Picosecond energy transfer in a transition metal dichalcogenide–graphene heterostructure revealed by transient Raman spectroscopy

Carino Ferrante, Giorgio Di Battista, Luis E. Parra López, Giovanni Batignani, Etienne Lorchat, Alessandra Virga, Stéphane Berciaud, and Tullio Scopigno

Abstract

Intense light–matter interactions and unique structural and electrical properties make van der Waals heterostructures composed by graphene (Gr) and monolayer transition metal dichalcogenides (TMD) promising building blocks for tunneling transistors and flexible electronics, as well as optoelectronic devices, including photodetectors, photovoltaics, and quantum light emitting devices (QLEDs), bright and narrow-line emitters using minimal amounts of active absorber material. The performance of such devices is critically ruled by interlayer interactions which are still poorly understood in many respects. Specifically, two classes of coupling mechanisms have been proposed, charge transfer (CT) and energy transfer (ET), but their relative efficiency and the underlying physics are open questions. Here, building on a time-resolved Raman scattering experiment, we determine the electronic temperature profile of Gr in response to TMD photoexcitation, tracking the picosecond dynamics of the G and 2D Raman bands. Compelling evidence for a dominant role of the ET process accomplished within a characteristic time of ~4 ps is provided. Our results suggest the existence of an intermediate process between the observed picosecond ET and the generation of a net charge underlying the slower electric signals detected in optoelectronic applications.

Significance

Hot carrier–based energy harvesting is critically implied in the performances of optoelectronic devices based on van der Waals heterostructures composed by graphene (Gr) and monolayer transition metal dichalcogenides (TMD). The way electron–hole couples initially photogenerated in the TMD are converted into an electric current in Gr is a controversial issue. In this work we identify the interlayer interaction occurring during the first picoseconds following photoexcitation as an energy transfer process that is much faster than (other) photogating phenomena implied in optoelectronic applications.
TMD/Gr junctions. This mechanism, however, becomes quite inefficient when the active TMD region is thinned down to a monolayer (4), likely due to picosecond nonradiative transfer to Gr.

A key technique for characterizing vdW heterostructures is Raman spectroscopy. The Raman spectrum of Gr is dominated by two features, known as the G and 2D modes, near 1,580 and 2,660 cm\(^{-1}\), respectively. The G- and 2D-mode processes arise from nonresonant inelastic scattering with one zone-center longitudinal and transverse optical phonon and from a resonant process involving a pair of near–zone edge transverse optical phonons, respectively (32–35). Both modes are sensitive to external perturbations and in particular to doping (36–38) and temperature rises (39–41). For instance, the integrated intensity of the 2D-mode feature \(I_{2D}\) is routinely used to determine the Fermi level in Gr (34, 38, 42) as well as of the defect density (43) and the electronic temperature (41, 44) as summarized in Eq. 1. Building upon this sensitivity, quantitative Raman-based methods have been developed to characterize Gr-based systems. Recently, a combination of photoluminescence (PL) spectroscopy and steady state spontaneous Raman spectroscopy have revealed the interplay between slow, extrinsic photoinduced net charge transfer and picosecond energy transfer, pointing toward the dominant role of energy transfer in TMD–Gr heterostructures (14), without, however, determining the associated temporal dynamics.

Extending such sensing capabilities to the out-of-equilibrium regime would offer a unique way to address the ultrafast response of Gr following photoexcitation of the TMD above its optical bandgap. Indeed, both charge and energy transfer processes may in principle generate hot carriers in Gr, leading to an increase in its electronic temperature \(T_e\), which can be quantified by monitoring hot PL and out-of-equilibrium Raman scattering. The former directly yields \(T_e\) (45), while the latter is sensitive to both transient Fermi level shifts and temperature rises (46, 47). Assuming that photoexcitation of the TMD–Gr heterostructure predominantly yields a net charge flow (i.e., charge transfer) across the coupled layers, a long-lived doped state is expected to develop in Gr after electron–phonon relaxation, affecting the line shape and the intensity of the G-mode and 2D-mode features. In contrast, in an energy transfer-dominated regime, Gr would only exhibit a transient increase of \(T_e\), the dynamics of which is governed by the interplay between energy transfer and phonon relaxation of the electron-hole (e-h) pairs injected in Gr occurred in the subpicosecond timescale (12, 45, 48, 49). This scenario would result in a subpicosecond decrease of \(I_{2D}\) combined with a broadening of the G-mode feature. Charge and energy transfer can thus be distinguished based on the transient dynamics of the Raman features of Gr.

