Unveiling valley lifetimes of free charge carriers in monolayer WSe$_2$

Manfred Ersfeld,[1] Frank Volmer,[1] Lars Rathmann,[1] Luca Kotewitz,[1] Maximilian Heithoff,[1] Mark Lohmann,[2] Bowen Yang,[3] Kenji Watanabe,[4] Takashi Taniguchi,[4] Ludwig Bartels,[3] Jing Shi,[2] Christoph Stampfer,[1,5] and Bernd Beschoten[1]

[1] *2nd Institute of Physics and JARA-FIT, RWTH Aachen University, 52074 Aachen, Germany*
[2] *Department of Physics and Astronomy, University of California, Riverside, California 92521, USA*
[3] *Department of Chemistry and Materials Science & Engineering Program,*
University of California, Riverside, California 92521, USA
[4] *National Institute for Materials Science, 1-1 Namiki, Tsukuba, 305-0044, Japan*
[5] *Peter Grünberg Institute (PGI-9), Forschungszentrum Jülich, 52425 Jülich, Germany*

We report on nanosecond long, gate-dependent valley lifetimes of free charge carriers in monolayer WSe$_2$, unambiguously identified by the combination of time-resolved Kerr rotation and electrical transport measurements. While the valley polarization increases when tuning the Fermi level into the conduction or valence band, there is a strong decrease of the respective valley lifetime consistent with both electron-phonon and spin-orbit scattering. The longest lifetimes are seen for spin-polarized bound excitons in the band gap region. We explain our findings via two distinct, Fermi level-dependent scattering channels of optically excited, valley polarized bright trions either via dark or bound states. By electrostatic gating we demonstrate that the transition metal dichalcogenide WSe$_2$ can be tuned to be either an ideal host for long-lived localized spin states or allow for nanosecond valley lifetimes of free charge carriers (> 10 ns).

With band gaps in the optically visible energy range[1–3] and spin-split valleys which allow the creation of valley- and spin-polarized excitons by circularly polarized light,[4,5] monolayer (ML) transition metal dichalcogenides (TMDs) are a very promising family of materials in the field of both spin- and valleytronics.[4,6–8] However, the full potential of TMD-based devices for valleytronic applications critically depends on their valley lifetime of free charge carriers. So far, there is a huge variation in experimentally reported valley lifetimes, which range from the picosecond up to the microsecond timescale.[9–18] Especially the longer lifetimes are not consistent with ab-initio studies that predict valley lifetimes in the picosecond range limited by electron-phonon coupling.[12,13] This discrepancy is partly due to the ambiguous use of the term valley polarization, which can be explained in respect to two measurement techniques typically used for exploring spin and valley dynamics in ML TMDs: Time-resolved photoluminescence (TRPL) and time-resolved Kerr rotation (TRKR). Both techniques rely on the valley-selective optical excitation of excitons by circularly polarized laser pulses (see Fig. 1a, circles represent the photo-excited electron-hole pair, shaded areas represent the filling of the bands with free charge carriers). The situation in Fig. 1a is sometimes called a valley polarization as, e.g., the numbers of holes in the valence band differ between the K and K' valleys. In this article we identify how such an exciton valley polarization can create a net valley polarization of free conduction and valence band states after the photo-excited charge carriers have recombined (see Fig. 1b) and determine the valley lifetimes of the respective free charge carrier valley polarization. Here, TRPL reaches its limitations as it is restricted to exciton lifetimes[5] and is therefore not capable to determine a valley polarization of free charge carriers after exciton recombination. However, the probe pulse in

TRKR can detect the temporal decay of the valley polarization from the Kerr rotation angle when tuned to the trion energy, as the creation of the trion depends on the availability of free charge carriers within each valley.

To identify the existence of a valley polarization of free charge carriers and to determine its valley lifetimes, we combine TRKR, electrical transport measurements, and photoluminescence (PL) spectroscopy on ML WSe$_2$ protected by hexagonal boron nitride and contacted via graphite electrodes (see Supporting Information (SI) for details on device fabrication). The combination of electrical transport and PL measurements (Fig. 1c) allows to assign the position of the Fermi level to the corresponding back-gate voltage (V_{BG}):[19] Between -80 V and -10 V the Fermi-level is pinned within the band gap of the TMD due to mid-gap states, resulting in a very low conductance with a current of some tens of pA due to residual hopping transport[20,21] and the dominance of bound exciton states (X$_{\mathrm{bound}}$) in the PL spectrum.[5,22,23] Between -10 V and 10 V the current undergoes an exponential increase as the Fermi level moves through the tail states of the conduction band.[24,25] Above $V_{\mathrm{BG}} = 20$ V the Fermi level is tuned into the conduction band, resulting in a linear gate voltage dependence of the current in accordance to diffusive charge transport of free electrons and the complete disappearance of the neutral exciton (X$_0$) emission in PL as enough free charge carriers are present for the formation of the energetically favourable charged excitons (trions, X$_{+/-}$).[5] At even higher electron densities we observe the appearance of the X$'_-$ feature, which can either be attributed to the interaction of an exciton with the Fermi sea of free electrons or to the onset of filling the energetically higher, spin-split conduction band.[26–28]

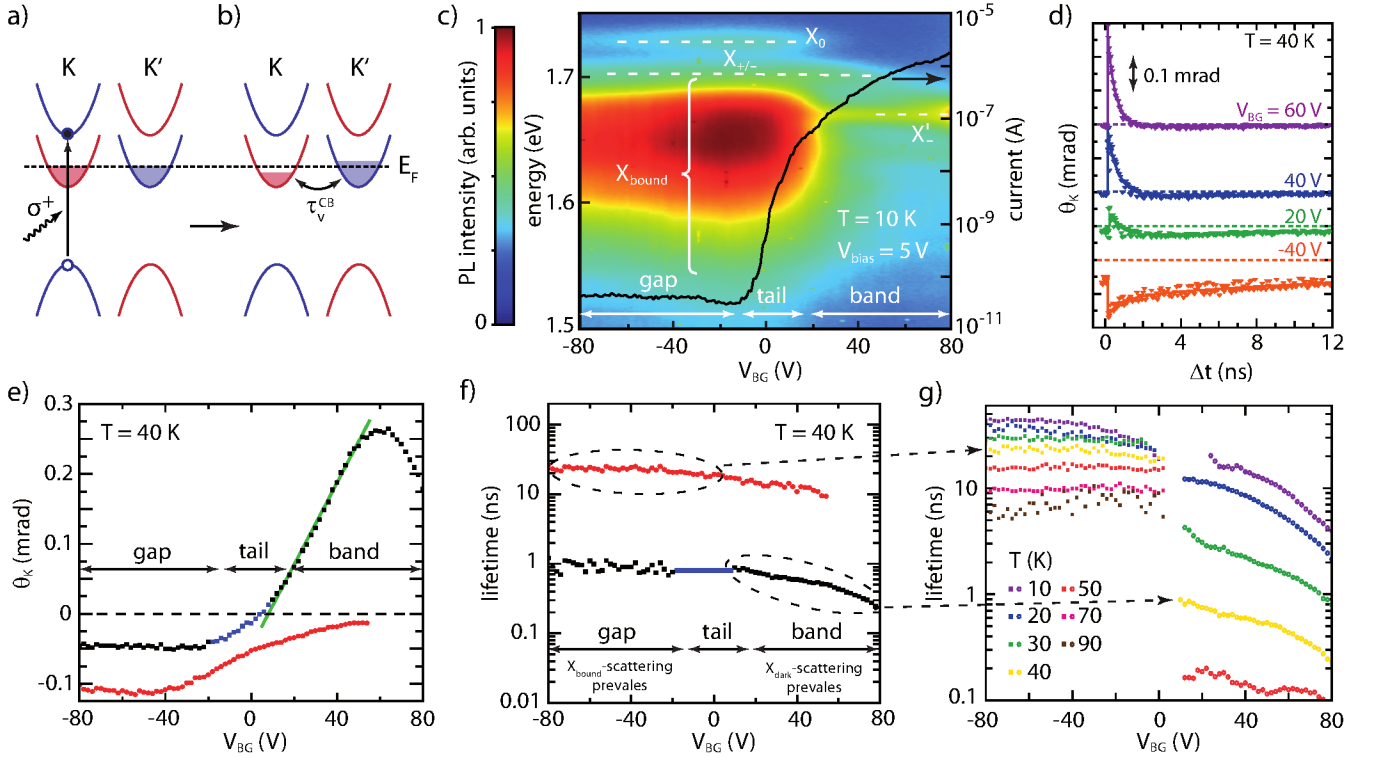Fig. 1d shows TRKR traces for representative gate voltages measured with pump and probe energies in the

Figure 1. Valley polarization of free charge carriers in monolayer WSe$_2$. (a) The circularly polarized pump pulse in time-resolved Kerr rotation (TRKR) measurements valley-selectively creates electron-hole pairs (circles) in an n-doped WSe$_2$ monolayer (shaded areas represent free charge carriers within the bands). (b) The important question in TMD-based valleytronics is how the initial exciton polarization in (a) can create a valley polarization of free charge carriers with a corresponding valley lifetime τ_v^{CB}. (c) Gate-dependent photoluminescence and electrical transport measurements are used to determine the position of the Fermi level within the band structure of the TMD as a function of gate-voltage. (d) TRKR curves for respective gate-voltages with bi-exponential fit functions (solid lines). Dashed lines represent the corresponding $\theta_K = 0$ axes. (e) Gate-dependent Kerr rotation amplitudes θ_K and (f) lifetimes. The black data points for positive gate voltages represent the valley polarization of free charge carriers, identified by 1.) the linear increase of its amplitude with increasing charge carrier density, i.e. increasing gate voltage, once the Fermi level crosses into the conduction band (green line in (e) is a guide to the eye), 2.) its maximum amplitude appearing at the same gate voltages as the X$'_-$ emission in the PL spectra in (c), which indicates the onset of the filling of the upper conduction band, and 3.) the decrease of the respective valley lifetimes in (f) for higher charge carrier densities. The red data points in (e,f) correspond to a polarization from polarized excitons bound to band gap states. (g) Lifetimes of polarized bound excitons ($V_{BG} < 0$ V) and conduction band valley polarization ($V_{BG} > 0$ V) at different temperatures.

trion regime at a temperature of $T = 40$ K, showing two exponentially decaying signals. Within the band gap ($V_{BG} = -40$ V, orange data points, dashed line represents the axis with the Kerr rotation angle θ_K being zero) a long-lived polarization with a lifetime exceeding the laser repetition interval of 12.5 ns can be observed together with a short-lived polarization of around 1 ns. In the transition regime between tail states and conduction band ($V_{BG} = 20$ V, green curve) the long-lived polarization starts to vanish, whereas the short-lived polarization first undergoes a sign reversal and then shows an increasing amplitude towards higher charge carrier densities at larger V_{BG} values (compare green, blue, and violet curves within the first two ns). We fitted the data over the whole gate-voltage range by the sum of two exponential decays (solid lines) and plotted the extracted amplitudes and lifetimes in Figs. 1e and 1f, respectively.

The Kerr rotation amplitude θ_K of the long-lived polarization is almost constant over the whole gap regime and starts to disappear as soon as the Fermi level reaches the tail regime (red data points in Fig. 1e), demonstrating its connection to band gap states. On the other hand, the amplitude of the short-lived polarization (black data points) shows a linear increase with increasing charge carrier density between $V_{BG} = 10$ V and 50 V (see green line in Fig. 1e as a guide to the eye). As this increase goes hand in hand with the linear increase in current in this gate voltage range, it is clearly connected to the increasing number of free charge carriers in the conduction band and, hence, can be attributed to a valley polarization of these free charge carriers. This assignment is backed up by the fact that the amplitude reaches a maximum at around the same gate voltages where the X$'_-$ feature appears in the PL map of Fig. 1c. The decrease of θ_K is

expected at even larger V_{BG} values as the filling of both the upper, spin-inverted conduction band and the bands at the Q-valleys (Λ-valleys) will reduce the overall net valley polarization of free charge carriers.[5] Finally, the strong decrease in lifetime towards higher gate-voltages (see black data points in Fig. 1f) also supports our assignment of the measured valley polarization to the free band carriers, as such a decay is expected from both wave vector dependent electron-phonon and spin-orbit scattering mechanisms.[12,29–33] Especially the electron-phonon scattering is expected to change as soon as the Q-valleys (Λ-valleys) come into play at higher electron densities.[13,34]

The free carrier valley polarization and the band gap polarization show distinctively different temperature dependencies of their respective lifetimes as depicted in Fig. 1g, where for $V_{BG} < 0$ V only the band gap polarization and for $V_{BG} > 0$ V only the valley polarization is plotted for simplicity (see dashed lines in Fig. 1f). The valley lifetime of the conduction band polarization strongly decreases from 20 ns at 10 K to below 200 ps at 70 K. Such a strong temperature dependence is expected if electron-phonon scattering limits the valley lifetimes.[12,13,33] In contrast, the long-lived polarization in the band gap is far more robust against temperature and retains lifetimes of 10 ns at 70 K and can be well observed up to temperatures of around 100 K. This is the temperature at which the bound exciton emission in PL typically disappears,[18,38,39] highlighting a possible connection between this long-lived polarization and bound excitons. This notion is backed up by comparing other publications showing either long-lived TRKR signals or bound exciton features in PL, respectively, showing that both signals often have similar temperature dependencies.[10,14,15,18,38–40]

The mechanisms which explain both the formation of the free carrier valley polarization and the long-lived polarization in the gap region are depicted in Fig. 2. We start with the Fermi level tuned into the conduction band ($V_{BG} > 10$ V) at equilibrium conditions, i.e. an equal number of free charge carriers in both valleys (indicated by the shaded areas in the band structure). A circularly polarized laser pump pulse now excites valley-selectively electron-hole pairs (circles in Fig. 2a).[5] By the interaction with the free conduction band carriers, the electron-hole pairs first create bright trions which, however, cannot directly be responsible for the TRKR signals in the ns range as their recombination times are typically in the ps range.[5,41,42]

If the photo-excited hole now recombines with the photo-excited electron, the system will end up in its initial state, which means that no valley polarization of free charge carriers can be created via this recombination channel. Instead, a net valley polarization can be created if, in a first step, one charge of the electron-hole pair scatters into the other valley. For WSe$_2$ this can happen by an intervalley transition of the photo-excited electron into the energetically lower conduction band of the same spin-orientation in the K'-valley (Fig. 2b), which

creates a dark exciton.[5,43] For the recombination of the dark exciton, the photo-excited hole has to recombine either with an electron from the lower conduction band of the K- or the K valley. The latter process is momentum-forbidden and, hence, is possible only via an interaction with a phonon. Although this process is possible, it will not lead to a valley polarization of free charge carriers, as the initial increase in electron number via the scattering of the photo-excited electron is undone by catching an electron from the same valley for the recombination process. Yet, a valley polarization of free charge carriers can be created, if the photo-excited hole recombines with an electron of the K-valley on the time-scale τ_{rec}^{dark} (see Fig. 2c). Although this process is spin-forbidden in case of dipole transitions, in-plane symmetry nevertheless dictates that it is possible via the coupling to photons with in-plane propagation and out-of-plane linear polarization,[36] which is demonstrated in several PL studies.[35,36,44–49] This kind of spin-flip recombination is reducing the number of carriers in the K-valley (compare Figs. 2c and 2d), whereas the electron number in the K'-valley is increased by the leftover photo-exited electron, creating a net valley polarization of conduction band electrons (see Fig. 2d).

