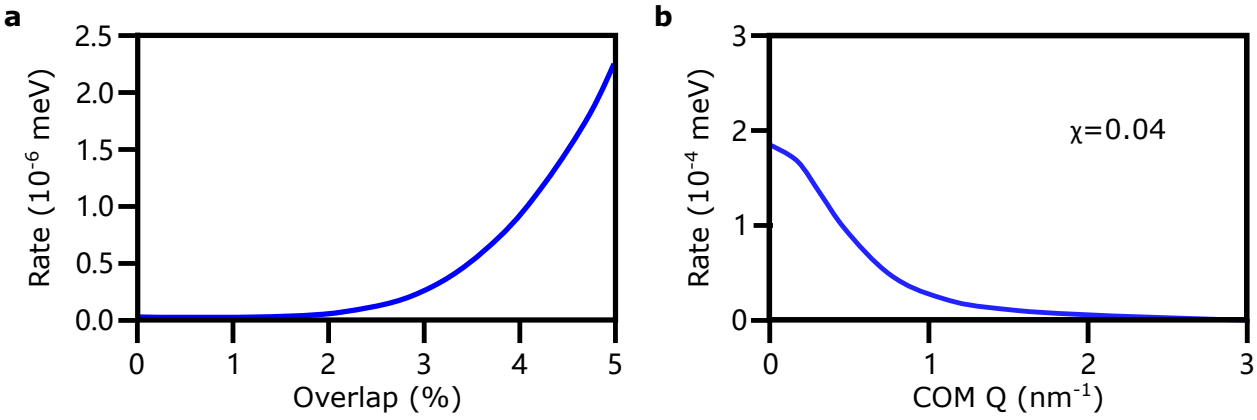
Similar to the Förster transfer, we can evaluate the Dexter induced scattering rate for WSe_2 excitons to graphene

$$\Gamma_{\mathbf{Q}} = 2\pi \sum_{\mathbf{q}} |D_{\mathbf{Q}}|^2 \delta(E^{\mathbf{q}}_{\mathbf{Q}} - E^{1s}_{\mathbf{Q}}) \tag{38}$$

with yields

$$\Gamma_{\mathbf{Q}} = \frac{|D_{\mathbf{Q}}|^2 A E_{\mathbf{Q}}}{4\hbar^2 v_F^2}. \tag{39}$$

Supplementary Fig. 9**a** illustrates the Dexter transfer rate from WSe_2 to graphene as a function of the wave function overlap for the same structure as considered for the Förster transfer. Assuming an overlap between the TMD and graphene wave functions of χ =0.039 (see estimation below, Supplementary Fig. 10**a**), we arrive at a Dexter rate of $1.0 \cdot 10^{-6}$meV. This number is small due to the mismatch of the K points of WSe_2 and graphene and due to the small overlap of the wave functions which enters with the fourth power. As long as the COM momentum $\mathbf{Q}$ is much smaller compared to the distance between the K points, the Dexter rate is independent of the COM momentum.



Supplementary Fig. 9: Estimation for the Dexter transfer rate. a, Dexter transfer

rate from WSe_2 to graphene as a function of the overlap of the electronic wave functions of graphene and WSe_2. **b**, Dexter transfer rate from WSe_2 to graphene as a function of the excitonic COM momentum $\mathbf{Q}$, with the $\mathrm{K}_{\mathrm{WSe_2}}$ valley and graphene being shifted on top of each others.

Maximum estimation for the Dexter process The reason for the very weak Dexter process is the large momentum mismatch between the K points in WSe_2 and graphene. This generates a momentum-bottleneck. To get an estimation of the maximally possible Dexter rate (and Dexter-like processes, such as Dexter-two-phonon processes), we remove this bottleneck artificially and move the K points of both layers on top of each other. The Hamiltonian reads

$$H_D = -\sum_{\mathbf{K},\mathbf{q},\nu} D_{\mathbf{q},\mathbf{K}} P^{\dagger\lambda}_{\mathbf{K}} R^{\mathbf{q}}_{\mathbf{K}} + h.c.. \tag{40}$$

with the coupling element

$$D_{\mathbf{q},\mathbf{K}} = \frac{1}{4\sqrt{A}} \chi^c \chi^v \sum_{\mathbf{k}} \varphi^{*\nu}_{\mathbf{k}} V^{Dex}_{\mathbf{K}+\mathbf{k}-\mathbf{q}}. \tag{41}$$