Here we provide direct mechanistic and temporal insights into interlayer coupling in monolayer tungsten disulfide (WS\(_2\))–Gr heterostructures by exploiting recent developments in ultrafast Raman spectroscopy (44, 46, 47, 50). Specifically, extracting the temporal profile of \(T_e\) in Gr following optical excitation of the WS\(_2\) monolayer slightly above its optical bandgap, we demonstrate highly efficient energy transfer from TMD to Gr in a 4-ps timescale, with no evidence for a net charge transfer.

**Results**

Our model system to track ultrafast interlayer transfers is a vdW heterostructure made from a Gr monolayer stacked onto WS\(_2\) monolayer supported by a silica substrate. In our experiments, the sample is probed using 1-ps pulses at 1.58 eV, i.e., around 400 meV below the WS\(_2\) optical bandgap. The role of the probe pulse is twofold: 1) it induces a combination of linear and nonlinear PL, which we use as an imaging tool for preliminary mapping the heterostructure, and 2) it serves as a time-delayed Raman-inducing beam following photoexcitation by 1-ps pulses at 2.07 eV, slightly above the WS\(_2\) optical bandgap. The time duration of the probe pulse is imposed by the necessary trade-off between a pulse duration \(\delta t\) as short as the exciton lifetime in the WS\(_2\) and a narrow excitation bandwidth \(\delta \nu\) required to yield a sufficient spectral resolution (these two quantities are fundamentally constrained by the Fourier relation: \(\delta \nu \delta t \geq 14.7 \text{ cm}^{-1}\text{ps}\)). The experiment concept is shown in Fig. 1 A and D and further detailed in *Materials and Methods*.

**PL.** PL spectroscopy is one of the main ways to characterize the TMD–Gr coupling. Bare TMD monolayers exhibit indeed strong PL: the radiative recombination upon optical excitation is favored by the presence of a direct bandgap in absence of competitive nonradiative decay channels (51, 52). Conversely, Gr does not exhibit significant PL (53) as efficient nonradiative decay channels (electron–phonon coupling) are active at any excitation wavelength owing to its gapless band structure (54). Accordingly, massive quenching of the TMD PL indicates efficient near-field coupling to Gr (4, 14–16, 55) (see also *SI Appendix*, Fig. S4).

A different scenario arises when photoexciting Gr with ultra-short laser pulses. Under large photon fluxes, the nonradiative recombination channel is significantly depleted, and hot PL (albeit weaker than the TMD PL) is observed from Gr (45, 56). Here we take advantage of such transient regime to introduce an imaging protocol, based on the PL detection upon pulsed laser excitation, for simultaneously mapping Gr and WS\(_2\). Specifically, photoexciting the WS\(_2\)–Gr heterostructure below the WS\(_2\) optical bandgap using our probe pulse, the WS\(_2\) PL stems from a lower-efficiency two-photon absorption process (*SI Appendix*, Fig. S6), whose intensity is comparable (as shown in Fig. 1C) with the Gr hot PL and is emitted in a different spectral region.

By recording a map of the integrated PL intensity in distinct spectral regions selected to maximize the contrast (Fig. 1B), we are able to isolate the regions with Gr and WS\(_2\) only, as well as the heterostructure with strongly and weakly coupled areas. Such simultaneous mapping, not achievable with continuous wave (cw) excitation, is a powerful method for imaging Gr-based heterostructures. The green contour of Fig. 1B highlights the region of coupled WS\(_2\)–Gr, characterized by PL quenching relative to bare WS\(_2\) and by the presence of hot PL from the coupled Gr monolayer (Fig. 1C). The Raman measurements presented hereafter are performed in this coupled region of the heterostructure.