The dark exciton can always undergo a momentum-forbidden recombination, because the number of available electrons in the K-valley, as given by the initially-present free charge carriers plus the scattered photo-excited electrons, is always larger than the number of photo-excited holes. This is not the case for the spin-forbidden recombination channel: especially for charge carrier densities lower than the density of excited dark excitons, the spin-forbidden recombination channel can completely deplete the K-valley from electrons. In such a situation, leftover photo-excited holes are limited to the momentum-forbidden recombination channel (the one that does not result in a valley polarization). Hence, the maximum number of spin-flip recombination events and, therefore, the maximum valley polarization is directly determined by the initial number of free charge carriers in the K-valley. Overall, we expect that an increasing number of free charge carriers in the K-valley will increase the chance that a photo-excited hole recombines via a spin-flip process. Therefore, we can assign the TRKR signal, which shows an increasing amplitude with increasing charge carrier density (black data points between $V_{BG} = 10$ V and 50 V in Fig. 1e), to a valley polarization of free charge carriers created by the scattering channel of bright trions via dark states as depicted in Figs. 2b-2d. Accordingly, the black data points in Fig. 1f for $V_{BG} > 10$ V represent the Fermi level dependent valley lifetimes τ_v^{CB}.

Next to the scattering via dark states, bright trions can also bind to localized states within the band gap caused e.g. by vacancies or dopant atoms. We identify these bound states as the origin of the long-lived polarization seen as the red data points in Fig. 1e and 1f. We note that this polarization is also measured with a laser probe
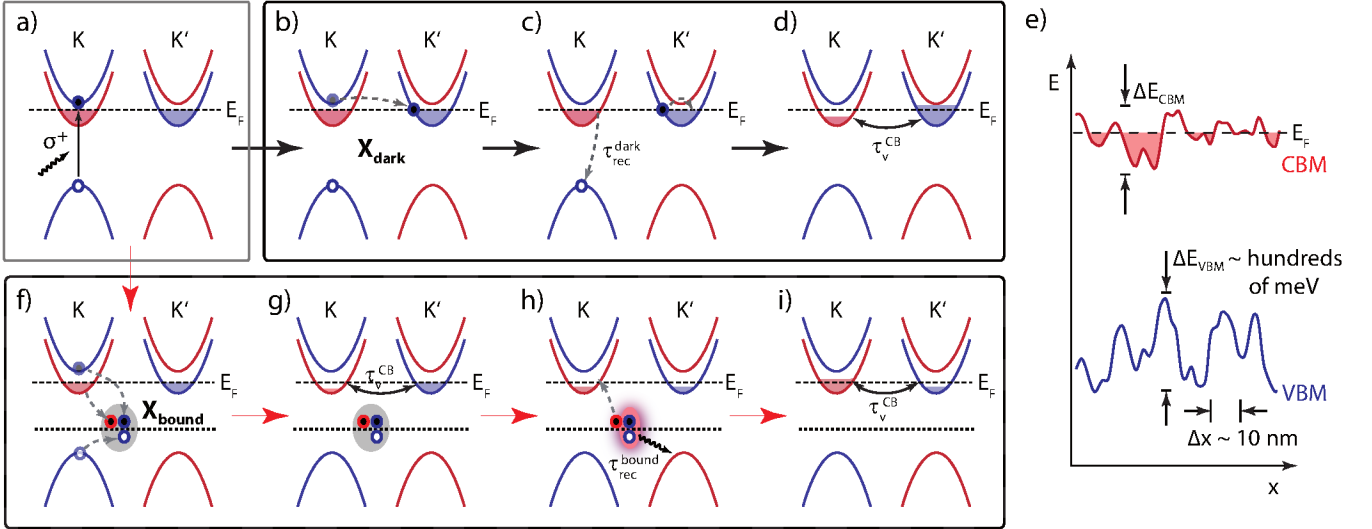
Figure 2. Formation of a valley polarization of free conduction band carriers by scattering of bright trions via dark or bound states. (a) Initial state of an exciton polarization created by a circularly polarized laser pulse in an n-doped WSe$_2$ monolayer. The photo-excited electron-hole pair (circles) interacts with free charge carriers in the bands (represented by shaded areas) to form a trion (interaction not shown for the sake of simplicity). (b) to (d) Scattering mechanism via dark excitons: The photo-excited electron can scatter into the K'-valley, creating a dark state (b). The dark exciton can relax via a spin-flip transition (c),[35,36] resulting in the recombination of the photo-excited hole with an initial free charge carrier, whereas the photo-excited electron now fills up the K'-valley. This results in a net valley-polarization of free charge carriers with a valley lifetime $\tau_\mathrm{v}^\mathrm{CB}$ (d). (e) Due to strain variations, monolayer TMDs can show conduction band minima (CBM) and valence band maxima (VBM) energies which can spatially vary by hundreds of meV over length-scales as small as 10 nm,[37] resulting in the appearance of charge puddles depicted as shaded areas. (f) to (i) Scattering mechanism via bound states: The photo-excited electron-hole pair can create a long-lived bound trion at a band gap state with an additional charge carrier from the charge puddles (f), creating at the same time a valley polarization (g). During the bound trion's recombination (h) the initially caught charge carrier can transfer back into its original valley and, hence, will again create a valley polarization (i).

pulse at the bright trion energy. When probing at the energetically lower bound exciton energies no Kerr rotation signal is observed. Therefore, we do not directly probe the polarization of the bound exciton states but rather have to consider a charge transfer process between bound excitons and the optically accessible conduction band states. This leads to a complication: On the one hand, the Fermi level has to be in the conduction band to have free carriers available both for the formation of the initially excited bright trions by the pump pulse and also for the creation of the final valley polarization. But at the same time, having the Fermi level in the conduction band also means that all band gap states are occupied by unpolarized charge carriers, preventing an interaction with bright trions.

In this context, it is important to consider the existence of charge puddles in the transition regions between gap and bands. Scanning tunneling spectroscopy studies revealed that monolayer TMDs can show conduction band minima (CBM) and valence band maxima (VBM) energies which can spatially vary by hundreds of meV over length-scales as small as 10 nm due to strain variations (see schematic in Fig. 2e).[37] In this situation, charge puddles are formed as depicted by the shaded areas in Fig. 2e. While bright trions are easily formed with the

charges in these puddles, their formation is diminished in nearby regions where the Fermi level is in the band gap. This picture accounts for the coexistence of both the exciton (X$_0$) and the trion feature (X$_{+/-}$) over a part of the gate voltage range in PL measurements as seen both in Fig. 1c and previous works.[5] Our model for the long-lived polarization is now based on the assumption that a photo-excited electron-hole pair can create a bound trion at a band gap state with an additional charge carrier taken from the conduction band (i.e. charges from the puddles) (Fig. 2f). For the sake of simplicity, we discuss the case that the additional electron originates from the K-valley, creating a bound exciton in a singlet state. This state is energetically more favorable than the triplet state, which would be created by an additional electron from the K'-valley (see SI for a more detailed discussion). The resulting imbalance of free charge carriers between K and K' valley will relax to equilibrium conditions within the time-scale of the valley lifetime $\tau_\mathrm{v}^\mathrm{CB}$ (Fig. 2g). This explains why we also see the valley polarization within the gap region (compare Figs. 2d with 2g) and that the lifetime of the valley polarization (black data points in Fig. 1f) shows a smooth transition over the whole gate voltage range.

The bound excitons can have recombination times

$\tau_{\text{rec}}^{\text{bound}}$ up to the μs range[22,50] and show a certain degree of polarization in PL,[22] indicating that a spin polarization of these bound excitons is not completely relaxed at the time of their recombination. Therefore, during the bound exciton's recombination (Fig. 2h), we argue that the initially caught charge carrier will predominately transfer back into the valley where it came from (the one with the same spin orientation) and, hence, will again create a valley polarization of free charge carriers (Fig. 2i). The measured long-lived lifetime in the TRKR data (red data points in Figs. 1e and 1f) is therefore a valley polarization that is constantly decaying with a valley lifetime in the lower ns range (the one which was unambiguously determined in the band regime via the process discussed in Figs. 2b-d), while being simultaneously replenished via the bound exciton recombination process over a timescale of tens of ns. We developed a model based on rate equations to simulate this competing decay and replenishment, and to simulate the resulting valley polarization over time (see Supporting Information). This model demonstrates that the resulting lifetime in a TRKR measurement (red data points in Fig. 1f) is mainly determined by the bound exciton recombination time and not by the much shorter valley lifetimes ($\tau_{\text{v}}^{\text{CB}} < \tau_{\text{rec}}^{\text{bound}}$). Although the recombination process via bound excitons leads to a valley polarization of free charge carriers, the respective lifetime in TRKR thus cannot be assigned to the genuine valley lifetime. In this respect, we note that the longest reported lifetimes of bound excitons[22,50] are interestingly close to the longest reported lifetimes in TRKR measurements.[10,11,13] Hence, our study highlights the importance for a thorough gate-dependent analysis of the TRKR signals combined with electrical transport measurements to untangle the different bright exciton scattering channels (Figs. 2b-2d vs. Figs. 2f-2i). This is the sole path to an unambiguous determination of the lifetime in the overall TRKR signal, which can be assign to the genuine valley polarization of free charge carriers.

The decrease in amplitude of the signal which originates from bound excitons (red data points in Fig. 1e) towards the conduction band can be well understood, as an increasing Fermi level will decrease the number of unoccupied localized states in the band gap at which a bright trion can be bound.[22] An important aspect of our model is that it predicts a sign reversal of the net valley polarization for the two scattering mechanisms via dark and bound excitons (compare Figs. 2d and 2i where the net valley polarization is carried by K' and by K states, respectively), which is indeed seen in the measurements (see black data points for positive and red data points for negative gate voltages in Fig. 1e).

To confirm that our model indeed describes the underlying valley- and spin-dynamics consistently, we study a device made from ML MoSe$_2$. In contrast to WSe$_2$, in MoSe$_2$ the formation of dark excitons is energetically not favorable due to its inverted conduction band spin ordering (see inset in Fig. 3a).[5,43] Hence, the scattering
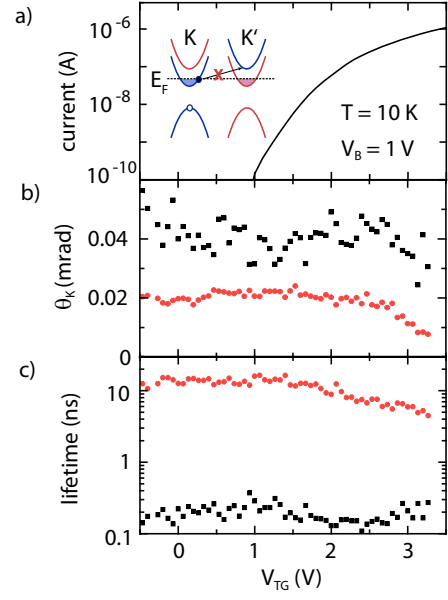

Figure 3. Valley polarization in monolayer MoSe$_2$. (a) Gate-dependent current showing a tuning of the Fermi level into the conduction band for top-gate voltages above $V_{\text{TG}} = 1$ V. Inset: Due to the inverted conduction band spin ordering, the scattering mechanism via dark states (Figs. 2b to 2d) is energetically not favourable in MoSe$_2$ and, therefore, there is no linear increase in the gate-dependent Kerr rotation amplitude (b) when tuning the Fermi level into the conduction band. (b) Kerr rotation amplitudes and (c) polarization lifetimes extracted from bi-exponential fits to TRKR traces.

mechanism via dark excitons (Figs. 2b to 2d), which is responsible for the linear increase of the valley polarization with increasing charge carrier density, is not expected. In contrast, the scattering mechanism via bound excitons is still feasible in MoSe$_2$ (Figs. 2f to 2i), but in that case the amplitude of the valley polarization is not limited by the amount of free charge carriers but rather by the number of bound excitons which can be created (see transition from Fig. 2f to 2g). And in fact, the TRKR data from the MoSe$_2$ sample shows both a long-lived and a short-lived polarization (see Figs. 3b and 3c) consistent to the scattering mechanism via bound states, but no increase of the Kerr rotation amplitude towards higher charge carrier densities, although the increase of the measured current by more than four orders of magnitude (Fig. 3a) clearly demonstrates that the Fermi level can be tuned into the conduction band of MoSe$_2$.

Finally, to confirm that our model also holds for the valence band we studied another WSe$_2$ sample in which we were able to tune the Fermi level all the way from the valence band into the conduction band. Figs. 4a and 4b show the Kerr rotation amplitude and the lifetime of the respective valley polarizations of both bands which got polarized by the scattering channel via dark trion states (Figs. 2b to 2d). Next to the conduction band, also the valence band shows the expected increase in the valley
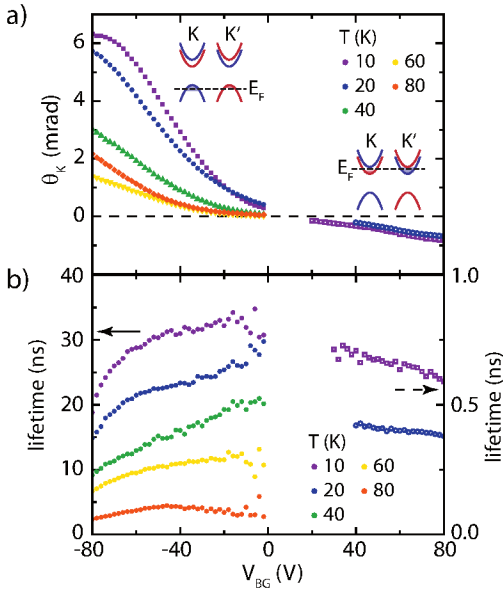
Figure 4. CVD-grown WSe$_2$ sample with valley polarization in conduction and valence band. (a) Gate-dependent Kerr rotation amplitudes for different temperatures and (b) corresponding lifetimes of the valley polarization of free charge carriers in the valence ($V_{\mathrm{BG}} < 0$ V) and the conduction band ($V_{\mathrm{BG}} > 20$ V). Plotted is the one polarization from a bi-exponential fit which shows the characteristic increase in amplitude with simultaneous decrease in lifetime towards higher charge carrier densities.

polarization and decrease in valley lifetimes with higher charge carrier densities. Interestingly, the valley lifetimes of the valence band states are a factor of 30 longer than the corresponding valley lifetimes of the conduction band states (see different y-axes).