The relaxation rate of excitons to graphene is then given by

$$\Gamma_{\mathbf{K}} = \frac{A}{2\pi\hbar} \int_0^{2\pi} d\phi \frac{E_{\mathbf{K}}}{\hbar^2 v_F^2} |D_{\frac{E_{\mathbf{K}}}{\hbar v_F}(cos\phi, sin\phi)^T, \mathbf{K}}|^2. \tag{42}$$

Supplementary Fig. 9**b** illustrates the maximum estimation of the Dexter rate as the function of the COM momentum. We find a relatively weak rate in the order of 10^{-3} meV due to the poor wave function overlap. If we would set $\chi = 1$, we would obtain 70 meV at the maximum. Last, we calculate the thermal average of the Dexter rate, i.e. integrate the momentum dependent Dexter rate together with a normalized Boltzmann distribution at 300 K. We obtain $1.5 \cdot 10^{-4}$ meV.

Microscopic calculation of interlayer phonon-assisted tunneling process

In this section we will derive an expression for the phonon-assisted tunneling of carriers between the layers. The Hamiltonian of phonon scattering and tunneling can be generally written as

$$H = H_0 + H_1, \tag{43}$$

with H_0 accounting for the dispersion of electrons and phonons.

$$H_0 = \sum_a \epsilon^a a_a^\dagger a_a + \sum_b \hbar\omega^b b_b^\dagger b_b. \tag{44}$$

The first term accounts for the dispersion of carriers with operators $a_a^{(\dagger)}$ and the second term accounts for the dispersion of phonons with operators $b_b^{(\dagger)}$. The quantum numbers a, b account for layer and momentum of the carriers. The interaction Hamiltonian H_1 reads

$$H_1 = \sum_{ab} t^{ab} a_a^\dagger a_b + \sum_{abc} g^{abc} a_a^\dagger a_b (b_c + b_{-c}^\dagger), \tag{45}$$

where the first term represents the tunneling and the second term the scattering of carriers with phonons. Here the notation $-c$ implies, that the momentum has to be inverted, but all other quantum numbers stay the same.

While we are interested in the second order processes of phonon-assisted tunneling, we apply a canocical transformation to the Hamiltonian

$$H' = e^{-S} H e^S = H_0 + \underbrace{(H_1 + [H_0, S])}_{\text{first order}} + \underbrace{\frac{1}{2}[H_1, S]}_{\text{second order}}, \tag{46}$$

and claim that the first order in the interaction vanishes. This holds true for the choice

$$S = \sum_{ab} \alpha_{ab} t^{ab} a^\dagger a_b + \sum_{abc} g^{abc} a_a^\dagger a_b (\beta_{abc} b_c + \gamma_{abc} b_{-c}^\dagger), \tag{47}$$

with coefficients

$$\alpha_{ab} = \frac{1}{\epsilon^b - \epsilon^a}, \tag{48}$$

$$\beta_{abc} = \frac{1}{\epsilon^b - \epsilon^a + \hbar\omega^c}, \tag{49}$$

$$\gamma_{abc} = \frac{1}{\epsilon^b - \epsilon^a - \hbar\omega^c}. \tag{50}$$

The second order Hamiltonian is now given as

$$H_2 = \frac{1}{2}[H_1, S]. \tag{51}$$

Restricting ourselves only to the tunneling-phonon contribution (besides this, also higher order tunneling terms, two-phonon processes as well as attractive electron-electron interaction through phonon interaction are contained in this Hamiltonian) we obtain

$$\begin{aligned} H = &\frac{1}{2} \sum_{abcd} t^{db} g^{adc} a_a^\dagger a_b \left(\left(\frac{1}{\epsilon^b - \epsilon^d} - \frac{1}{\epsilon^d - \epsilon^a + \hbar\omega^c} \right) b_c + \left(\frac{1}{\epsilon^b - \epsilon^d} - \frac{1}{\epsilon^d - \epsilon^a - \hbar\omega^{-c}} \right) b_{-c}^\dagger \right) \\ &- \frac{1}{2} \sum_{abcd} t^{ad} g^{dbc} a_a^\dagger a_b \left(\left(\frac{1}{\epsilon^d - \epsilon^a} - \frac{1}{\epsilon^b - \epsilon^d + \hbar\omega^c} \right) b_c + \left(\frac{1}{\epsilon^d - \epsilon^a} - \frac{1}{\epsilon^b - \epsilon^d - \hbar\omega^{-c}} \right) b_{-c}^\dagger \right) \end{aligned} \tag{52}$$