**Time-Resolved Raman Scattering.** In our experiments the Raman signal induced by probe pulses is used to monitor Gr response following optical excitation by the higher-energy pump photons absorbed by the WS\(_2\) monolayer. Probe pulses also lead to an out-of-equilibrium increase of the electronic temperature \(T_e^{\text{eq}} \approx 1,550\) K in Gr (*Materials and Methods*), due to the e-h pairs injected upon absorption \(A_{\text{eq}} \approx 2.3\%\).

In contrast, the pump pulses are mainly absorbed by the WS\(_2\) layer [optical absorptance \(A_X \approx 5.4\%\) at 2.07 eV in the same preparation condition (57, 58)], generating excitons in WS\(_2\). However, a fraction of the pump pulse is absorbed also by the Gr layer \([A = (1 - A_X)A_{\text{Gr}} \approx 2.2\%\) and induces a photoinduced electronic heating of Gr (45) (for 70 µW, \(T_e < 1,400\) K [*SI Appendix*, Fig. S8]), implying an e-h pair density \(n_{\text{eh}} < 4.5 \times 10^{11}\) cm\(^{-2}\) [*SI Appendix*, Fig. S9]. Importantly, probe pulses have a negligible interaction with WS\(_2\), as testified by the weak two-photon fluorescence (*SI Appendix*, Fig. S6).
Fig. 1. Optical characterization of a WS2–Gr heterostructure. (A) Time-delayed pump and probe beams focused onto a diffraction-limited WS2–Gr spot. (B) PL map generated by probe only (1.58 eV) after integrating the emission spectra in the ranges 1.88 to 2.05 eV (WS2, two-photon PL) and 1.65 to 1.85 eV (Gr, hot luminescence) and the corresponding optical image (Right). Contour lines indicate bare Gr (blue), few-layer Gr (yellow), WS2 (red), coupled (green), and uncoupled (black) WS2–Gr regions. (C) PL spectra for the spots indicated in the map with corresponding colors. (D) Time-resolved, probe-generated Raman spectra of Gr modes in the WS2–Gr heterostructure for two time delays, corrected from hot PL background (Materials and Methods).

Discussion

The PL quenching of WS2 may stem either from charge transfer or from energy transfer. Here we numerically simulate the Raman experimental results considering only energy transfer from WS2 to Gr with near-unity efficiency.

We first focus on the time-resolved Raman response of bare Gr. In this case, the fast drop and recovery of $I_{2D}$ can be ascribed to the e–h pair density instantaneously photogenerated upon pump absorption and relaxing on an ultrafast (resolution-limited) timescale, followed by a slower D phonon heating. In order to isolate the latter, the Gr experimental profiles have been preliminarily fitted as the sum of a Gaussian dip and an exponential saturation term (Fig. 2), which takes into account the electronic and subsequent lattice heating, respectively. The latter ($\Delta I_{2D}^{\text{ph}}$), expected to be weakly sensitive to the WS2 coupling, is reported in Fig. 2G. The genuine modification of the 2D area induced by the electronic heating ($I_{2D}^e$) is hence obtained for both Gr and WS2–Gr by subtracting the phononic term, as $I_{2D} = I_{2D}^e - \Delta I_{2D}^{\text{ph}}$.