Remarkably, the valley lifetimes in our samples are found to vary between 100 ps and tens of ns at 10 K, which is much longer than expected from ab-initio studies of pristine TMD monolayers that predict electron-phonon-limited valley lifetimes in the lower ps-range.[12,13] Accordingly, to achieve such long-lived valley lifetimes the intervalley electron-phonon scattering has to be suppressed, which is most likely accomplished by fabrication-induced local strain variations in our samples. In this

respect, it was shown that strain indeed has a significant impact on the band structure of TMDs,[37,51] being able to suppress the electron-phonon coupling[32,52] and, hence, can e.g. lead to an increase in charge carrier mobilities.[32,53]

Together with our model in Fig. 2 this possible strain-induced decrease in electron-phonon coupling can explain how an exciton valley polarization can create a nanosecond long valley polarization of free charge carriers. This valley polarization of free charge carriers can be clearly identified in our experiments, as 1.) its amplitude increases linearly with the gate-induced charge carrier density which goes hand in hand with a simultaneous increase in electrical conductance, 2.) its amplitude decreases as soon as the upper, spin-inverted conduction band starts to be filled, and 3.) its lifetime decreases towards higher charge carrier densities in accordance to theory about wave vector dependent electron-phonon and spin-orbit scattering mechanisms.[12,29–33] Overall, we have shown that by changing the Fermi level via gating, TMDs can be tuned to be either ideal hosts for long-lived localized spin states or allow valley lifetimes of conduction and valence band states exceeding 10 ns at 10 K, which are adequate timescales for both spin manipulation and valley transport.

Author contributions: M. E., F. V., and L. R. are equally contributing authors.

[1] Mann, J. et al. 2-Dimensional Transition Metal Dichalcogenides with Tunable Direct Band Gaps: $MoS_{2(1-x)}Se_{2x}$ Monolayers. *Adv. Mater.* **2014**, *26*, 1399–1404.

[2] Manzeli, S.; Ovchinnikov, D.; Pasquier, D.; Yazyev, O. V.; Kis, A. 2D transition metal dichalcogenides. *Nat. Rev. Mater.* **2017**, *2*, 17033.

[3] Mak, K. F.; Shan, J. Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides. *Nature Photonic* **2016**, *10*, 216–226.

[4] Xu, X.; Yao, W.; Xiao, D.; Heinz, T. F. Spin and pseudospins in layered transition metal dichalcogenides. *Nat. Phys.* **2014**, *10*, 343–350.

[5] Wang, G.; Chernikov, A.; Glazov, M. M.; Heinz, T. F.; Marie, X.; Amand, T.; Urbaszek, B. Colloquium: Excitons in atomically thin transition metal dichalcogenides. *Rev. Mod. Phys.* **2018**, *90*, 021001.

[6] Schaibley, J. R.; Yu, H.; Clark, G.; Rivera, P.; Ross, J. S.; Seyler, K. L.; Yao, W.; Xu, X. Valleytronics in 2D materials. *Nat. Rev. Mater.* **2016**, *1*, 16055.

[7] Jin, C.; Ma, E. Y.; Karni, O.; Regan, E. C.; Wang, F.;

Heinz, T. F. Ultrafast dynamics in van der Waals heterostructures. *Nat. Nanotechnol.* **2018**, *13*, 994–1003.

[8] Jin, C.; Kim, J.; Utama, M. I. B.; Regan, E. C.; Kleemann, H.; Cai, H.; Shen, Y.; Shinner, M. J.; Sengupta, A.; Watanabe, K.; Taniguchi, T.; Tongay, S.; Zettl, A.; Wang, F. Imaging of pure spin-valley diffusion current in WS$_2$-WSe$_2$ heterostructures. *Science* **2018**, *360*, 893–896.

[9] Yang, L.; Sinitsyn, N. A.; Chen, W.; Yuan, J.; Zhang, J.; Lou, J.; Crooker, S. A. Long-lived nanosecond spin relaxation and spin coherence of electrons in monolayer MoS$_2$ and WS$_2$. *Nat. Phys.* **2015**, *11*, 830–834.

[10] Dey, P.; Yang, L.; Robert, C.; Wang, G.; Urbaszek, B.; Marie, X.; Crooker, S. A. Gate-Controlled Spin-Valley Locking of Resident Carriers in WSe$_2$ Monolayers. *Phys. Rev. Lett.* **2017**, *119*, 137401.

[11] Goryca, M.; Wilson, N. P.; Dey, P.; Xu, X.; Crooker, S. A. Detection of thermodynamic "valley noise" in monolayer semiconductors: Access to intrinsic valley relaxation time scales. *Sci. Adv.* **2019**, *5*, eaau4899.

[12] Molina-Sánchez, A.; Sangalli, D.; Wirtz, L.; Marini, A. Ab Initio Calculations of Ultrashort Carrier Dynamics in Two-Dimensional Materials: Valley Depolarization in Single-Layer WSe$_2$. *Nano Lett.* **2017**, *17*, 4549.

[13] Ersfeld, M.; Volmer, F.; de Melo, P. M. M. C.; de Winter, R.; Heithoff, M.; Zanolli, Z.; Stampfer, C.; Verstraete, M. J.; Beschoten, B. Spin States Protected from Intrinsic Electron-Phonon Coupling Reaching 100 ns Lifetime at Room Temperature in MoSe$_2$. *Nano Lett.* **2019**, *19*, 4083–4090.

[14] Hsu, W.-T.; Chen, Y.-L.; Chen, C.-H.; Liu, P.-S.; Hou, T.-H.; Li, L.-J.; Chang, W.-H. Optically initialized robust valley-polarized holes in monolayer WSe$_2$. *Nat. Commun.* **2015**, *6*, 8963.

[15] Song, X.; Xie, S.; Kang, K.; Park, J.; Sih, V. Long-Lived Hole Spin/Valley Polarization Probed by Kerr Rotation in Monolayer WSe$_2$. *Nano Lett.* **2016**, *16*, 5010–5014.

[16] Schwemmer, M.; Nagler, P.; Hanninger, A.; Schüller, C.; Korn, T. Long-lived spin polarization in n-doped MoSe$_2$ monolayers. *Appl. Phys. Lett.* **2017**, *111*, 082404.

[17] McCormick, E. J.; Newburger, M. J.; Luo, Y. K.; McCreary, K. M.; Singh, S.; Martin, I. B.; Cichewicz, E. J.; Jonker, B. T.; Kawakami, R. K. Imaging spin dynamics in monolayer WS$_2$ by time-resolved Kerr rotation microscopy. *2D Mater.* **2017**, *5*, 011010.

[18] Zhu, C. R.; Zhang, K.; Glazov, M.; Urbaszek, B.; Amand, T.; Ji, Z. W.; Liu, B. L.; Marie, X. Exciton valley dynamics probed by Kerr rotation in WSe$_2$ monolayers. *Phys. Rev. B* **2014**, *90*, 161302.

[19] Wang, Z.; Shan, J.; Mak, K. F. Valley- and spin-polarized Landau levels in monolayer WSe$_2$. *Nat. Nanotechnol.* **2017**, *12*, 144–149.

[20] Ghatak, S.; Pal, A. N.; Ghosh, A. Nature of Electronic States in Atomically Thin MoS$_2$ Field-Effect Transistors. *ACS Nano* **2011**, *5*, 7707–7712.

[21] Qiu, H.; Xu, T.; Wang, Z.; Ren, W.; Nan, H.; Ni, Z.; Chen, Q.; Yuan, S.; Miao, F.; Song, F.; Long, G.; Shi, Y.; Sun, L.; Wang, J.; Wang, X. Hopping transport through defect-induced localized states in molybdenum disulphide. *Nat. Commun.* **2013**, *4*, 1–6.

[22] Moody, G.; Tran, K.; Lu, X.; Autry, T.; Fraser, J. M.; Mirin, R. P.; Yang, L.; Li, X.; Silverman, K. L. Microsecond Valley Lifetime of Defect-Bound Excitons in Monolayer WSe$_2$. *Phys. Rev. Lett.* **2018**, *121*, 057403.

[23] Tongay, S.; Suh, J.; Ataca, C.; Fan, W.; Luce, A.; Kang, J. S.; Liu, J.; Ko, C.; Raghunathanan, R.; Zhou, J.; Ogletree, F.; Li, J.; Grossman, J. C.; Wu, J. Defects activated photoluminescence in two-dimensional semiconductors: interplay between bound, charged, and free excitons. *Sci. Rep.* **2013**, *3*, 2657.

[24] Zhu, W.; Low, T.; Lee, Y.-H.; Wang, H.; Farmer, D. B.; Kong, J.; Xia, F.; Avouris, P. Electronic transport and device prospects of monolayer molybdenum disulphide grown by chemical vapour deposition. *Nat. Commun.* **2014**, *5*, 1–8.

[25] Wang, L.; Li, Y.; Gong, X.; Thean, A. V.; Liang, G. A Physics-Based Compact Model for Transition-Metal Dichalcogenides Transistors With the Band-Tail Effect. *IEEE Electron Device Letters* **2018**, *39*, 761–764.

[26] Wang, Z.; Mak, K. F.; Shan, J. Strongly Interaction-Enhanced Valley Magnetic Response in Monolayer WSe$_2$. *Phys. Rev. Lett.* **2018**, *120*, 066402.

[27] Efimkin, D. K.; MacDonald, A. H. Many-body theory of trion absorption features in two-dimensional semiconductors. *Phys. Rev. B* **2017**, *95*, 035417.

[28] Sidler, M.; Back, P.; Cotlet, O.; Srivastava, A.; Fink, T.; Kroner, M.; Demler, E.; Imamoglu, A. Fermi polaron-polaritons in charge-tunable atomically thin semiconductors. *Nature Physics* **2016**, *13*, 255.

[29] Ochoa, H.; Roldán, R. Spin-orbit-mediated spin relaxation in monolayer MoS$_2$. *Phys. Rev. B* **2013**, *87*, 245421.

[30] Ochoa, H.; Finocchiaro, F.; Guinea, F.; Fal'ko, V. I. Spin-valley relaxation and quantum transport regimes in two-dimensional transition-metal dichalcogenides. *Phys. Rev. B* **2014**, *90*, 235429.

[31] Habe, T.; Koshino, M. Spin relaxation in a hole-doped transition-metal dichalcogenide monolayer and bilayer with crystal defects. *Phys. Rev. B* **2016**, *93*, 075415.

[32] Ge, Y.; Wan, W.; Feng, W.; Xiao, D.; Yao, Y. Effect of doping and strain modulations on electron transport in monolayer MoS$_2$. *Phys. Rev. B* **2014**, *90*, 035414.

[33] Gunst, T.; Markussen, T.; Stokbro, K.; Brandbyge, M. First-principles method for electron-phonon coupling and electron mobility: Applications to two-dimensional materials. *Phys. Rev. B* **2016**, *93*, 035414.

[34] Sohier, T.; Ponomarev, E.; Gibertini, M.; Berger, H.; Marzari, N.; Ubrig, N.; Morpurgo, A. F. Enhanced Electron-Phonon Interaction in Multivalley Materials. *Phys. Rev. X* **2019**, *9*, 031019.

[35] Slobodeniuk, A. O.; Basko, D. M. Spin-flip processes and radiative decay of dark intravalley excitons in transition metal dichalcogenide monolayers. *2D Mater.* **2016**, *3*, 035009.

[36] Wang, G.; Robert, C.; Glazov, M. M.; Cadiz, F.; Courtade, E.; Amand, T.; Lagarde, D.; Taniguchi, T.; Watanabe, K.; Urbaszek, B.; Marie, X. In-Plane Propagation of Light in Transition Metal Dichalcogenide Monolayers: Optical Selection Rules. *Phys. Rev. Lett.* **2017**, *119*, 047401.

[37] Shin, B. G.; Han, G. H.; Yun, S. J.; Oh, H. M.; Bae, J. J.; Song, Y. J.; Park, C.-Y.; Lee, Y. H. Indirect Bandgap Puddles in Monolayer MoS$_2$ by Substrate-Induced Local Strain. *Advanced Materials* **2016**, *28*, 9378–9384.

[38] Yan, T.; Qiao, X.; Liu, X.; Tan, P.; Zhang, X. Photoluminescence properties and exciton dynamics in monolayer WSe$_2$. *Appl. Phys. Lett.* **2014**, *105*, 101901.

[39] Huang, J.; Hoang, T. B.; Mikkelsen, M. H. Probing the origin of excitonic states in monolayer WSe$_2$. *Sci. Rep.* **2016**, *6*, 22414.

[40] Volmer, F.; Pissinger, S.; Ersfeld, M.; Kuhlen, S.;

Stampfer, C.; Beschoten, B. Intervalley dark trion states with spin lifetimes of 150 ns in WSe$_2$. *Phys. Rev. B* **2017**, *95*, 235408.

[41] Robert, C.; Lagarde, D.; Cadiz, F.; Wang, G.; Lassagne, B.; Amand, T.; Balocchi, A.; Renucci, P.; Tongay, S.; Urbaszek, B.; Marie, X. Exciton radiative lifetime in transition metal dichalcogenide monolayers. *Phys. Rev. B* **2016**, *93*, 205423.

[42] Fang, H. H.; Han, B.; Robert, C.; Semina, M. A.; Lagarde, D.; Courtade, E.; Taniguchi, T.; Watanabe, K.; Amand, T.; Urbaszek, B.; Glazov, M. M.; Marie, X. Control of the Exciton Radiative Lifetime in van der Waals Heterostructures. *Phys. Rev. Lett.* **2019**, *123*, 067401.

[43] Malic, E.; Selig, M.; Feierabend, M.; Brem, S.; Christiansen, D.; Wendler, F.; Knorr, A.; Berghäuser, G. Dark excitons in transition metal dichalcogenides. *Phys. Rev. Materials* **2018**, *2*, 014002.