Now we insert the compounds: $a = (\mathbf{k}_a, \lambda_a, l_a)$ for electrons, where $\mathbf{k}_a$ accounts for the momentum, λ_a accounts for the band and l_a accounts for the layer quantum number. For phonons

we insert the compounds $c = (\mathbf{k}_c, l_c, \xi_c)$, with momentum $\mathbf{k}_c$, layer l_c and branch ξ_c and apply the selection rules from the matrix elements:

$$t^{\lambda_b\lambda_d l_b l_d}_{\mathbf{k}_b\mathbf{k}_d} = t^{\lambda_b\lambda_d l_b l_d}\delta^{l_b\bar{l_d}}_{\mathbf{k}_b\mathbf{k}_d}\delta^{\lambda_b\lambda_d}, \tag{53}$$

i.e. the tunneling conserves momentum, and band but changes the layer index, and

$$g^{\lambda_a\lambda_d l_a l_d l_c}_{\mathbf{k}_a\mathbf{k}_d\mathbf{k}_c} = g^{\lambda_a\lambda_d l_a l_d l_c}_{\mathbf{k}_c}\delta^{l_a l_d}\delta^{l_d l_c}_{\mathbf{k}_c,\mathbf{k}_a-\mathbf{k}_d}\delta^{\lambda_a\lambda_d}. \tag{54}$$

i.e. phonon scattering conserves the layer index, but changes the momentum of the carriers.

As a result we obtain

$$\begin{aligned} H = &\frac{1}{2}\sum_{\mathbf{k}\mathbf{K}\lambda l\xi}\lambda^{\dagger\bar{l}}_{\mathbf{k}+\mathbf{K}}\lambda^{l}_{\mathbf{k}}\left(\underbrace{t^{\lambda l\bar{l}}g^{\lambda\bar{l}\xi}_{\mathbf{K}}(\alpha^{\lambda l\bar{l}}_{\mathbf{k}}+\gamma^{\lambda\bar{l}\xi}_{\mathbf{k},\mathbf{K}})}_{s^{\lambda l\bar{l}\xi}_{\mathbf{k},\mathbf{K}}}b^{\bar{l}\xi}_{\mathbf{K}}+\underbrace{t^{\lambda l\bar{l}}g^{\lambda\bar{l}\xi}_{\mathbf{K}}(\alpha^{\lambda l\bar{l}}_{\mathbf{k}}+\beta^{\lambda\bar{l}\xi}_{\mathbf{k},\mathbf{K}})}_{\tilde{s}^{\lambda l\bar{l}\xi}_{\mathbf{k},\mathbf{K}}}b^{\dagger\bar{l}\xi}_{-\mathbf{K}}\right)\\ &-\frac{1}{2}\sum_{\mathbf{k}\mathbf{K}\lambda l\xi}\lambda^{\dagger\bar{l}}_{\mathbf{k}+\mathbf{K}}\lambda^{l}_{\mathbf{k}}\left(\underbrace{t^{\lambda\bar{l}l}g^{\lambda l\xi}_{\mathbf{K}}(\alpha^{\lambda l\bar{l}}_{\mathbf{k}+\mathbf{K}}+\gamma^{\lambda l\xi}_{\mathbf{k},\mathbf{K}})}_{u^{\lambda l\bar{l}\xi}_{\mathbf{k},\mathbf{K}}}b^{l\xi}_{\mathbf{K}}+\underbrace{t^{\lambda\bar{l}l}g^{\lambda l\xi}_{\mathbf{K}}(\alpha^{\lambda l\bar{l}}_{\mathbf{k}+\mathbf{K}}+\beta^{\lambda l\xi}_{\mathbf{k},\mathbf{K}})}_{\tilde{u}^{\lambda l\bar{l}\xi}_{\mathbf{k},\mathbf{K}}}b^{\dagger l\xi}_{-\mathbf{K}}\right), \end{aligned} \tag{55}$$

with

$$\alpha_{\mathbf{k}}^{\lambda ij} = \frac{1}{\epsilon_{\mathbf{k}}^{\lambda i} - \epsilon_{\mathbf{k}}^{\lambda j}} \tag{56}$$

$$\beta_{\mathbf{k},\mathbf{K}}^{\lambda i\xi} = \frac{1}{\epsilon_{\mathbf{k}+\mathbf{K}}^{\lambda i} - \epsilon_{\mathbf{k}}^{\lambda i} + \hbar\omega_{-\mathbf{K}}^{i\xi}} \tag{57}$$

$$\gamma_{\mathbf{k},\mathbf{K}}^{\lambda i\xi} = \frac{1}{\epsilon_{\mathbf{k}+\mathbf{K}}^{\lambda i} - \epsilon_{\mathbf{k}}^{\lambda i} - \hbar\omega_{\mathbf{K}}^{i\xi}} \tag{58}$$