In order to quantitatively evaluate the e–h pair density injected by the probe pulse and the corresponding $T_e$ increase, the relation between the Gr electronic properties and the $I_{2D}$ can be exploited (42):

$$I_{2D}^e \propto \left( \frac{\gamma_K(E)}{\gamma(K, T_e)} \right)^2 = \left( \frac{\gamma_K(E)}{\gamma(K) + \gamma_T(E) + \gamma_{\text{def}}(E) + \gamma_{\text{ee}}(E, T_e)} \right)^2,$$

where $E$ denotes the energy of the electronic state and $\gamma$ is the scattering rate of electrons and holes, equal to the sum of $\gamma_K$, $\gamma_T$, $\gamma_{\text{def}}$, and $\gamma_{\text{ee}}$, which are the scattering rate of electrons with 2D-mode and G-mode phonons near the $K$ and $\Gamma$ points of the Brillouin zone, defects, and other electrons, respectively. Notably, $\gamma_{\text{ee}}$ is the only term that depends on $T_e$ (36, 37, 39, 42, 43, 59). Specifically (60),

$$\gamma_{\text{ee}}(E, T_e) = \alpha T_e,$$

with $\alpha = 0.51$ cm$^{-1}$ K$^{-1}$ for $E > 100$ meV (46).

From the full width at half maximum (FWHM) of 2D mode ($\Gamma_{2D}$) measured without pump pulse, we can estimate $\gamma \approx 195$ meV for WS2–Gr at $T_e = T_{2D}^{\text{ph}}$ (35). In the presence of pump pulse, an increase of $T_e$ is expected. However, for negative or large pump–probe delays (> 4 ps) the electronic heating of Gr...
induced by the direct interaction with pump pulse is not observed within the probe pulse duration due to the ps recombination of the e–h pairs. The electronic heating effect can only be observed around the time overlap of pump and probe pulse, generating a drop in $I_{2D}$ (Fig. 2G). Taking into account this experimental evidence and Eq. 1, the temporal profile of $T_e$ in WS$_2$–Gr and in bare Gr can be extracted (Fig. 3A and Materials and Methods).

As shown in Fig. 3A, the additional energy flux injected through energy transfer from the WS$_2$ layer to Gr leads to a different $T_e$ profile as compared to reference measurements in bare Gr, taking into account that the energy associated with a WS$_2$ exciton transfer is $\sim 2$ eV, and using the Fermi distribution with the density of states in Gr, we can derive the relation between the additional e–h pair density induced by the pump pulse and the relative increase of $T_e$ above $T_e^{\text{ref}}$ (right axis of Fig. 3A and Materials and Methods).

At this point, we model the dynamics of e–h pairs in Gr taking into account the transfer pathways reported in Fig. 3B. The pump pulse creates a population of e–h pairs in the WS$_2$ and Gr monolayers proportional to their absorbances. The WS$_2$ layer with Gr on top exhibits a fast decay channel ($\tau_T$), which inhibits its much longer decay $\tau_0 \sim 100$ ps $- 1$ ns (14–16, 61) and exciton–exciton recombination (6). Due to the suppression of this slow linear recombination in bare WS$_2$, the temporal evolution of the exciton density in WS$_2$, $N_X(t)$, can be described by the following differential equation:

$$\frac{dN_X}{dt} = -\frac{N_X}{\tau_T} + A_X I_{\text{PUP}}(t),$$

where $I_{\text{PUP}}(t)$ is the temporal profile of the pump pulse, assumed Gaussian with a FWHM of 1 ps.

On the other hand, Gr exhibits an intrinsic decay time for the thermalization of the electronic system with the phononic bath and substrate ($\tau_G$). Interlayer energy transfer induces e–h pair injection in Gr ($N_G(t)$), leading to a rise in $T_e$. Its temporal evolution can be expressed as

$$\frac{dN_G}{dt} = -\frac{N_G}{\tau_T} + \frac{N_X}{\tau_G} + A I_{\text{PUP}}(t),$$

where the coupling term $N_X/\tau_G$ obviously vanishes in the bare Gr case. Performing a best fit of the experimental data reported in Fig. 3A with $N_G(t)$ obtained from Eqs. 3 and 4, convoluted with the 1-ps probe pulse temporal profile, we obtain the black and blue solid lines for bare Gr and WS$_2$–Gr, respectively, with characteristic times $\tau_G \approx 0.6$ ps and $\tau_T \approx 4.3$ ps.