[44] Tang, Y.; Mak, K. F.; Shan, J. Long valley lifetime of dark excitons in single-layer WSe$_2$. *Nat. Commun.* **2019**, *10*, 4047.

[45] Zhou, Y.; Scuri, G.; Wild, D. S.; High, A. A.; Dibos, A.; Jauregui, L. A.; Shu, C.; De Greve, K.; Pistunova, K.; Joe, A. Y.; Taniguchi, T.; Watanabe, K.; Kim, P.; Lukin, M. D.; Park, H. Probing dark excitons in atomically thin semiconductors via near-field coupling to surface plasmon polaritons. *Nat. Nanotechnol.* **2017**, *12*, 856.

[46] Liu, E.; van Baren, J.; Lu, Z.; Altaiary, M. M.; Taniguchi, T.; Watanabe, K.; Smirnov, D.; Lui, C. H. Gate Tunable Dark Trions in Monolayer WSe$_2$. *Phys. Rev. Lett.* **2019**, *123*, 027401.

[47] Robert, C.; Amand, T.; Cadiz, F.; Lagarde, D.; Courtade, E.; Manca, M.; Taniguchi, T.; Watanabe, K.; Urbaszek, B.; Marie, X. Fine structure and lifetime of dark excitons in transition metal dichalcogenide monolayers. *Phys. Rev. B* **2017**, *96*, 155423.

[48] Borghardt, S.; Tu, J.-S.; Taniguchi, T.; Watanabe, K.; Kardynal, B. E. The Effect of Exciton Localization on Exciton Dynamics in *p*-Doped WSe$_2$ Monolayers. *arXiv* **2019**, 1911.01092.

[49] Borghardt, S.; Sonntag, J.; Tu, J.-S.; Taniguchi, T.; Watanabe, K.; Beschoten, B.; Stampfer, C.; Kardynal, B. E. Radially polarized light beams from spin-forbidden dark excitons and trions in monolayer WSe$_2$. *arXiv* **2020**, 2001.08043.

[50] Goodman, A. J.; Willard, A. P.; Tisdale, W. A. Exciton trapping is responsible for the long apparent lifetime in acid-treated MoS$_2$. *Phys. Rev. B* **2017**, *96*, 121404.

[51] Conley, H. J.; Wang, B.; Ziegler, J. I.; Haglund, R. F.; Pantelides, S. T.; Bolotin, K. I. Bandgap Engineering of Strained Monolayer and Bilayer MoS$_2$. *Nano Lett.* **2013**, *13*, 3626–3630.

[52] Niehues, I. et al. Strain Control of Exciton–Phonon Coupling in Atomically Thin Semiconductors. *Nano Lett.* **2018**, *18*, 1751–1757.

[53] Liu, T.; Liu, S.; Tu, K.-H.; Schmidt, H.; Chu, L.; Xiang, D.; Martin, J.; Eda, G.; Ross, C. A.; Garaj, S. Crested two-dimensional transistors. *Nat. Nanotechnol.* **2019**, *14*, 223–226.

[54] *Research Center Jülich GmbH, HNF - Helmholtz Nano Facility, J. Large Scale Res. Facil. (JLSRF)* **2017**, *3*, A112.

Supporting Information: Unveiling valley lifetimes of free charge carriers in monolayer WSe$_2$

Manfred Ersfeld,[1] Frank Volmer,[1] Lars Rathmann,[1] Luca Kotewitz,[1] Maximilian
Heithoff,[1] Mark Lohmann,[2] Bowen Yang,[3] Kenji Watanabe,[4] Takashi Taniguchi,[4]
Ludwig Bartels,[3] Jing Shi,[2] Christoph Stampfer,[1,5] and Bernd Beschoten[1]

[1]*2nd Institute of Physics and JARA-FIT, RWTH Aachen University, 52074 Aachen, Germany*
[2]*Department of Physics and Astronomy, University of California, Riverside, California 92521, USA*
[3]*Department of Chemistry and Materials Science & Engineering Program,*
University of California, Riverside, California 92521, USA
[4]*National Institute for Materials Science, 1-1 Namiki, Tsukuba, 305-0044, Japan*
[5]*Peter Grünberg Institute (PGI-9), Forschungszentrum Jülich, 52425 Jülich, Germany*

This Supporting Information contains a description of the device fabrication in section I, a discussion on how to diminish the impact of the laser-induced screening of the gate electric field in section II, a discussion of the fitting method in section III, a discussion that the lifetime which we attribute to the valley polarization cannot be explained by a dark exciton lifetime in section IV, additional explanations to the figures shown in the main manuscript in section V, and a schematic of the used TRKR setup in section VI.

I. DETAILED INFORMATION ABOUT DEVICE FABRICATION

Fig. S1a shows an optical image and Fig. S1b the corresponding schematics of the WSe$_2$ sample used for the measurement presented in Fig. 1 of the main manuscript. It consists of a monolayer WSe$_2$ flake encapsulated by hBN and electrically contacted with graphitic leads. The sample is placed on a Si^{++}/SiO$_2$ substrate with an oxide thickness $d_{ox} \approx 300$ nm, which was used to apply a back-gate voltage. The ML WSe$_2$ flake was obtained by a standard exfoliation technique from a commercially available bulk crystal (HQ Graphene) onto a polydimethysiloxane (PDMS) stamp. The ML was identified from optical contrast, Raman spectroscopy and in-situ PL response. Afterwards, the flake was covered with a top-hBN flake using a standard dry-transfer process. By means of reactive ion etching (RIE) we etched the flake into a rectangular shape. RIE was done with a gas mixture of O$_2$/CF$_4$ at a ratio of 1:4.

The subsequent fabrication steps, which are necessary to incorporate graphitic contacts, are depicted in Fig. S2. First, we built an hBN/graphite heterostructure on a Si^{++}/SiO$_2$ substrate. Then, the inverse image of the contact geometry was written by means of electron beam lithography into a spin-coated layer of polymethylmethacrylate (PMMA), which after development served as an etching mask for the following RIE step. We employed a selective RIE process using only O$_2$ that allows to etch graphite at a much higher rate than hBN. However, this process still is supposed to attack the hBN, resulting in a rougher surface inducing more strain than

normally expected from an atomically flat hBN crystals.

In the next step, the WSe$_2$/hBN half-sandwich was picked up at 130°C with a PDMS/polycarbonate (PC) stamp and transferred onto the pre-patterned contact geometry by melting the PC membrane to the substrate. Finally, after dissolving the PC membrane with chloroform, the metallic leads to the graphitic contacts were patterned in a second electron beam lithography step before depositing Cr(5 nm)/Au(50 nm) contacts.

The MoSe$_2$ device used for the measurement discussed in Fig. 3 of the main manuscript, is a half-sandwiched ML MoSe$_2$ device contacted with graphitic leads used for transport measurements. Fig. S1c shows an optical image of this device before contacting one side of the transparent graphitic top-gate via a Cr/Au electrode. The MoSe$_2$ ML is highlighted by the orange dashed line and the top hBN flake by the green dashed line. The gate modulation in this device was achieved by a transparent graphitic top-gate (white dashed line). The MoSe$_2$ monolayer was exfoliated and covered with hBN in the same manner as described for the WSe$_2$ ML.

The fabrication steps for patterning the graphitic contacts are depicted in Fig. S3. The graphite was exfoliated onto a Si^{++}/SiO$_2$ substrate. Afterwards, the contact geometry was written by e-beam lithography as a positive image into a PMMA layer to define an aluminum hard mask. This hard mask was used for RIE etching with pure O$_2$. The aluminum mask was removed with tetramethylammoniumhydroxid (TMAH) and the MoSe$_2$/hBN half-sandwich was transferred onto the contacts in the same way as described for the WSe$_2$ device. The graphitic top-gate was added to the device with a similar hot-pick-up and dry transfer method. Finally, the graphitic top-gate and contacts were contacted by Cr/Au leads.

Fig. S1d shows an optical image of the CVD-grown WSe$_2$ device used for the measurements shown in Fig. 4 of the main manuscript. Directly after growth, an hBN flake was placed on the WSe$_2$ to protect it from degradation. Using the before-mentioned hot-pickup technique, the hBN/WSe$_2$ stack was transferred onto a Si^{++}/SiO$_2$(300 nm) substrate with prepatterned Cr/Au electrodes.

We note that all fabrication steps were down under ambient clean room conditions.
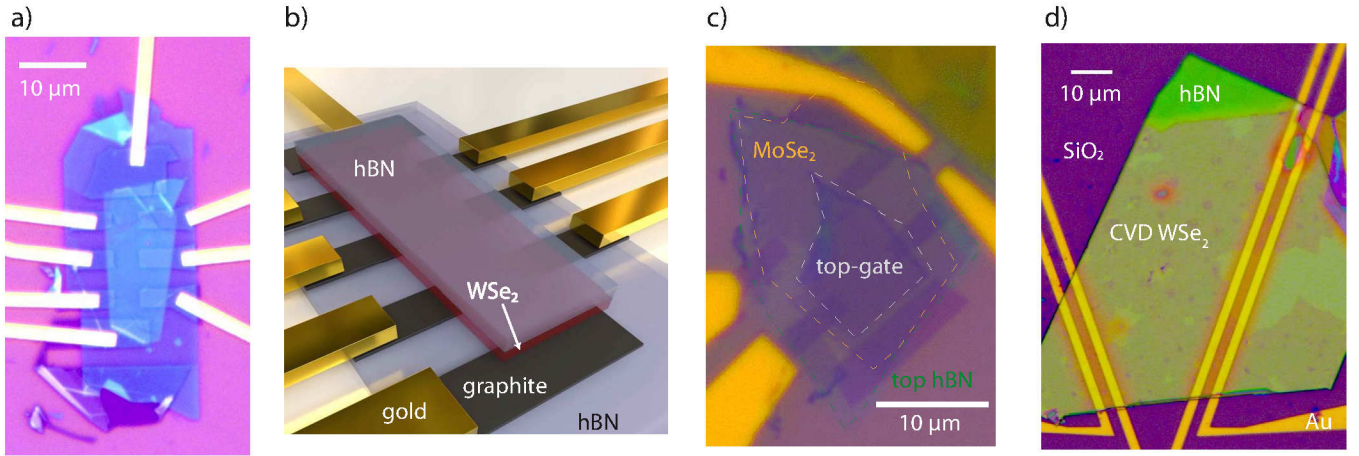
FIG. S1. (a) Optical microscope image of the WSe$_2$ device used for the measurements in Fig. 1 of the main manuscript and (b) schematic representation of the device structure. (c) Optical microscope image of the MoSe$_2$ sample used for the measurement in Fig. 3 of the main manuscript. The image was taken before contacting the transparent graphitic top-gate via a Cr/Au electrode. (c) Optical microscope image of the CVD-grown WSe$_2$ sample used for the measurements in Fig. 4 of the main manuscript.
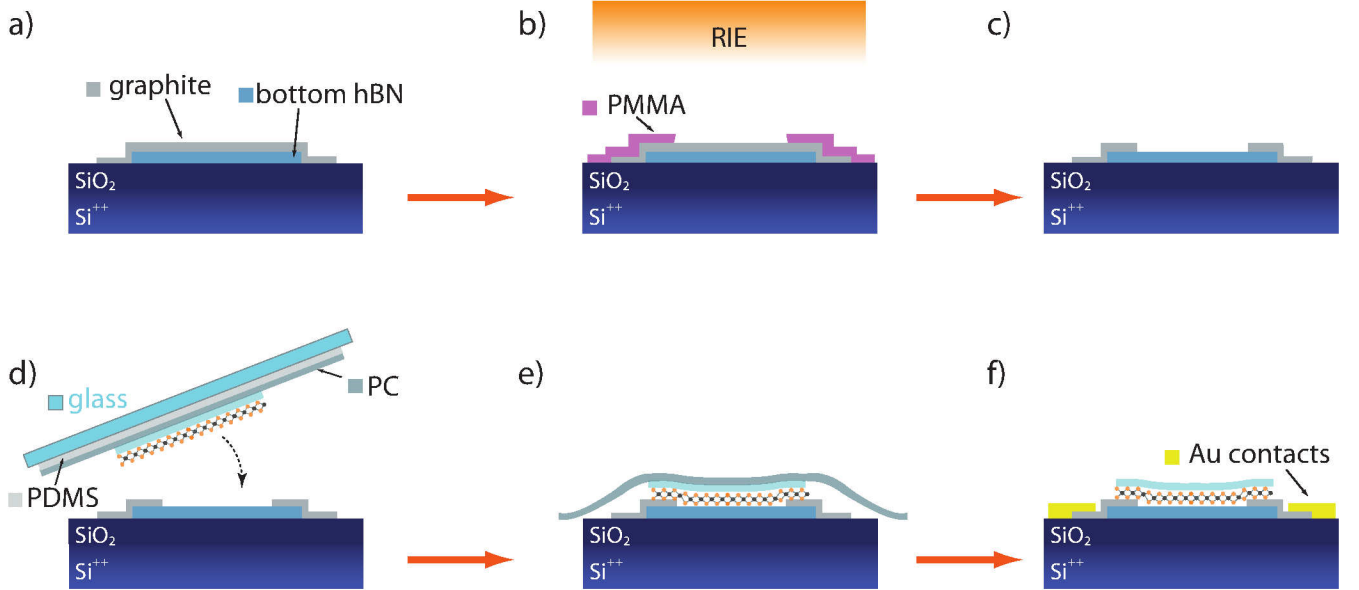