Both lines describe phonon assisted tunneling from $\mathbf{k}, b$ to $\mathbf{k}+\mathbf{K}, \bar{b}$. However, in the first line first the tunneling and next the phonon scattering takes place, whereas in the second line, first the phonon scattering and next the tunneling takes place. Considering the intermediate states$\mathbf{k}, \bar{b}$ and $\mathbf{k}+\mathbf{K}, b$ being much larger in energy, in both lines, all contributions have the same sign and add up. The relative sign between both lines $(-)$ is compensated by the opposite signs of the appearing u and s functions. To further evaluate the phonon assisted tunnel Hamiltonian, we carry out the summation over the layer

$$\begin{aligned}
H = &\frac{1}{2}\sum_{\mathbf{K}\lambda G\xi} \lambda_{\mathbf{k}+\mathbf{K}}^{\dagger W}\lambda_{\mathbf{k}}^{G}\left(\underbrace{t^{\lambda GW}g_{\mathbf{K}}^{\lambda W\xi}(\alpha_{\mathbf{k}}^{\lambda GW}+\gamma_{\mathbf{k},\mathbf{K}}^{\lambda W\xi})}_{s_{\mathbf{k},\mathbf{K}}^{\lambda GW\xi}} b_{\mathbf{K}}^{W\xi} + \underbrace{t^{\lambda GW}g_{\mathbf{K}}^{\lambda W\xi}(\alpha_{\mathbf{k}}^{\lambda GW}+\beta_{\mathbf{k},\mathbf{K}}^{\lambda W\xi})}_{\tilde{s}_{\mathbf{k},\mathbf{K}}^{\lambda GW\xi}} b_{-\mathbf{K}}^{\dagger W\xi}\right) \\
&-\frac{1}{2}\sum_{\mathbf{K}\lambda W\xi} \lambda_{\mathbf{k}+\mathbf{K}}^{\dagger G}\lambda_{\mathbf{k}}^{W}\left(\underbrace{t^{\lambda GW}g_{\mathbf{K}}^{\lambda W\xi}(\alpha_{\mathbf{k}+\mathbf{K}}^{\lambda WG}+\gamma_{\mathbf{k},\mathbf{K}}^{\lambda W\xi})}_{u_{\mathbf{k},\mathbf{K}}^{\lambda WG\xi}} b_{\mathbf{K}}^{W\xi} + \underbrace{t^{\lambda GW}g_{\mathbf{K}}^{\lambda W\xi}(\alpha_{\mathbf{k}+\mathbf{K}}^{\lambda WG}+\beta_{\mathbf{k},\mathbf{K}\xi}^{\lambda W})}_{\tilde{u}_{\mathbf{k},\mathbf{K}}^{\lambda WG\xi}} b_{-\mathbf{K}}^{\dagger W\xi}\right) \\
&+\frac{1}{2}\sum_{\mathbf{K}\lambda\xi} \lambda_{\mathbf{k}+\mathbf{K}}^{\dagger G}\lambda_{\mathbf{k}}^{W}\left(\underbrace{t^{\lambda WG}g_{\mathbf{K}}^{\lambda G\xi}(\alpha_{\mathbf{k}}^{\lambda WG}+\gamma_{\mathbf{k},\mathbf{K}}^{\lambda G\xi})}_{s_{\mathbf{k},\mathbf{K}}^{\lambda WG\xi}} b_{\mathbf{K}}^{G\xi} + \underbrace{t^{\lambda WG}g_{\mathbf{K}}^{\lambda G\xi}(\alpha_{\mathbf{k}}^{\lambda WG}+\beta_{\mathbf{k},\mathbf{K}}^{\lambda G\xi})}_{\tilde{s}_{\mathbf{k},\mathbf{K}}^{\lambda WG\xi}} b_{-\mathbf{K}}^{\dagger G\xi}\right) \\
&-\frac{1}{2}\sum_{\mathbf{K}\lambda\xi} \lambda_{\mathbf{k}+\mathbf{K}}^{\dagger W}\lambda_{\mathbf{k}}^{G}\left(\underbrace{t^{\lambda WG}g_{\mathbf{K}}^{\lambda G\xi}(\alpha_{\mathbf{k}+\mathbf{K}}^{\lambda GW}+\gamma_{\mathbf{k},\mathbf{K}}^{\lambda G\xi})}_{u_{\mathbf{k},\mathbf{K}}^{\lambda GW\xi}} b_{\mathbf{K}}^{G\xi} + \underbrace{t^{\lambda WG}g_{\mathbf{K}}^{\lambda G\xi}(\alpha_{\mathbf{k}+\mathbf{K}}^{\lambda GW}+\beta_{\mathbf{k},\mathbf{K}}^{\lambda G\xi})}_{\tilde{u}_{\mathbf{k},\mathbf{K}}^{\lambda GW\xi}} b_{-\mathbf{K}}^{\dagger G\xi}\right)
\end{aligned} \tag{59}$$