The decay time $\tau_G$ in bare Gr is close to the temporal resolution of our setup and is consistent with previous reports on Gr (62–64). Most importantly, we are able to determine the relevant energy transfer timescale from WS$_2$ to Gr at room temperature. Remarkably, $\tau_T \approx 4$ ps is longer than $\tau_G$, indicating that energy transfer–mediated e–h pair injection in Gr (the term $N_X/\tau_G$ in Eq. 4) is responsible for the longer decay of $N_G(t)$ observed in WS$_2$–Gr.

Since we are investigating a system separated by a subnanometer vdW gap, Förster and Dexter energy transfer are both likely to occur. The $\tau_T$ value that we directly determine here, however, provides a possible rationale to previous observations of PL quenching (14) and TA spectroscopy (15–17) and agrees with theoretical calculations based on a dominant Förster-type energy transfer mechanism (25).

Notably, the WS$_2$ PL quenching and the $I_{2D}$ drop (42) may be compatible with a charge transfer mechanism involving a net transient flux of electrons (or holes) to Gr. Critically, in this scenario, the doping generated in Gr upon charge transfer would be expected to relax on a timescale much longer than the picosecond dynamics reported in Fig. 2G (14, 31). Moreover, the measured temporal profile of the G band FWHM is also not compatible with a charge transfer–dominated scenario, as detailed in SI Appendix.

The energy transfer nature of the process established here suggests that the photogating effects (14, 28) and associated photodetection capabilities demonstrated in TMD–Gr-based optoelectronic devices (27, 31) result from an additional conversion mechanism into a much slower, less efficient, net charge transfer. Moreover, this evidence confirms the dominant contribution of energy transfer, measured with Raman spectroscopy (14), and at the same time the ps simultaneous transfer of e–h of the exciton provides an explanation for the low internal quantum efficiency of photocurrent generation in photodetectors made from TMD monolayers (4).

**Conclusion**

We have measured the time-resolved Raman spectrum of Gr in a WS$_2$–Gr heterostructure upon resonant photoexcitation of band-edge WS$_2$ excitons. By comparing the Raman response...
of bare and WS2-coupled Gr, we have unveiled a slower relaxation of the 2D mode integrated intensity occurring on a 4-ps timescale in WS2–Gr. Our experimental data are rationalized by a kinetic model based solely on energy transfer. Although photocurrent generation occurs with relatively high efficiency in vertically biased TMD–Gr heterojunctions (4, 65) made from few-layer TMD films, our results strongly suggest that net photoinduced charge transfer has negligible efficiency in ultimately thin monolayer TMD–Gr heterostructures, explaining the low efficiency of associated photodevices. Instead, we demonstrate that a TMD monolayer can boost carrier injection in Gr through picosecond energy transfer, with a room-temperature transfer yield approaching unity considering the longer excitonic lifetimes in bare TMD monolayers. Besides showing the possibility to exploit photothermionic effects for efficient hot carrier injection from Gr to TMDs under subbandgap excitation (5, 17), our results provide an essential step for the microscopic understanding of ultrafast interlayer coupling in TMD–Gr heterostructures, whose relevance for optoelectronics is thus consolidated. They also indicate as a key technological challenge the efficient funneling of hot carriers generated in Gr before they release their energy into heat, either through Gr phonon emission or through coupling with the substrate.

Materials and Methods

Sample Preparation. Our WS2–Gr heterostructure was fabricated as described previously (14) using a dry, viscoelastic transfer technique (66). In brief, a WS2 monolayer and a Gr monolayer were mechanically exfoliated from bulk crystals and stacked onto a fused silica substrate using a home-built transfer station. The sample was characterized using cw PL and Raman spectroscopies in ambient conditions (see SI Appendix, Figs. S4 and S5, respectively) before performing time-resolved studies.

PL Measurements. The experimental setup is the same as the time-resolved Raman experiment, but in this case, only the probe pulse is sent on the sample. The PL spectra reported in Fig. 1 B and C are collected in 2 s per pixel, and the sample scan is performed with two mechanical stages. This experimental scheme allows us to perform a simultaneous imaging of the two monolayers. Moreover, using the nonlinear two-photon PL for WS2, the spatial resolution is higher as compared to measurements in the cw regime performed using the same optical elements.