FIG. S2. Fabrication process of the WSe$_2$ device used for the measurements shown in Fig. 1 of the main manuscript. (a) Graphite/hBN stack on a Si^{++}/SiO$_2$ substrate. (b) Reactive ion etching of the contact geometry as defined in an electron-beam lithography step. The PMMA layer serves as an etching mask. (c) Device geometry after etching. The PMMA mask layer was removed with aceton. (d) Transfer of the WSe$_2$/hBN half-sandwich onto the contact geometry. The half-sandwich was picked up with a hot pick-up method using a PDMS/PC stamp. (e) Deposition of the half-sandwich by melting the transfer PC membrane to the substrate. Afterwards, the PC was dissolved in chloroform. (f) Growth of Cr/Au for contacting the graphitic leads to finalize the device.
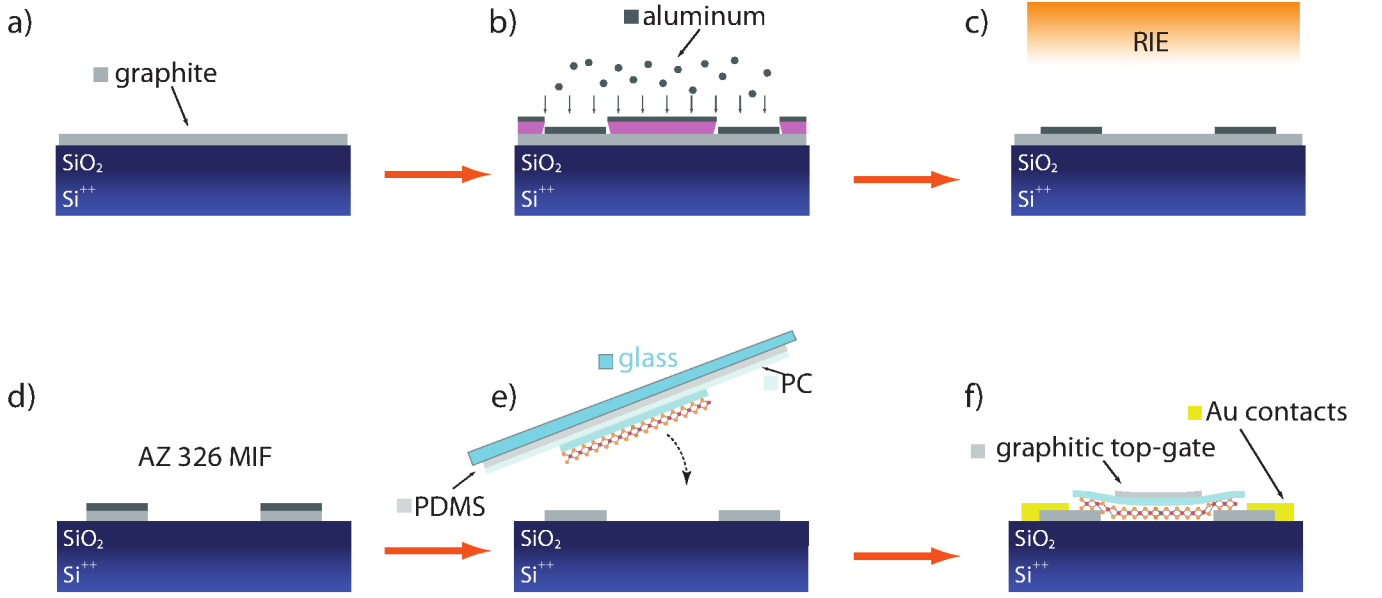
FIG. S3. Fabrication process of the MoSe$_2$ sample which measurement results are shown in Fig. 3 of the main manuscript. (a) Graphite exfoliated onto a Si^{++}/SiO$_2$ substrate. (b) Aluminum hard mask deposition (25 nm) onto a mask written with electron-beam lithography into a PMMA layer. (c) After the aluminum lift-off the contacts were etched with pure O$_2$-RIE. (d) Device geometry before aluminum hard mask removal with TMAH (AZ 326 MIF). (e) Deposition of the hBN/MoSe$_2$ half-sandwich onto the contacts by using a hot-pick up method with a PDMS/PC stamp. The membrane is molten to the substrate and dissolved with chloroform. (f) Adding of a graphitic top-gate similar to the technique in (e) before depositing Cr/Au contacts to the graphitic leads to finalize the device.

II. HOW TO DIMINISH THE PHOTO-INDUCED SCREENING OF THE GATE-ELECTRIC FIELD

To unveil the true gate-dependent spin and valley dynamics, the standard optical measurement techniques have to be modified to account for the screening of the gate electric field by photo-excited charged defects in the dielectric layer.[1–3] We demonstrate that disregarding this procedure can lead to erroneous conclusions drawn from gate-dependent measurements.

Donor and acceptor like defect states in hBN and SiO$_2$ (or defect states at the interface between these insulators and a 2D material) can be optically active for photons energies in the visible spectral range.[1–3] In the presence of a gate electric field, which is applied across the dielectric layer, photo-excited charge carriers can either tunnel towards the 2D material or the gate electrode, leaving behind a charged defect state. This charged defect state will then screen the gate-electric field, which in turn will change the effective gate-induced charge carrier density in the 2D-material.[1–3]

Furthermore, photo-induced changes in the properties of the TMD device can occur on significantly different time-scales.[4–6] This complicates the optical measurement of the gate-dependent valley and spin dynamics: Once the gate-voltage is set to a certain value, the time-dependent photo-induced screening of the gate-electric field will lead to an unwanted change in gate-induced charge carrier densities and therefore a shift of the Fermi level over time. In particular, this is an issue in the "conventional" measurement technique where the gate-voltage is set to a fixed value and then the full TRKR curve gets recorded by varying the time delay Δt between pump and probe pulses over laboratory time. This is schematically depicted in Fig. S4a, where we assume that for each TRKR curve the Kerr rotation amplitude is measured at ten successive time delays before the next gate voltage is applied.

In this method, data point #1 is recorded right after setting the gate voltage, whereas the last data point of this curve (#10 in this example) may be recorded minutes later (this is the time we need to measure our actual TRKR curves, which consist of up to 200 data point measured at different time delays). But during this measurement time, the photo-doping effect will continuously shift the Fermi-level. As a result, the data points within a single TRKR curves are recorded at different charge carrier densities.

To circumvent this problem, we adapted our measurement scheme by first setting a time delay and then record the Kerr rotation signal for a full gate voltage sweep (see numbering of data points in Fig. S4b). For maximum comparability, we use the same gate sweep direction and velocity for all TRKR, PL, and electrical measurements, which ensures the photo-doping effect impacts all measurements in the same way.

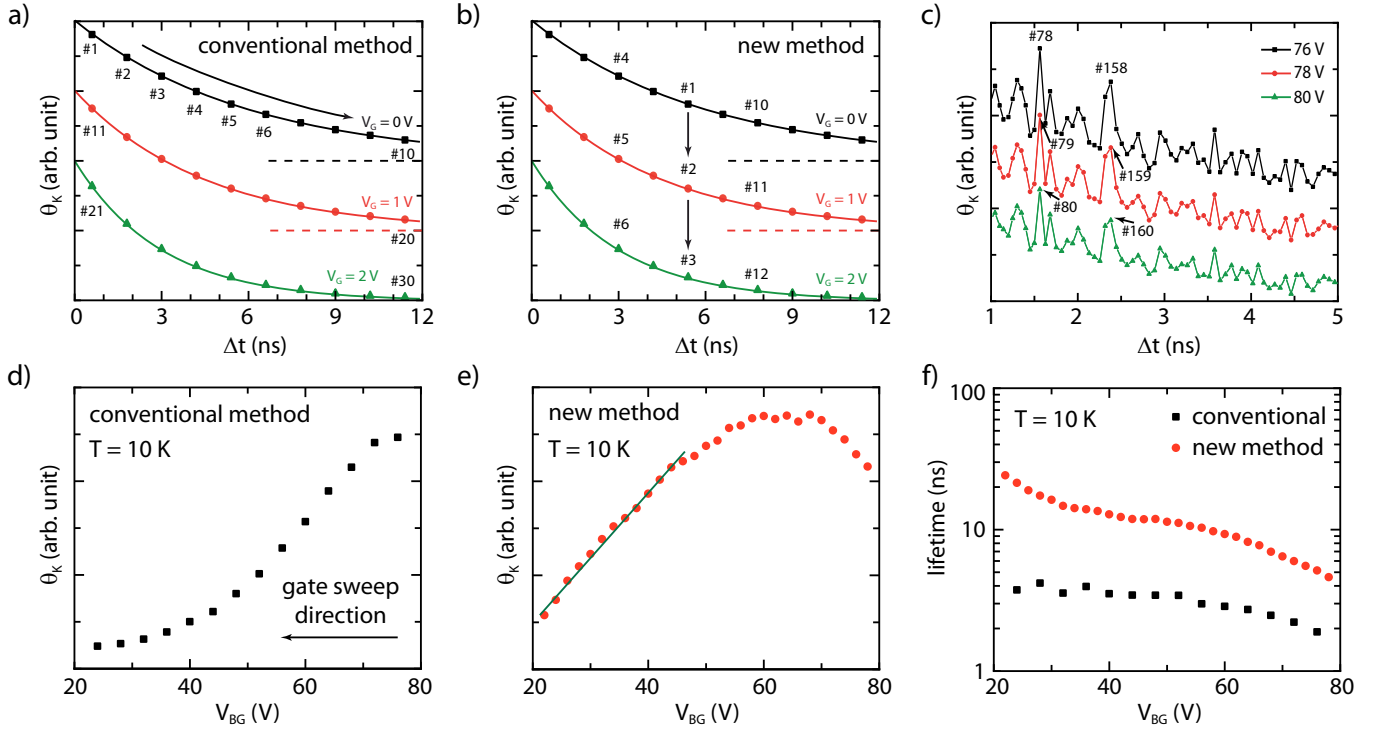With this method, each data point of a specific TRKR

FIG. S4. (a) Schematic representation of the conventional TRKR measurement scheme where the time delay between pump and probe laser pulses is successively increased from $\Delta t = 0$ ns to 12 ns before the next TRKR trace gets recorded for the next gate voltage (see numbering of data points). Due to temporal changes in the screening of the gate-electric field on laboratory time scales, subsequent data point measured at the same gate-voltage might be recorded at a slightly different charge carrier densities. (b) Schematic representation of our new TRKR measuring scheme, in which a full gate-sweep is performed for each time delay (see numbering of data points). (c) As the time delays in this new method are distributed randomly, a change of device properties over lab time can be identified by the same apparent "noise" signal for different gate voltages. (d) to (f) Direct comparison between the conventional and the new measurement method. Both gate-dependent TRKR measurements were conducted right after each other on the same day, but yield significantly different amplitudes, lifetimes and overall gate-dependent changes.

curve has the same "history", i.e., each data point is not only recorded exactly at the same time after setting the corresponding gate voltage, but also the temporal course of both setting and staying at previous gate voltages during the continuous illumination of the sample is exactly the same. Therefore, we can be certain that all data point of a specific TRKR curve are in fact recorded at the same charge carrier density.

In our new method we also randomize the sequence of the time delays (see numbering of data points in Fig. S4b). This is an important step for identifying changes of the device properties on laboratory time-scales, which are especially pronounced within the first few hours of each new cooling cycle. We suspect that this temporal change is due to an degassing of initially adsorbed molecules on top of the devices due to laser illumination.[7] In this context, we note that one cooling cycle usually extends over five days before we have to change the liquid helium container. At this point, the device necessarily has to be warmed up to room temperature and residuals from gas permeation through the elastomer sealing gaskets, micro-leakages or other gas loads my absorb on the devices.

When using the conventional measurement scheme, the temporal change of the device properties due to the degassing of absorbed molecules over laboratory time may emerge in a TRKR measurement as an exponentially decaying signal and therefore can lead to a misinterpretation of the measurement. Instead, in our new method this change over laboratory time is randomly distributed over the whole range of measured delay times. Therefore, a change in device properties over lab time can easily be identified as an apparent, highly reproducible "noise" signal when comparing delay scans at different gate voltages. This is depicted in Fig. S4c where we show a gate-dependent TRKR measurements recorded right after cooling down the device from ambient conditions. The TRKR curves for the three gate voltages are vertically shifted for better visibility. It becomes obvious that the apparent noise of these curves is almost identical, showing that in fact the device slowly changes its properties over time and that the change in Kerr rotation amplitude due to photo-induced effects is more pronounced than the actual noise. Especially the first two measurements at time delays of around 1.6 ns and 2.4 ns are clearly visible as peaks because the exponentially decreasing amplitude over laboratory time has its strongest impact. We discard

measurements as shown in Fig. S4c and wait until the device properties have settled to such an extent that the random noise is visible between different traces.

A direct comparison between both methods is depicted in Figs. S4d to S4f, where we plot the Kerr rotation amplitude and lifetime of the Kerr rotation signal we assign to the valley polarization of free charge carriers. Both gate dependent TRKR measurements were taken directly one after the other on the first day of a new cooling cycle, first by the conventional method, then by the new method. For the former, the first TRKR curve was measured at a gate voltage close to 80 V and then the gate voltage was continuously decreased to 20 V. The extracted Kerr rotation amplitude almost decays exponential in the gate sweep direction (see Fig. S4d). This is a clear indication of the exponentially decaying change in device properties due to the photo-induced effects. In contrast, with our new method we clearly observe the linear increase in polarization with increasing gate voltage (see green line in Fig. S4e) which is expected from a valley polarization of free charge carriers.

Next to the gate-dependent Kerr rotation amplitudes, also the extracted lifetimes differ significantly between both measurement methods as shown in Fig. S4f. The apparent lifetimes in case of the conventional method are much shorter than the ones obtained from the new method. This discrepancy is due to the fact that in case of the conventional method the exponentially decreasing change in device properties over laboratory time is projected onto the sweep velocity of the delay-time: A slow measurement of the time-delay trace would project the largest change in the exponentially decreasing change of device properties over laboratory time into the first few measured time delays, yielding an apparent short lifetime. A fast measurement of the time-delay trace would instead distribute the same change in device properties over a larger span of measured time-delays, therefore yielding a longer apparent lifetime.

Finally, we note that the photo-induced screening of the gate field leads to a minor complication in the comparison of gate-dependent PL and TRKR data. As stated before, we use the same gate sweep direction and velocity for all TRKR and PL measurements, which ensures the photo-induced effects impact all measurements in a similar way. The only difference stems from the laser excitation energy which was around 1.7 eV in case of TRKR and 2.33 eV in case of PL measurements. But a higher laser energy yields a stronger photo-doping effect as additional, energetically lower defect states can be excited.[1] This explains why the three regions identified in both TRKR and PL data (the gap region, the tail region, and the conduction band region) seem to slightly differ by a few volts in the back gate voltage dependent measurements in Fig. 1 of the main manuscript.