Here, the first two lines involve phonons in WSe_2 whereas the last two lines involve phonons

in graphene. The $-$ signs in front of line two and line four are compensated by the opposite signs of the s and u functions. To investigate the different contributions to the tunnel-phonon coupling in more detail, we evaluate the functions s and u in the different terms: (a) we fix the momenta in WSe_2 to the vicinity of the K point. (b) While all scattering processes conserve the energy, the approximation (a) settles the energy and momentum range of involved carriers. (c) In the next step, we approximate the prefactors $\alpha_{\mathbf{k}}^{\lambda ij}$ by their values in the region of interest, where we find $\alpha_{\mathbf{k}\approx K_W}^{cGW} = \frac{1}{1\text{eV}}$ and $\alpha_{\mathbf{k}\approx K_G}^{cWG} = \frac{1}{250\text{meV}}$, which are read out from the DFT calculation in the main manuscript. To evaluate the prefactors $\beta_{\mathbf{k},\mathbf{K}}^{\lambda i\xi}$ and $\gamma_{\mathbf{k},\mathbf{K}}^{\lambda i\xi}$ we first realize, that that ΔE_{K_G} appears whenever WSe_2 phonons are involved, and ΔE_{K_W} appears whenever graphene phonons are involved. $\Delta E_{K_G} \approx 250\,\text{meV}$ is large in comparison to typical phonon energies of 30 meV in WSe_2 and $\Delta E_{K_W} \approx 1\,\text{eV}$ is large in comparison to typical phonon energies of 200 meV in graphene. Consequently we ignore the appearing phonon energies in $\beta_{\mathbf{k},\mathbf{K}}^{\lambda i\xi}$ and $\gamma_{\mathbf{k},\mathbf{K}}^{\lambda i\xi}$. As a result, the Hamiltonian simplifies to

$$\begin{aligned} H = &\sum_{\mathbf{K}\mathbf{k}\lambda,\xi} \underbrace{\frac{t^{\lambda GW} g_{\mathbf{K}}^{\lambda W\xi}}{\epsilon_{K_{gr}}^{W} - \epsilon_{K_{gr}}^{G}}}_{h_{\mathbf{K}}^{W\xi}\approx const.} \lambda_{\mathbf{k}+\mathbf{K}}^{\dagger W} \lambda_{\mathbf{k}}^{G} \left(b_{\mathbf{K}}^{W\xi} + b_{-\mathbf{K}}^{\dagger W\xi} \right) + h.c. \\ &+ \sum_{\mathbf{K}\mathbf{k}\lambda\xi} \underbrace{\frac{t^{\lambda WG} g_{\mathbf{K}}^{\lambda G\xi}}{\epsilon_{K_W}^{G} - \epsilon_{K_W}^{W}}}_{h_{\mathbf{K}}^{G}\approx const.} \lambda_{\mathbf{k}+\mathbf{K}}^{\dagger W} \lambda_{\mathbf{k}}^{G} \left(b_{\mathbf{K}}^{G\xi} + b_{-\mathbf{K}}^{\dagger G\xi} \right) + h.c.\,. \end{aligned} \tag{60}$$

For a carrier in graphene, the relaxation rate to WSe_2 via phonon-assisted tunneling is given

as

$$\Gamma^G_{\mathbf{k}} = 2\pi \sum_{\pm,\mathbf{K},\xi,i\in\{W,G\}} |h^{i\xi}_{\mathbf{K}}|^2 \left(\frac{1}{2} \pm \frac{1}{2} + n^{i\xi}_{\mathbf{K}}\right) \delta(\epsilon^G_{\mathbf{k}} - \epsilon^W_{\mathbf{k}+\mathbf{K}} \mp \hbar\omega^{i\xi}_{\mathbf{K}}) \tag{61}$$

Assuming $h^{i\xi}_{\mathbf{K}} \approx h^{i\xi}$ and $\hbar\omega^{i\xi}_{\mathbf{K}} \approx \hbar\omega^{i\xi}$, we obtain

$$\Gamma^G_{\mathbf{k}} = A \sum_{i\in\{W,G\},\xi,\pm} \frac{m^W}{\hbar^2} |h^i|^2 \left(\frac{1}{2} \pm \frac{1}{2} + n^{i\xi}\right) 1_{\epsilon^G_{\mathbf{k}} \mp \hbar\omega^{i\xi} - \epsilon^W_{\mathbf{0}} > 0}, \tag{62}$$

which is constant for graphene electrons which have at least the energy of the conduction band plus the phonon energy in graphene. This reflects the constant density of states in WSe_2. The area A cancels with the area in the phonon coupling element which is contained in h^i.