Gr, having no band gap (67), is not expected to have any radiative decay channel for charge carriers. However, under high excitation densities it emits light (45) over a wide spectral range owing to the inhibition of a nonradiative recombination channel. Indeed, in an out-of-equilibrium configuration, the large density of charge carriers in the conduction band cannot fully relax down to the Fermi level via electron-phonon decay pathways. Thus, the electronic subsystem also relaxes radiatively through hot PL which is well described by Planck’s law. Hence, in order to extract the electronic temperature $T_e$, we fit the hot PL spectra of Gr with Planck’s law (black lines in SI Appendix, Fig. S7):

$$I(\hbar\omega, T_e) = \eta_{\text{em}} T_e \exp \left( \frac{\hbar \omega}{k_B T_e} \right)$$

where $\eta_{\text{em}}$ is the emissivity of the blackbody, $T_e$ is the time emission of the black-body, and $E(\hbar\omega)$ is the responsivity of the detection chain taking into account the efficiency of the charge-coupled device, which depends on the photon energy. $\hbar\omega$ is the energy of the emitted photon, $T_e$ is the electronic temperature, and $k$ is Boltzmann’s constant.

In SI Appendix, Fig. S7, we show the hot PL spectra obtained performing a probe pulse only experiment on Gr (b) and on WS2–Gr (a) at different powers of the probe pulse. The peak located near 1.8 eV is present in both Gr and in WS2–Gr and is assigned to the anti-Stokes Raman G mode, while the broad peak at around 2 eV is assigned to PL from the A-exciton of WS2 excited by a two-photon absorption process.

In SI Appendix, Fig. S8, we show the temperature extracted from fit as a function of PRP powers for both bare Gr (diamond) and WS2–Gr (square), from the same figure, we estimate that the probe pulses (with power of 210 μW) used in our time-resolved Raman scattering experiment lead to an electronic temperature $T_e \sim 1,550$ K. The temporally overlapped pump pulses induce an additional temperature increase, which originated from both direct electronic heating of Gr and energy transfer from WS2.

Time-Resolved Raman Experiments. We use a two-module Toptica FemtoFiber Pro source as in ref. 50. This setup is able to generate, at a repetition rate of 40 MHz, 1-ps probe pulses at 1.58 eV (PRP) and a supercontinuum (SC) output between 0.89 and 1.38 eV. The SC spectral intensity can be tuned with a motorized Si–prism pair compressor. A periodically poled lithium niobate crystal with a fan-out grating (a poling period changing along the transverse direction) is exploited to produce broadly tunable (from 1.97 to 2.30 eV) narrowband 1-ps pump pulses, with a power < 10 mW (68). A dichroic mirror is used to combine the two beams, whose relative temporal delay is tuned with a motorized optical delay line. A long-working distance 20× objective (numerical aperture NA = 0.4) focuses the pulses onto the same ~2 μm spot on the sample. Before objective, the probe and pump power are 210 and 70 μW, respectively. The Raman emission on the Stokes side is collected by a condenser (numerical aperture 0.75), and the pump-probe pulses are filtered out using interference filters. The Raman spectrum is measured with a monochromator (Acton Spectra Pro-2500i) coupled to a cooled charge-coupled device array (CCD Princeton Instruments Pixis 100).

In Fig. 2, the Stokes Raman signals of Gr generated by PRP are reported at several time delays, first on Gr and then on WS2–Gr. For each delay we acquire Raman spectra for 10 min and repeat the acquisition six times. In SI Appendix, Fig. S10, we show a typical spectrum in WS2–Gr with and without pump. Raman spectra
are affected by an intense background due to the substrate. In order to remove this background, we collect, with same acquisition time, the signal spectrum from the substrate, and we subtract it (with a scaling of the intensity) from the Raman signal. In **SI Appendix, Fig. S11A**, we show two selected spectra (pump off and 0.1 ps) in which we have removed the substrate background. In order to better clean data we remove a baseline (red) and get the spectra as in **SI Appendix, Fig. S11B**. We do the same procedure on the spectra obtained on Gr alone. The two peaks at \( \sim 1,585 \) and \( \sim 2,620 \text{ cm}^{-1} \) are the 6 and 2D modes, respectively.