III. FITTING METHOD AND RESULTS

The lifetimes τ_i ($i = v$ for valley or $i = b$ for bound exciton) and Kerr rotation amplitudes Θ_i in the main manuscript were acquired by fitting the TRKR curves by a two-exponential function of the form:

$$\Theta_{\mathrm{K}}(t) = \Theta_{\mathrm{v}} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{v}}}\right) + \Theta_{\mathrm{b}} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{b}}}\right) \quad (1)$$

However, when fitting the TRKR curves, we have to consider that at low temperatures the long-lived bound exciton polarization is much longer than the laser repetition interval of $T_{\mathrm{rep}} = 12.5$ ns. This leads to a non-zero Kerr rotation signal slightly before the next pump pulse reaches the sample at $\Delta t = 0$ ns (see e.g. orange curve in Fig. 1d or the schematic in Fig. S5(a)). Therefore, we now consider the influence of the subsequent pump pulse on an already existing exciton population stemming from previous pulses. For this, we assume that the pump pulse will have a negligible effect on the already existing excitons. Accordingly, the TRKR data can be fitted by a sum over several pulses and a bi-exponential decay of the form:

$$\Theta_{\mathrm{K}}(t) = \Theta_{\mathrm{v}} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{v}}}\right) \quad (2)$$
$$+ \sum_n \Theta'_{\mathrm{b}} \cdot \exp\left(-\frac{t + nT_{\mathrm{rep}}}{\tau_{\mathrm{b}}}\right)$$

We note that equations 1 and 2 yield exactly the same lifetimes τ_i as the terms which depend on the time t are mathematically identical. The fitting procedures only results in different Kerr rotation amplitudes of the bound exciton polarization. Factoring out the exponential term for the bound exciton signal in equation 2 yields the following relationship:

$$\Theta_{\mathrm{b}} = \Theta'_{\mathrm{b}} \sum_n \exp\left(-\frac{n \cdot T_{\mathrm{rep}}}{\tau_{\mathrm{b}}}\right) \quad (3)$$

The two amplitudes Θ_{b} and Θ'_{b} are schematically depicted in Fig. S5a. Here the solid black curve is the TRKR curve which is the sum of the 1 ns long signal stemming from the valley polarization (black dashed line) and the longer-lived signal originating from the polarized bound excitons (red dashed line). In the main manuscript we plot Θ_{b}, i.e. the overall Kerr rotation amplitude of the bound exciton signal measured from $\Theta_{\mathrm{K}} = 0$ as this amplitude is the most intuitive one when comparing the fitted values to the raw data. According to equation 3 the amplitude Θ_{b} not only considers the impact of the last pump pulse, but is the weighted sum over all previous pulses. Instead the amplitude Θ'_{b} is the actual increase in the long-lived TRKR signal created by each individual pulse (see green arrow in Fig. S5a). Fig. S5b shows the same data as Fig. 1e of the main manuscript, which was obtained by using the fitting function of equation 1. Instead, Fig. S5c depicts results of the
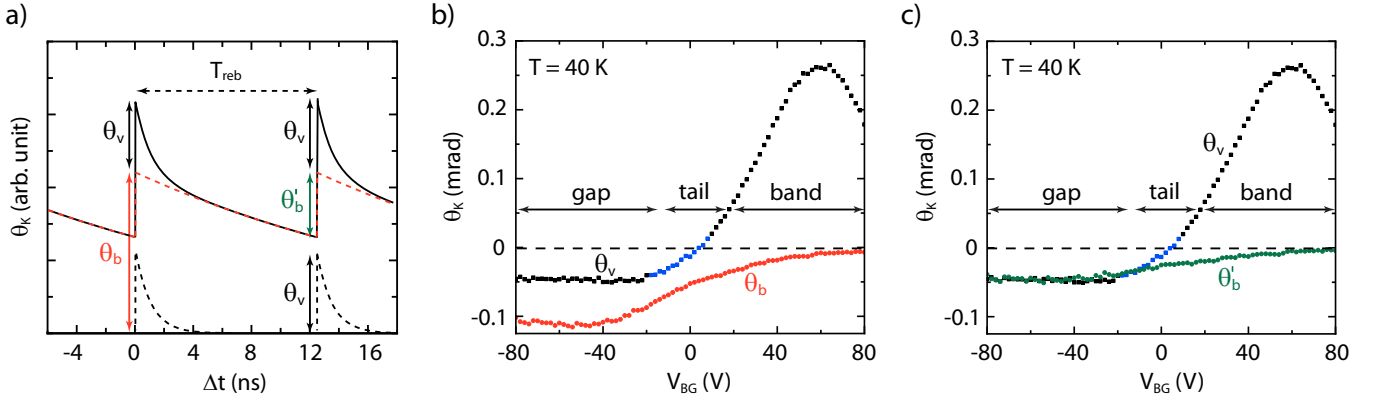
FIG. S5. (a) Schematic representation of the different Kerr rotation amplitudes obtained by the two fitting functions in equations 1 and 2, respectively. Θ'_b (green arrow) only depicts the increase in the long-lived signal induced by the current pump pulse, whereas Θ_b represents the total amplitude, which also includes leftover polarization from previous pump pulses. (b) and (c) Direct comparison of the results using the two fitting functions.

same set of raw data when using equation 2 for fitting. It is important to emphasize that using equation 2 yields $\Theta'_\mathrm{b} \approx \Theta_\mathrm{v}$ within the band gap region. This is relevant when considering our model depicted in Fig. 2 of the main manuscript. There we assumed that the electron which was caught during the formation of the bound trion (Fig. 2f) will go back into the same valley from which it came during the bound trions recombination (Fig. 2h). In this case, the valley polarization created by these two processes should be comparable (Figs. 2g and 2i). Of course, the physical situation depicted in Figs. 2g and 2i is quite different, as there are additional bound excitons in case of Fig. 2g. These additional quasi-particles will most likely have an impact on the Kerr rotation amplitude, as they are not only polarized but also may energetically shift the Kerr resonance due to band-gap renormalization effects. This makes a direct comparison between the expected Kerr rotation amplitudes between the two situations depicted in Figs. 2g and 2i quite challenging. Nevertheless, we note one common connection seen in all of our samples shown in this study: Next to long-lived polarization in the band gap region, there is always a shorter-lived signal with an amplitude which is similar to the amplitude increase of the long-lived Kerr signal created by the new pump pulse, i.e. Θ'_b. This observation is in favor of our proposed model.

Next, we note that the amplitude of the valley polarization Θ_v is most likely underestimated in the band gap region by the used fitting model. To illustrate this, we simulate the data shown in Figs. 1e and 1f based on the model depicted in Fig. 2. We use a simplified three level rate equation system representing the bound excitons $|B\rangle$, the valley polarization $|V\rangle$ and a ground state $|G\rangle$ (see Fig. S6a). The coupled rate equation reads:

$$\frac{dN_\mathrm{B}}{dt} = -\frac{1}{\tau_\mathrm{rec}^{bound}} N_\mathrm{B} \tag{4}$$

$$\frac{dN_\mathrm{V}}{dt} = -\frac{1}{\tau_\mathrm{v}} N_\mathrm{V} + \frac{1}{\tau_\mathrm{rec}^{bound}} N_\mathrm{B} \tag{5}$$

with N_i (i = B, V) denoting the occupation numbers of the respective levels, $\tau_\mathrm{rec}^{bound}$ the recombination lifetime of the bound excitons, and τ_v the valley lifetime.

We solve this system of coupled rate equations by numerical integration. First, we focus on the process responsible for the creation of the valley polarization during the recombination of the bound excitons (illustrated in Figs. 2h to 2i). We therefore initialize the system by putting 10^5 excitons in the $|B\rangle$ level and setting the occupation numbers of the other levels to zero (see Fig. S6a). The lifetime $\tau_\mathrm{rec}^{bound}$ of the bound exciton level is set to 24 ns in accordance to the measurements in the band gap region of Fig. 1f. The resulting temporal occupation of $|V\rangle$ is simulated for different valley lifetimes τ_v (see Fig. S6b). We note that the occupation of the level $|V\rangle$ is most relevant for the Kerr rotation signal. This becomes obvious as a Kerr rotation can only be measured for probe pulse energies tuned into the bright trion regime. As explained in the main manuscript, this is due to the fact that the additional charge carrier, which is needed for the formation of the trion, has to come from the bands. Hence, the valley-selective excitation of trions can directly probe a valley polarization of free charge carriers. Consistent to this, no Kerr rotation signal is observed when probing at the energetically lower bound exciton energies. Hence, we only plot the occupation of the level $|V\rangle$ and not the one of $|B\rangle$.

The corresponding temporal evolution of the $|V\rangle$ occupations in Fig. S6b show three distinct regimes: As the level $|V\rangle$ is empty at $\Delta t = 0$, there is an initial increase in occupation number with the time constant of τ_v. The occupation number then shows a maximum after approximately $3 \cdot \tau_\mathrm{v}$. At this point the filling of this level through level $|B\rangle$ with the time constant $\tau_\mathrm{rec}^{bound}$ is equal to its emptying to the ground state with the time constant τ_v. Afterwards, the occupation number of $|V\rangle$ slowly decreases. We fit this decrease by an exponential function and confirm that for $\tau_\mathrm{v} \ll \tau_\mathrm{rec}^{bound}$ the decay time
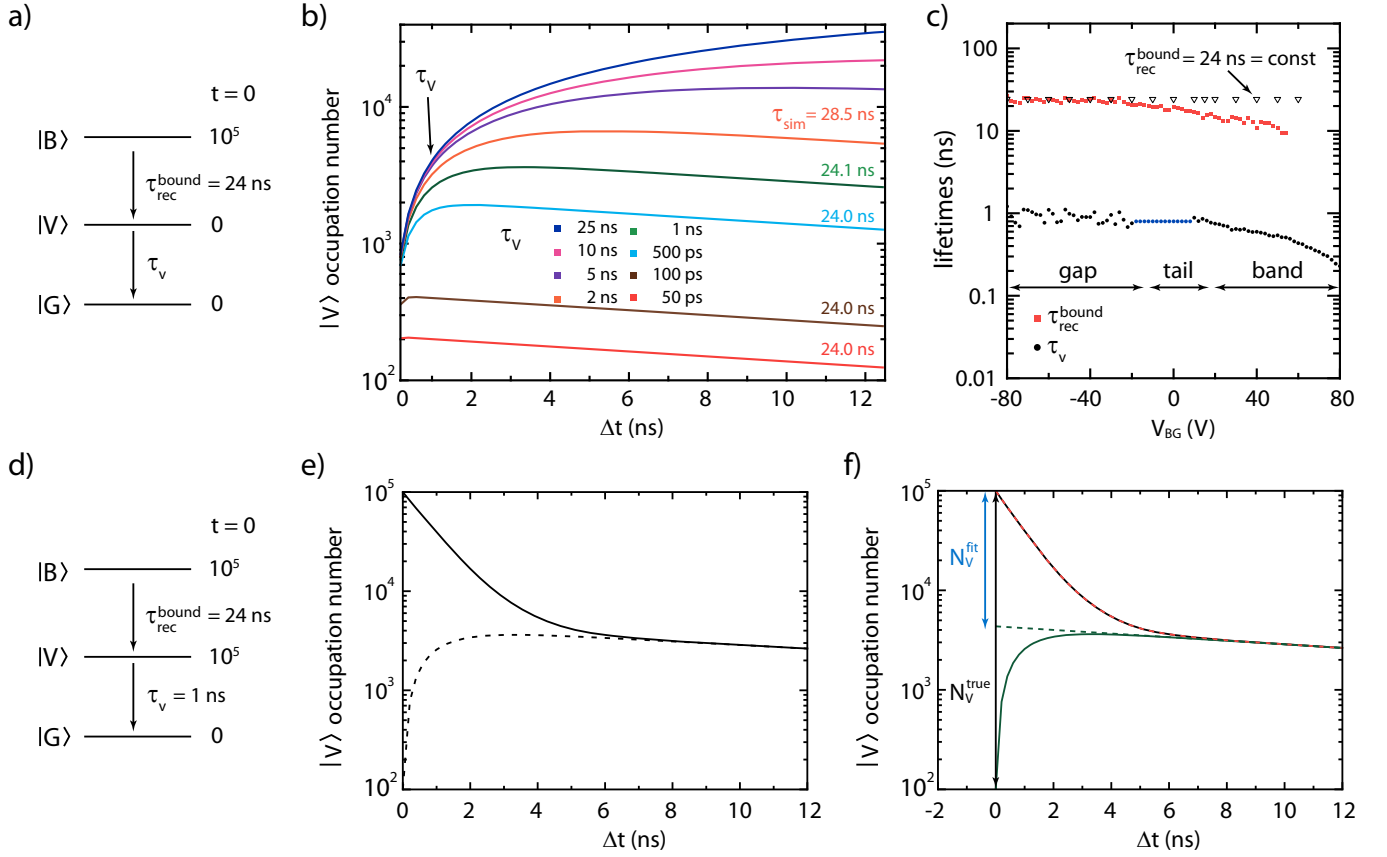
FIG. S6. (a) A three state model used to simulate the lifetime of the Kerr rotation originating from the bound exciton states. Bound excitons recombine with a lifetime of $\tau_{\mathrm{rec}}^{bound}$ and polarize the valleys during the recombination process (see Fig. 2h). The valley polarization then relaxes into the ground state via a rate of $1/\tau_{\mathrm{v}}$. (b) Results from numerical simulation for an initial configuration of $N_V = 0$ and $N_B = 10^5$ at $t = 0$ for different τ_{v}. (c) Same data as in Fig. 1f. (d) to (f) Results of the three state model by additionally assuming an initial occupation of the $|V\rangle$ state at $t = 0$. Further explanation see text.

is equal to the recombination time of bound excitons used for the simulation ($\tau_{\mathrm{rec}}^{bound} = 24$ ns). This is also expected, as under such conditions the number of particles in $|V\rangle$ is only determined by the number of most recent transitions from $|B\rangle$ to $|V\rangle$ via the time $\tau_{\mathrm{rec}}^{bound}$. All other particles in $|V\rangle$ have already decayed into the ground state. The fitted recombination time only deviates from the input value if τ_{v} approaches $\tau_{\mathrm{rec}}^{bound}$. As the measured valley lifetime in Fig. 1f ($\tau_{\mathrm{v}} = 1$ ns) is one order of magnitude shorter than the long-lived bound exciton lifetimes (24 ns), we are confident that this long-lived Kerr rotation signal represents the bound exciton recombination time $\tau_{\mathrm{rec}}^{bound}$ (see green curve in Fig. S6b).

This conclusion is also important for the interpretation of the gate-dependent lifetimes of bound excitons (red data points in Fig. 1f and Fig. S6c), as the decrease in lifetimes starting in the tail region cannot be due to the decreasing valley lifetime. This is illustrated in Fig. S6c by the open triangular data points, for which we assume a gate-voltage independent bound exciton lifetime of 24 ns and simulated the impact of the decreasing valley lifetimes on the fitted values. As it is also depicted in Fig. S6b, a decreasing valley lifetime from 1 ns (green curve) to

50 ps (red curve) changes the fitted exciton lifetimes only marginally by a maximum of 0.1 ns. Accordingly, the actual bound exciton recombination time $\tau_{\mathrm{rec}}^{bound}$ has to decrease towards higher gate voltages (i.e. higher free charge carrier densities). This might be due to both screening and band gap renormalization effects.