In contrast for carriers initially located in WSe_2, we get

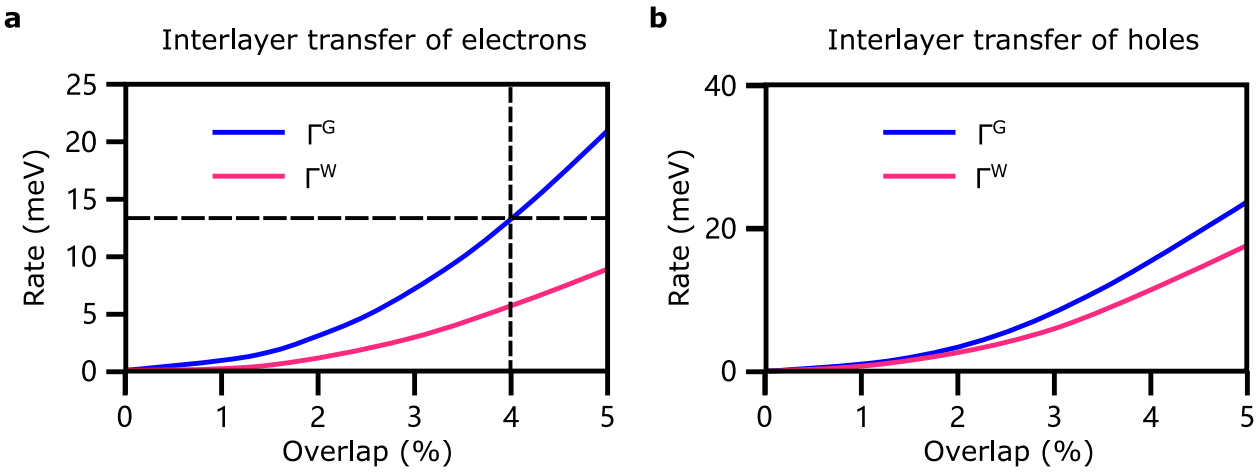
$$\Gamma^W_{\mathbf{k}} = 2\pi \sum_{\pm,\mathbf{K},\xi,i\in\{W,G\}} |h^{i\xi}_{\mathbf{K}}|^2 \left(\frac{1}{2} \pm \frac{1}{2} + n^{i\xi}_{\mathbf{K}}\right) \delta(\epsilon^W_{\mathbf{k}} - \epsilon^G_{\mathbf{k}+\mathbf{K}} \mp \hbar\omega^{i\xi}_{\mathbf{K}}). \tag{63}$$

With similar approximation as above, we end up at

$$\Gamma^W_{\mathbf{k}} = A \sum_{i\in\{W,G\},\xi,\pm} \frac{\epsilon_{\mathbf{k}} \mp \hbar\omega^{i\xi}}{\hbar^2 v_F^2} |h^i|^2 \left(\frac{1}{2} \pm \frac{1}{2} + n^{i\xi}\right), \tag{64}$$

To calculate the relaxation rates, we assume a potential barrier of E_B =5 eV corresponding to the energy of the WSe_2 conduction band w.r.t. to the vacuum level. Then we approximate the tunneling element as $t = \chi E_B$, with χ being the wavefunction overlap[22]. For graphene, we include two optical phonon branches with energies of $\hbar\omega$ =200 meV and the coupling strength of g =200 meV[26]. For WSe_2, we include two optical phonon branches with energies of $\hbar\omega$ =30 meV

and the coupling strength of $g = 10\,\mathrm{meV}$[27]. Supplementary Fig. 10**a** illustrates the approximate relaxation rates of electrons from graphene to WSe_2 (Γ^G) and WSe_2 to graphene (Γ^W) as a function of the overlap of the wavefunctions between WSe_2 and graphene. We find for both a quadratic increase as a function of the overlap, since the latter enters quadratic in both relaxation rates. The difference between both relaxation rates arise from different final densities of states of the carrier relaxation. In the experiment, a delayed rise of the WSe_2 signal w.r.t. to the graphene signal of about 50 fs (13 meV) was found, which indicates an overlap of 4.0%. The order of magnitude of this value appears reasonable, since the overlap between two neighboring graphene atoms is 7% as an example[28].