**T\(_e\) and e–h Pair Density Calculation.** To extract the \( T\(_e\) \) from Eq. 1, the value of \( \gamma (\nu_T^\text{e}) \) is required, which can be extracted from the FWHM of the 2D mode feature, \( \Gamma_{2D} \) (35), which writes

\[
\Gamma_{2D} = 4 \nu_T^\text{e} / 21 - 1 \left( \frac{\partial \nu_T^\text{e}}{\partial \nu_T^\text{e}} \right) \gamma (\nu_T^\text{e}),
\]

where \( \nu_T^\text{e} \) is the 2D mode frequency, \( \partial \nu_T^\text{e} / \partial (\nu_T^\text{e}) \left( \frac{1}{2 \nu_T^\text{e}} \right) \approx 100 \text{ cm}^{-1} \) (32, 69), i.e., the ratio between the phonon and Fermi velocity, defined as the slope of phononic electronic dispersion at the phonon (electron) momentum corresponding to a given excitation laser energy \( h\nu_{\text{max}} \) (32). \( \nu_T^\text{e} \) and \( \nu_T^\text{e} \) indicate the speed of light and the Planck constant, respectively. However, Eq. 6 does not consider the bimodal spectral feature of the 2D mode, due to the contributions of so-called inner and outer processes (70, 71). For this reason, we isolate each contribution, following the procedure reported in ref. 69. The fitted bandwidth of the probe pulse alone induces \( T\(_e\) \approx T\(_e\) \) in Gr, further reducing the upper bound for the Fermi level down to 30 meV (72). In **SI Appendix, Fig. S9**, we report \( n_{\text{e}} \) as a function of \( T\(_e\) \) for \( E_g = 0 \) meV and \( E_g = 30 \) meV. Such a doping effect mainly results in a minor vertical shift, not relevant for the incremental carrier generation \( N\(_e\) \) additional to the probe-induced \( n_{\text{e}} \) associated with a temperature rise from \( T\(_e\) \) to the pump-induced time-dependent \( T\(_e\) \).

**Data Availability.** All study data are included in the article and/or **SI Appendix**.

**ACKNOWLEDGMENTS.** TS acknowledges the support from the PRIN 2017 Project 2017955BA3 - HARVEST. T.S. and G.B. are grateful to the ‘Progetti di Ricerca Medi 2019’ grant by Sapienza Università di Roma. We acknowledge Francesco Mauri for fruitful discussions. We are grateful to the SiNano clean room staff and to M. Romeo for technical support. S.B. acknowledges support from Institut Universitaire de France (IUF) and benefited from a “Sapienza” University visiting professorship grant 2017. This project has received funding from Agence Nationale de la Recherche under grants 2D-POEM ANR 18-ERC-0009, and ATOEMS ANR-20-CE24-0010. This work of the Interdisciplinary Thematic Institute QMat, as part of the IIT 2021-2026 program of the University of Strasbourg, CNRS and Inserm, was supported by IdEx Unistra (ANR 10 IDEX 0002), and by SIRI STRATUS project (ANR 20 SFRI 0012) and EUR QMAT ANR 17-EURE-0024 under the framework of the French Investments for the Future Program. T.S. acknowledges the support from the European Union’s Horizon 2020 research and innovation programme Graphene Flagship under grant agreement No 881603 and the PRIN 2017 Project 2017955BA3 - HARVEST.