Of course, we do not observe the increase in polarization seen in Fig. S6b in the raw data (see Fig. 1d). This is due to the fact that we have neglected that the creation of the bound trions also directly creates a valley polarization (see Fig. 2f and 2g). Hence, we extent our simulation by initializing both the $|V\rangle$ and the $|B\rangle$ states with the same initial occupation number (Fig. S6d) and assume a valley lifetime of 1 ns. The resulting time evolution of the $|V\rangle$ state is depicted in Fig. S6e as a solid line and shows an exponential decrease like in our TRKR data. For comparison, we also include the curve from Fig. S6f for the same valley lifetime as a dashed line.

The simulated curve can be fitted perfectly by the bi-exponential function of equation 1 (see dashed red line in Fig. S6f). Most importantly, the fit yields exactly the two lifetimes put into the simulation, i.e. 1 ns for the valley lifetime and 24 ns for the bound exciton lieftime. This is

not self-evident, as the fit function assumes an erroneous time-dependence of the long-lived signal (green dashed line in Fig. S6f) compared to the actual time dependent amplitude according to our model (solid green line). The most important impact of this analysis is that the Kerr rotation amplitude of the valley polarization at $t = 0$ gets underestimated: Instead of $N_{\mathrm{V}}^{\mathrm{true}}$, we only determine the amplitude $N_{\mathrm{V}}^{\mathrm{fit}}$. Hence, the real amplitude of the valley polarization Θ_{V} is most likely larger than Θ_{b}' in Fig. S5c.

A. Additional TRKR signals for $\Delta t < 1$ ns

The rate equation system in the last section, which described the valley polarization created by the recombination of bound excitons (Figs. 2f-2i in the main manuscript), can also be applied to the valley polarization created by the recombination of dark excitons (Figs. 2b-2d). Accordingly, the latter signal should exhibit a similar initial increase as the one created by the recombination of bound excitons (see dashed line for $\Delta t < 2$ ns in Fig. S6e). This initial increase would occur on the timescale of the dark exciton recombination time, which is reported to be between 100 ps and 1 ns at cryogenic temperatures,[8,9] and is therefore expected to be well below 1 ns for the data shown in Fig. 1d of the main manuscript which was measured at 40 K.

As for the case of the bound excitons, such an increase is not observed in our measurements (see Fig. 1d). Similar to the argumentation of the last section, we believe that this initial increase is masked by other signals with exponentially decreasing amplitudes and similar time constants. In this context we note that in our study we focus on the nanosecond time-scale to discuss the valley and bound exciton dynamics. Especially the fast bright exciton dynamics in the ps-range are not the scope of our study. This is the reason why the time step between two adjacent data points in Fig. 1d is around 50 ps. We nevertheless conducted some measurements at increased temporal resolution and frequently observed additional TRKR signals with lifetimes ranging between tens and hundreds of picoseconds (see e.g. Fig. S7). This is consistent to varying values of measured bright exciton lifetimes (both neutral and charged excitons) reported in literature.[10–13]

There is also another aspect, which may explain the absence of the initial increase in the TRKR signal belonging to the valley polarization. If dark trions were also contributing to the TRKR signal, their decreasing population would perfectly mask the increasing valley polarization. Of course, due to their nature we do not expect that dark trions directly interact with the photons of the probe pulse. Nevertheless, we speculate that a polarized dark trion population may influence the Kerr signal of both bright excitons and the valley polarization via two mechanisms: On the one hand, a pump-beam-induced valley-polarized dark trion population will bind a different amount of resident charge carriers in both K
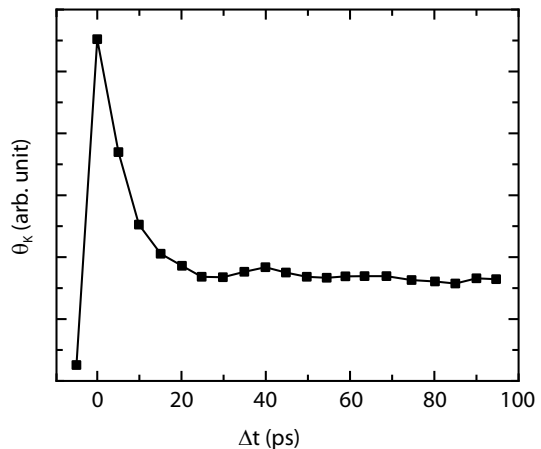

FIG. S7. Additional short-lived signal in the TRKR measurements, which can be detected by increasing the temporal resolution of the measurement. The signal is attributed to a lifetime of bright excitons.

and K' valleys. As these carriers are no longer available for the creation of bright trions, the excitation probability of bright trions between both valleys may be affected. This will lead to different absorption of right and left circularly polarized light and, hence, a Kerr signal (see detailed discussion of this model in Ref. 14). On the other hand, a dark exciton population may also shift the Kerr resonances of both the bright exciton and the valley polarization due to band-gap renormalization effects.[15,16] An exponentially decaying shift in the resonance curve due to an exponentially decaying dark exciton population can manifest itself in a decaying Kerr signal.

IV. DARK EXCITONS

We now discuss that the TRKR signal, which we attribute to the valley polarization (Fig. 1d) cannot be assigned to dark excitons. Overall, there are three arguments against such an assignment:

1.) By far the most important argument is that gate dependent PL measurements show a decreasing amplitude of dark exciton emission for increasing charge carrier densities.[9,17,18] This is in complete contradiction to the increase of the valley polarization with increasing charge carrier density as seen in Fig. 1e and 4a.

2.) Even at the lowest temperatures, gate-dependent dark exciton recombination times (at zero magnetic field) barely reach 1 ns.[8,9] Instead, we measure lifetimes of up to 30 ns (see Fig. 1f and 4b) for the Kerr rotation signal we attribute to the valley polarization.

3.) A recent publication based on spin-noise measurements excluded dark excitons as a source of long-lived signals in TRKR measurements.[19]

V. COMMENTS AND ADDITIONAL EXPLANATIONS TO THE FIGURES OF THE MAIN MANUSCRIPT

A. Fig. 1c

Schottky barriers: One main problem in contacting 2-dimensional semiconductors is the formation of Schottky barriers between the TMD flake and the contacts.[20–22] Depending on the sign of the applied dc bias voltage, either the drain or the source contact is therefore reversed biased, whereas the other contact is forward biased. In this respect, our graphitic contacts are no different and we observe pronounced Schottky barrier characteristics in all of our devices. Therefore, the main drop of the applied 5 V bias voltage will occur over the reversed biased Schottky barrier and not across the TMD channel.

Furthermore, we note that the measured resistance at a constant bias voltage varies significantly between different contacts even on the same device. Unfortunately, the impedance of some contacts are too high to even use these contacts as voltage probes. This is the reason that no Hall-effect measurement could be conducted on the device shown in Fig. S1(a) to determine the free charge carrier density as a function of gate-voltage.

X′_ feature: At high electron densities we observe the appearance of the so-called X′_ feature in the PL data, whose physical origin has not fully been determined. One possible explanation of this PL feature is a different interaction of excitons with low and high charge carrier densities, respectively. It is argued that a trion can only be considered as a three-particle object at low enough charge carrier densities, whereas at higher charge carrier densities the interaction between excitons and a Fermi sea of free electrons leads to an exciton-polaron picture.[23,24] However, we believe that this explanation is rather unlikely to explain the X′_ feature in our measurements, as the transition between the three-particle and the exciton-polaron picture is argued to occur at quite low charge carrier densities (the Fermi level position E_F measured form the conduction band minimum has to be much smaller than the trion binding energy E_T).[23] In contrast, we argue that the X′_ feature is due to the onset of the upper spin-split conduction band.[25] To support this picture, we estimate the position of the Fermi level E_F for the gate voltage range of the X′_ emission.

First, we model the ML TMD as a two-dimensional electron gas (2DEG). This assumption is justified as the electron's Fermi wavelength λ_F is much larger than the thickness of the monolayer TMD. As long as E_F is close enough to the conduction band minimums at the K/K' valleys the bands can be approximated to follow a nearly quadratic dispersion relation, giving rise to a well-defined effective mass m^*. Under these circumstances the position of the Fermi level as a function of charge carrier density is given as:[26]

$$E_\mathrm{F} = \frac{2\pi\hbar^2 n}{m^* g_s g_v} \,, \tag{6}$$

with the spin degeneracy g_s, the valley degeneracy g_v, the effective mass m^* and the reduced Planck's constant $\hbar$. To express the charge carrier density n as a function of gate voltage V_G, we use the standard plate capacitor model. As soon as the Fermi level is within the conduction or valence band of the monolayer TMD, we consider all effects due to the band gap, mid-gap states, tail states, and quantum capacitance by introducing the so-called threshold voltage V_th and write:[27–29]

$$n = \frac{\varepsilon_0 \varepsilon_r}{d \cdot e} \left(V_\mathrm{G} - V_\mathrm{th}\right) \,, \tag{7}$$

with ε_0 being the vacuum permittivity, ε_r the relative permittivity of the dielectric, e the elementary charge, and d the dielectric thickness. For the device shown in Fig. 1, we estimate the threshold voltage to be $V_\mathrm{th} \approx 10$ V. This is the gate voltage at which the current starts to increase linearly with increasing back gate voltage. Putting in all values, we end up with a charge carrier density of around $4 \cdot 10^{12}$ cm^{-2} at a gate voltage of 80 V.

With this result the position of the Fermi energy at this gate voltage can now be calculated from equation 6. In case of a ML TMD the spin degeneracy is lifted ($g_s = 1$) and only valley degeneracy has to be considered ($g_v = 2$). We furthermore assume $m^* = 0.28 \cdot m_\mathrm{e}$,[30] which finally yields a Fermi energy of about 40 meV. This energy is well within the range of the assumed spin splitting between lower and upper conduction band.[30,31]

B. Figs. 1d, 1e and 1f

The reason for showing 40 K **data:** At 40 K the lifetime of the valley polarization is significantly shorter than the lifetime of the bound exciton signal over the whole gate voltage range (see Fig. 1f). This is necessary to reliable extract both lifetimes from a two-exponential fit to the raw data. However, at lower temperatures (see Fig. 1g) the lifetime of the valley polarization approaches the lifetime of the bound excitons especially at low charge carrier densities (i.e. low gate voltages). Therefore, it gets difficult to distinguish both signals in a TRKR curve and fitting gets challenging. We e.g. attribute the small decrease in the bound exciton lifetimes towards 0 V backgate voltage seen for 10 K and 20 K to this fitting issue. Instead, for all other temperatures, at which valley and bound exciton lifetimes are far enough apart, the bound exciton lifetimes are almost constant over the depicted gate voltage range.

Blue data points: As seen in Fig. 1e, the Kerr rotation signal which we attributed to the valley polarization of free charge carriers undergoes a sign reversal. Because

of its small amplitude, the fitting gets very error-prone in the region where it crosses the $\theta_K = 0$ axis. Nevertheless, the fitting routine yields the smooth transition of the Kerr rotation amplitude as seen in Fig. 1e, if the degree of freedom for the fit is reduced. This was done by assuming a fixed valley lifetime in this gate voltage regime (see blue data points in Fig. 1f) which allows for a continuous transition between the positive and negative gate voltages.

C.   Fig. 2

We note that Figs. 2b-i only show the kind of relaxation mechanisms that are responsible for the formation of a valley polarization of free charge carriers. These are the ones most relevant for the explanation of our experimental data. We note, however, that the bright exciton population created in Fig. 2a can also follow a few other scattering and relaxation channels, such as direct recombination, Auger-type exciton-exciton annihilation or the formation of other exciton complexes such as biexcitons.[10]

Furthermore, in Figs. 2b-i we only depict those scattering and relaxation processes which we deem to be the most important ones with the highest transition probabilities. So far, we neglected other possible transitions which in some cases can create a valley polarization with an opposite sign to the polarizations discussed in Figs. 2b-i. One such case would concern the transition shown in Fig. 2f. In principle, the bound exciton can capture the excess electron not from the K-valley, as it is shown and discussed so far, but instead from the K'-valley. This would result in an inverted valley polarization both in Figs. 2g and 2i. But in this case, the bound trion would be in a triplet state (all three spins would be orientated in the same direction). This state has a higher energy than the singlet state depicted in Fig. 2g. We therefore assume the transition probability to the singlet state to exceed that to the triplet state. Although both transitions are possible, any imbalance between the transitions probabilities will result in a measurable net valley polarization with a sign corresponding to the dominating transition.

D.   Figs. 2d and 2g

As it is seen by the black data points in Fig. 1e, the measured valley polarization shows opposite sign in the gap and conduction band regimes. At first sight, this may appear to contradict the model proposed in Fig. 2: Fig. 2d depicts the valley polarization in the case that the Fermi level is tuned into the conduction band, whereas Fig. 2g depicts the case that the valley polarization is in the gap region. As the polarizations in the conduction bands show the same sign in Figs. 2d and 2g, the change in sign in the experimental data is unexpected. However, the physical situation depicted in Figs. 2g and 2i is quite different, as there are additional bound excitons in case

of Fig. 2g. These additional quasi-particles will have an impact on the Kerr rotation amplitude of the valley polarization, as the bound excitons are not only polarized but will also shift the Kerr resonance energetically due to band-gap renormalization effects. In this context, we note that in our new measurement technique, which diminishes the photo-induced screening of the gate-electric field (see section II), the same pump and probe energies have to be used for every gate voltage. This makes it impossible to account for a shift in the Kerr resonance dependent on the gate voltage and, hence, renders a direct comparison between the expected Kerr rotation amplitudes in the two situations depicted in Figs. 2g and 2i challenging.
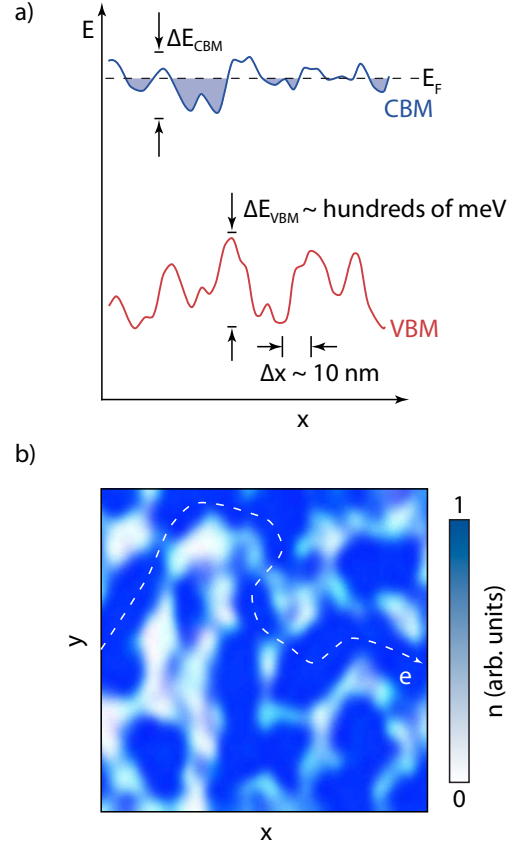

FIG. S8. (a) Due to strain variations and substrate corrugations, monolayer TMD flakes can show conduction band minima (CBM) and valence band maxima (VBM) energies which can spatially vary by hundreds of meV over lengthscales as small as 10 nm,[32] resulting in the appearance of charge puddles depicted as shaded areas. (b) Two-dimensional representation of the resulting charge puddles with a percolating electron path around the areas in which the Fermi level is still in the band gap (white areas).