Supplementary Fig. 10: Tunnel transfer for electrons and holes. a, The ICT-induced relaxation rate of electrons in graphene towards WSe_2 (blue) as a function of the electronic wavefunction overlap between the involved conduction bands. The relaxation rate of electrons in WSe_2 towards graphene (pink). The dashed lines indicate the overlapping values which can be expected from the experiment results. **b**, The relaxation rate of

holes in graphene, which have larger energies compared to the valence band maximum in WSe_2 (blue) and the relaxation rate of holes in WSe_2 (pink).

To evaluate the phonon-assisted tunneling rates for holes, we assume similar phonon coupling elements in conduction and valence band in WSe_2[27] and graphene but account for the different dispersion of the valence band[29]. Supplementary Fig. 10**b** illustrates the tunneling rates of holes. We find qualitatively similar tunneling rates as for electrons. However, the tunneling of holes from WSe_2 to graphene is stronger compared to the electrons. This arises from the larger density of states of graphene for the involved final states. The reason for this, is that the Fermi energy of the system is closer to the conduction band minimum compared to the valence band maximum, cp. Fig. 2**a** in the manuscript.

Table 1: Parameters used in the computation. * exemplary value at $z = 0$ determined numerically by evaluating the Wannier equation for WSe_2 on a SiC substrate. [19,30]. ** taken as double distance between the chalcogen atoms. The Fermi velocity taken the experimental result.

Param.		Param.		Ref.
$\hbar$	0.658 eV fs	d_G	0.25 e nm	[31]
e	1 e	v_F	1.8 nm fs^{-1}	exp
ϵ_0	5.5·10^{-2} e^{2}eV^{-1}nm^{-1}	d_{WSe_2}	0.32 e nm	[19]
k_B	8.6·10^{-5} eV K^{-1}	$\lvert\varphi_{WSe_2}(\mathbf{r}=0)\rvert$	0.36 nm^{-1}	*
ϵ_{SiC}	9.6	$E^{1s}_{WSe_2}$	1.7 eV	[32]
		M_{WSe_2}	3.7 eVfs^{2}nm^{-2}	[33]
ϵ_G	6[34]	ϵ_{WSe_2}	13.36	
a_G	0.33 nm [35]	a_{WSe_2}	0.67 nm	** [36]

References

1. Zhao, W. *et al.* Lattice dynamics in mono-and few-layer sheets of WS_2 and WSe_2. *Nanoscale* **5**, 9677–9683 (2013).

2. Terrones, H. *et al.* New first order Raman-active modes in few layered transition metal dichalcogenides. *Sci. Rep.* **4**, 4215 (2014).

3. Tonndorf, P. *et al.* Photoluminescence emission and Raman response of monolayer MoS_2, $MoSe_2$, and WSe_2. *Opt. Express* **21**, 4908–4916 (2013).

4. He, K. *et al.* Tightly bound excitons in monolayer WSe_2. *Phys. Rev. Lett.* **113**, 026803 (2014).

5. Froehlicher, G., Lorchat, E. & Berciaud, S. Charge versus energy transfer in atomically thin graphene-transition metal dichalcogenide van der Waals heterostructures. *Phys. Rev. X* **8**, 011007 (2018).

6. Maklar, J. *et al.* A quantitative comparison of time-of-flight momentum microscopes and hemispherical analyzers for time-resolved ARPES experiments *Rev. Sci. Instrum.* **91**, 123112 (2020).

7. Nakamura, H. *et al.* Spin splitting and strain in epitaxial monolayer WSe_2 on graphene. *Phys. Rev. B* **101**, 165103 (2020).

8. Gierz, I., Henk, J., Höchst, H., Ast, C. R. & Kern, K. Illuminating the dark corridor in graphene: Polarization dependence of angle-resolved photoemission spectroscopy on graphene. *Phys. Rev. B* **83**, 121408 (2011).

9. Hwang, C. *et al.* Fermi velocity engineering in graphene by substrate modification. *Sci. Rep.* **2**, 590 (2012).

10. Siegel, D. A. *et al.* Many-body interactions in quasi-freestanding graphene. *PNAS* **108**, 11365–11369 (2011).

11. Yuan, L., Wang, T., Zhu, T., Zhou, M. & Huang, L. Exciton dynamics, transport, and annihilation in atomically thin two-dimensional semiconductors. *J. Phys. Chem. Lett.* **8**, 3371–3379 (2017).

12. Wang, H., Zhang, C. & Rana, F. Ultrafast dynamics of defect-assisted electron–hole recombination in monolayer MoS_2. *Nano Lett.* **15**, 339–345 (2015).

13. Dendzik, M. *et al.* Observation of an excitonic Mott transition through ultrafast core-*cum*-conduction photoemission spectroscopy. *Phys. Rev. Lett.* **125**, 096401 (2020).