**Author affiliations:** 1. Graphene Labs, Istituto Italiano di Tecnologia, I-16163 Genova, Italy; 2. Innovative Nuclear Systems Laboratory, Fusion and Technology for Nuclear Safety and Security Department, “Italian National Agency for New Technologies, Energy and Sustainable Economic Development,” Casaccia, 00123 Roma, Italy; 3. Center for Life Nano Science @Sapienza, Istituto Italiano di Tecnologia, I-00161 Roma, Italy; 4. Departamento de Física, Universidade de Roma “La Sapienza,” 00185 Roma, Italy; and 5. “Institut de Physique et Chimie des Matériaux de Strasbourg, UMR 7504, Université de Strasbourg, CNRS, F-67000 Strasbourg, France

1. T. Georgiou et al., Vertical field-effect transistor based on graphene WS\(_2\) heterostructures for flexible and transparent electronic. Nat. Nanotechnol. 8, 100–103 (2013).
2. L. Britnell et al., Strong light-matter interactions in heterostructures of atomically thin films. Science 340, 1311–1314 (2013).
3. C. Palacios-Berraquero et al., Atomically thin quantum-light-emitting diodes. Nat. Commun. 7, 12978 (2016).
4. M. Massiotte et al., Picoscopic photoreponse in van der Waals heterostructures. Nat. Nanotechnol. 11, 42–46 (2016).
5. M. Massiotte et al., Photo-thermic effect in vertical graphene heterostructures. Nat. Commun. 7, 12174 (2016).
6. E. Lorchat et al., Filtering the photoluminescence spectra of atomically thin semiconductors with graphene. Nat. Nanotechnol. 15, 283–288 (2020).
7. G. Wang et al., Colloquium: Excitons in atomically thin transition metal dichalcogenides. Rev. Mod. Phys. 90, 021001 (2018).
8. C. Traverso et al., Optical parametric amplification by monolayer transition metal dichalcogenides. Nat. Photon. 15, 6–10 (2021).
9. E. A. Pogna et al., Photo-induced bandgap renormalization governs the ultrafast response of single-layer mos2. ACS Nano 10, 1182–1188 (2016).
10. L. M. Malaíl, T. A. Alcântara, A. P. M. Barbosa, K. F. Mak, A. M. de Paula, Observation of intense second harmonic generation from mos2 atomic crystals. Phys. Rev. B 87, 201401(R) (2013).
11. K. F. Mak, J. Shan, Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides. Nat. Photonics 10, 216–226 (2016).
12. K. Tubaro et al., Photocatalysis cascading and multiple hot carrier generation in graphene. Nat. Phys. 9, 248–252 (2013).
13. F. H. L. Koppens et al., Photodetectors based on graphene, other two-dimensional materials and hybrid systems. Nat. Nanotechnol. 9, 780–793 (2014).
14. G. Froehlicher, E. Lorchat, S. Becuau, Charge versus energy transfer in atomically thin transition metal dichalcogenide van der waals heterostructures. Phys. Rev. B 98, 011007 (2018).
15. L. Yuan et al., Photocarrier generation from interlayer charge-transfer transitions in WS\(_2\) graphene heterostructures. Sci. Adv. 4, e1700324 (2018).
16. J. He et al., Electron transfer and coupling in graphene-tungsten disulfide double van der Waals heterostructures. Nat. Commun. 5, 5625 (2014).
17. Y. Chen, Y. Li, Y. Zhao, H. Zhou, H. Zhu, Highly efficient hot electron harvesting from graphene before electron-hole thermalization. Sci. Adv. 5, eaax958 (2019).
18. A. Govorov, P. L. H. Martinez, H. V. Demir, Understanding and Modeling Förster-Type Resonance Energy Transfer (FRET) (Springer, 2016).
19. D. M. Adams et al., Charge transfer on the nanoscale: Current status. J. Phys. Chem. B 107, 6668–6697 (2003).
20. D. Koszwa et al., Evidence for fast interlayer energy transfer in MoS\(_2\)/WS\(_2\) heterostructures. Nano Lett. 16, 4087–4093 (2016).
21. S. Aeschimann et al., Direct evidence for efficient ultrastable charge separation in epitaxial WS\(_2\)/graphene heterostructures. Sci. Adv. 6, eaay7611 (2020).