14. Liu, F., Ziffer, M. E., Hansen, K. R., Wang, J. & Zhu, X. Direct determination of band-gap renormalization in the photoexcited monolayer MoS_2. *Phys. Rev. Lett.* **122**, 246803 (2019).

15. Ruppert, C., Chernikov, A., Hill, H. M., Rigosi, A. F. & Heinz, T. F. The role of electronic and phononic excitation in the optical response of monolayer WS_2 after ultrafast excitation. *Nano Lett.* **17**, 644–651 (2017).

16. Harb, M. *et al.* Carrier relaxation and lattice heating dynamics in silicon revealed by femtosecond electron diffraction. *J. Phys. Chem. B* **110**, 25308–25313 (2006).

17. Hong, J., Senga, R., Pichler, T. & Suenaga, K. Probing exciton dispersions of freestanding monolayer WSe_2 by momentum-resolved electron energy-loss spectroscopy. *Phys. Rev. Lett.* **124**, 087401 (2020).

18. Selig, M. *et al.* Ultrafast dynamics in monolayer transition metal dichalcogenides: interplay of dark excitons, phonons, and intervalley exchange. *Phys. Rev. Res.* **1**, 022007 (2019).

19. Selig, M. *et al.* Excitonic linewidth and coherence lifetime in monolayer transition metal dichalcogenides. *Nat. Commun.* **7**, 13279 (2016).

20. Selig, M. *et al.* Dark and bright exciton formation, thermalization, and photoluminescence in monolayer transition metal dichalcogenides. *2D Mater.* **5**, 035017 (2018).

21. Dong, S. *et al.* Measurement of an excitonic wave function. *arXiv:2012.15328* (2020).

22. Ovesen, S. *et al.* Interlayer exciton dynamics in van der Waals heterostructures. *Commun. Phys.* **2**, 1–8 (2019).

23. Katsch, F., Selig, M. & Knorr, A. Theory of coherent pump–probe spectroscopy in monolayer transition metal dichalcogenides. *2D Mater.* **7**, 015021 (2019).

24. Selig, M., Malic, E., Ahn, K. J., Koch, N. & Knorr, A. Theory of optically induced Förster coupling in van der Waals coupled heterostructures. *Phys. Rev. B* **99**, 035420 (2019).

25. Dexter, D. L. A theory of sensitized luminescence in solids. *J. Chem. Phys.* **21**, 836–850 (1953).

26. Piscanec, S., Lazzeri, M., Mauri, F., Ferrari, A. & Robertson, J. Kohn anomalies and electron-phonon interactions in graphite. *Phys. Rev. Lett.* **93**, 185503 (2004).

27. Jin, Z., Li, X., Mullen, J. T. & Kim, K. W. Intrinsic transport properties of electrons and holes in monolayer transition-metal dichalcogenides. *Phys. Rev. B* **90**, 045422 (2014).

28. Reich, S., Maultzsch, J., Thomsen, C. & Ordejon, P. Tight-binding description of graphene. *Phys. Rev. B* **66**, 035412 (2002).

29. Kormányos, A., Zólyomi, V., Drummond, N. D. & Burkard, G. Spin-orbit coupling, quantum dots, and qubits in monolayer transition metal dichalcogenides. *Phys. Rev. X* **4**, 011034 (2014).

30. Berghäuser, G. & Malic, E. Analytical approach to excitonic properties of MoS_2. *Phys. Rev. B* **89**, 125309 (2014).

31. Malic, E., Winzer, T., Bobkin, E. & Knorr, A. Microscopic theory of absorption and ultrafast many-particle kinetics in graphene. *Phys. Rev. B* **84**, 205406 (2011).

32. Christiansen, D. *et al.* Phonon sidebands in monolayer transition metal dichalcogenides. *Phys. Rev. Lett.* **119**, 187402 (2017).

33. Kormányos, A. *et al.* k· p theory for two-dimensional transition metal dichalcogenide semiconductors. *2D Mater.* **2**, 022001 (2015).

34. Bessler, R., Duerig, U. & Koren, E. The dielectric constant of a bilayer graphene interface. *Nanoscale Adv.* **1**, 1702–1706 (2019).

35. Ye, S. *et al.* Thickness-dependent strain effect on the deformation of the graphene-encapsulated Au nanoparticles. *J. Nanomater.* **2014** (2014).

36. Berkelbach, T. C., Hybertsen, M. S. & Reichman, D. R. Theory of neutral and charged excitons in monolayer transition metal dichalcogenides. *Phys. Rev. B* **88**, 045318 (2013).