E.   Fig. 3 in general

As discussed in the main manuscript, there are no dark excitons in $MoSe_2$ and therefore the only way to create a

valley polarization is the scattering channel via excitons which are bound to defect states within the band gap (Figs. 2f to 2i). It is expected that this mechanism is highly suppressed as soon as the Fermi level enters the conduction band (see decrease in the amplitude of the red data points in Fig. 1e). This argument seems to be in conflict with the trend of the Kerr amplitude in Fig. 3b, which only shows a slight decrease starting at $V_{\mathrm{TG}} = 2.5$ V long after a measurable current appears at $V_{\mathrm{TG}} = 1$ V.

There are two important differences in the device fabrication of the MoSe$_2$ sample which explain this apparent conflict. On the one hand, the MoSe$_2$ flake is not fully encapsulated by hBN, but instead is in direct contact with the SiO$_2$ substrate (see Fig. S3). In this condition, variations in the conduction band minima (CBM) and valence band maxima (VBM) energies are especially high as illustrated in Fig. S8a.[32] The second difference is the kind of gating: Whereas the WSe$_2$ samples are gated via a backgate consisting of a 300 nm thick SiO$_2$ layer, the MoSe$_2$ sample is gated with a graphitic top gate using an h-BN flake of approximately 20 nm thickness as the gate dielectric. Hence, the capacitive coupling strength between gate and TMD is quite different between these devices:

$$\frac{\varepsilon_r(\mathrm{hBN})}{d(\mathrm{hBN})} \cdot \frac{d(\mathrm{SiO_2})}{\varepsilon_r(\mathrm{SiO_2})} \approx 14 \; , \qquad (8)$$

with ε_r being the relative permittivity of the dielectric and d being the thickness of the dielectric. Once the TMD is fully tuned into either conduction or valence band, the change in charge carrier density n with gate voltage V_G is derived as:[27]

$$n(V_{\mathrm{G}}) = \frac{\varepsilon_0 \varepsilon_r}{d \cdot e} \left(V_{\mathrm{G}} - V_{\mathrm{th}} \right) \; , \qquad (9)$$

with ε_0 being the vacuum permittivity, e the elementary charge, and V_{th} the threshold voltage. But this is only true because the so-called quantum capacitance can be neglected within the bands. Instead, in both the gap and tail region the quantum capacitance plays a significant role in the gate voltage dependent shift of the Fermi level.[28,29,33,34] For example, the increase ΔV_{G} in gate voltage necessary to tune the Fermi level from the valance band maximum to the conduction band minimum is given in a first order approximation by:

$$\Delta V_{\mathrm{G}} = \frac{E_{\mathrm{G}}}{e} + \frac{n_{\mathrm{D}} \cdot e \cdot d}{\varepsilon_0 \varepsilon_r} \; , \qquad (10)$$

with E_{G} being the band gap energy and n_{D} being the density of mid-gap states. Interestingly, only the second term of this formula depends on the capacitive coupling strength. Therefore, to overcome the band gap in case of no mid-gap states, the same amount of voltage has to be applied both to a back-gated or a top-gated device of the same material. But due to the much higher capacitive coupling strength of the top-gated device structure, this change in gate voltage represents a much larger part of the overall measured gate voltage range (a maximum gate voltage of 3.5 V in case of the MoSe$_2$ sample compared to 80 V in case of the WSe$_2$ samples).

Considering now the much larger fluctuations of the band gap energy in case of the MoSe$_2$ sample (Fig. S8a) because of its interaction with the rough SiO$_2$ substrate,[32] makes it clear that there is a much more pronounced overlap of the three different regimes (the gap region, the tail region, and the conduction band region) of this device for a given gate voltage compared to e.g. the device shown in Fig. 1. In this respect it is important to note, that the current shown in Fig. 3a is plotted on a logarithmic scale and is quite small at a gate voltage of $V_{\mathrm{TG}} = 1$ V. This small, exponentially increasing current can best be explained by percolating transport. Fig. S8b represent the charge puddles shown in Fig. S8a in two dimensions together with a percolating electron path. This picture explains how a small charge current can be measured, whereas at the same time the TRKR still shows signals, which can be attributed to a Fermi-level in or near the band-gap (white areas in Fig. S8a).

F.   Figs. 3b and 3c

The approximately 200 ps long Kerr rotation signal shown in Fig. 3 of the main manuscript for the monolayer MoSe$_2$ device was attributed to the valley lifetime of this sample. Inconsistent to this assignment is the fact that its lifetime does not decrease for higher charge carrier densities (i.e. higher positive gate voltages), which would be expected from both electron-phonon and spin-orbit scattering mechanisms.[35–40] At the same time, the 200 ps lie within the upper range of reported bright exciton lifetimes.[10] Therefore, we cannot exclude that this signal could also result from relatively long recombination times of bright excitons.

VI.   TIME-RESOLVED KERR ROTATION SETUP

Fig. S9 schematically depicts the setup used to conduct the TRKR experiments. Two wavelength tunable mode-locked Ti:sapphire lasers (SPECTRA PHYSICS Tsunami) operating at a 80 MHz repetition rate were used as pump and probe laser sources with typical pulse widths of about 5 ps. Both lasers can be synchronized via a lock-to-clock (LTCL) laser synchronization unit to set a variable time delay between pump and probe pulses ranging from 0 ns to 12.5 ns. The polarization of the pump beam can be switched between σ^+, σ^- and linear by applying a voltage to the liquid crystal variable retarder. The laser beams are focused onto the sample by a 15 mm focal length aspheric lens (0.66 NA). Typical spot diameters are 8 μm FWHM. In order to distinguish and filter the pump from the probe beam, the pump beam enters and leaves the lens in a distance of 5 mm with respect to the probe beam
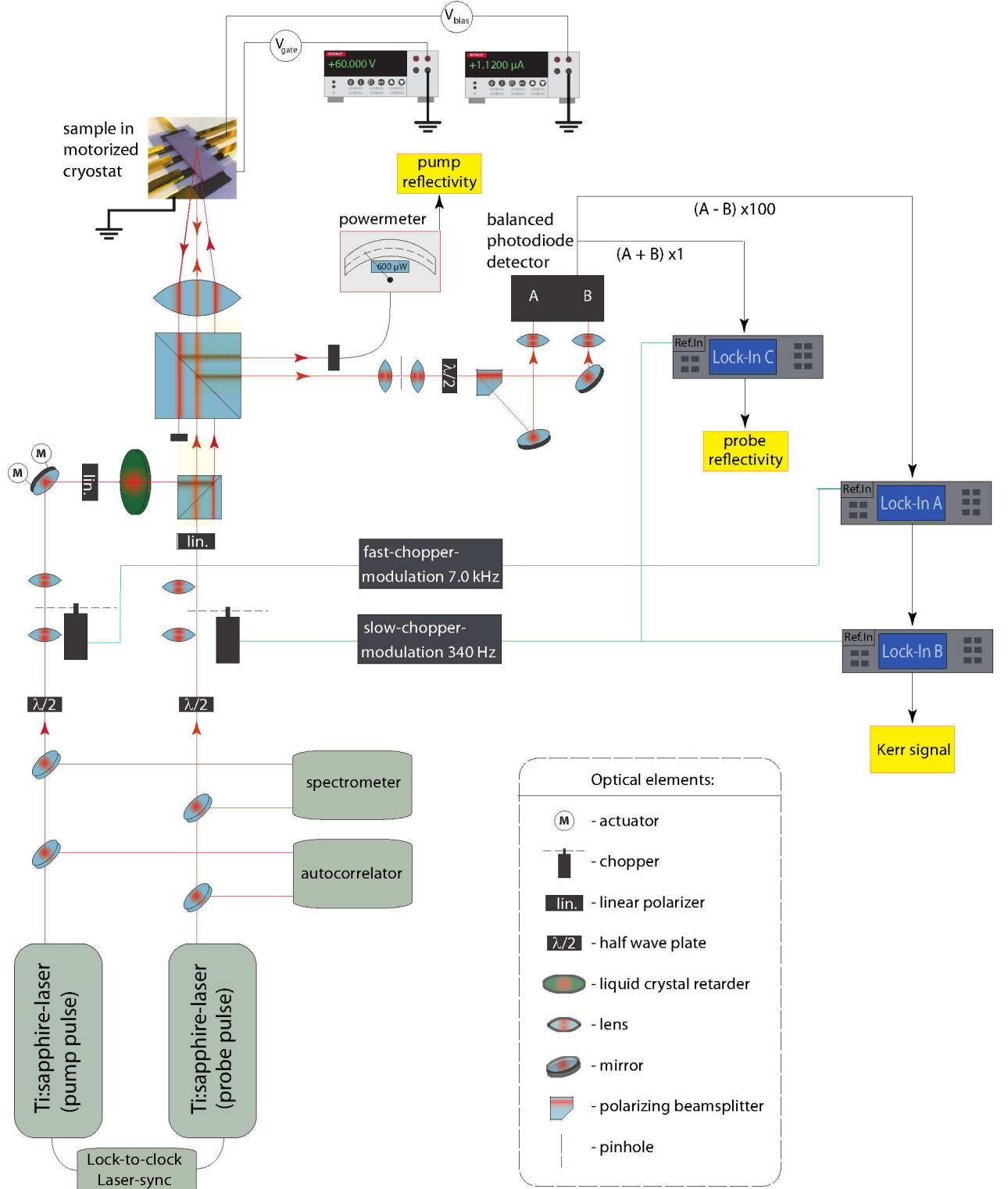
FIG. S9. Schematic of the experimental setup. Pump and probe laser pulses stem from two independently tunable Ti-sapphire ps lasers, which are synchronized and electronically delayed with a temporal jitter of less than 1 ps. The Kerr rotation signal is detected by a polarization bridge with balanced photodiodes.[41]

position, the latter passing the lens in its center. Therefore, the probe beam hits the sample perpendicularly and gets reflected into itself before being guided towards the detection arm. Here, a pin-hole of 150 μm diameter is placed to block any off-axis beam reflections that might disturb the signal. After passing another half-wave plate serving to equilibrate the detection diodes when the pump is blocked, finally the probe beam gets split into its per-pendicularly polarized components which are detected by the diodes A and B, respectively.[41]

For measuring the Kerr rotation angle we modulate the pump and probe beams with optical choppers driven at a fast modulation frequency of 7 kHz (pump) and a slower frequency of 340 Hz (probe). The pre-amplified A-B diode signal can then be demodulated using two lock-in-amplifier stages yielding the TRKR signal.

[1] Ju, L.; Velasco, J.; Huang, E.; Kahn, S.; Nosiglia, C.; Tsai, H.-Z.; Yang, W.; Taniguchi, T.; Watanabe, K.; Zhang, Y.; Zhang, G.; Crommie, M.; Zettl, A.; Wang, F. Photoinduced doping in heterostructures of graphene and boron nitride. *Nat. Nanotechnol.* **2014**, *9*, 348–352.

[2] Neumann, C.; Rizzi, L.; Reichardt, S.; Terrés, B.; Khodkov, T.; Watanabe, K.; Taniguchi, T.; Beschoten, B.; Stampfer, C. Spatial Control of Laser-Induced Doping Profiles in Graphene on Hexagonal Boron Nitride. *ACS Appl. Mater. Interfaces* **2016**, *8*, 9377–9383.

[3] Luo, X.; Andrews, K.; Wang, T.; Bowman, A.; Zhou, Z.; Xu, Y.-Q. Reversible photo-induced doping in WSe$_2$ field effect transistors. *Nanoscale* **2019**, *11*, 7358–7363.

[4] Furchi, M. M.; Polyushkin, D. K.; Pospischil, A.; Mueller, T. Mechanisms of Photoconductivity in Atomically Thin MoS$_2$. *Nano Lett.* **2014**, *14*, 6165–6170.

[5] Amit, I.; Octon, T. J.; Townsend, N. J.; Reale, F.; Wright, C. D.; Mattevi, C.; Craciun, M. F.; Russo, S. Role of Charge Traps in the Performance of Atomically Thin Transistors. *Advanced Materials* **2017**, *29*, 1605598.

[6] Yin, L.; He, P.; Cheng, R.; Wang, F.; Wang, F.; Wang, Z.; Wen, Y.; He, J. Robust trap effect in transition metal dichalcogenides for advanced multifunctional devices. *Nat. Commun.* **2019**, *10*, 1–8.

[7] Venanzi, T.; Arora, H.; Erbe, A.; Pashkin, A.; Winnerl, S.; Helm, M.; Schneider, H. Exciton localization in MoSe$_2$ monolayers induced by adsorbed gas molecules. *Appl. Phys. Lett.* **2019**, *114*, 172106.

[8] Robert, C.; Amand, T.; Cadiz, F.; Lagarde, D.; Courtade, E.; Manca, M.; Taniguchi, T.; Watanabe, K.; Urbaszek, B.; Marie, X. Fine structure and lifetime of dark excitons in transition metal dichalcogenide monolayers. *Phys. Rev. B* **2017**, *96*, 155423.

[9] Liu, E.; van Baren, J.; Lu, Z.; Altaiary, M. M.; Taniguchi, T.; Watanabe, K.; Smirnov, D.; Lui, C. H. Gate Tunable Dark Trions in Monolayer WSe$_2$. *Phys. Rev. Lett.* **2019**, *123*, 027401.

[10] Wang, G.; Chernikov, A.; Glazov, M. M.; Heinz, T. F.; Marie, X.; Amand, T.; Urbaszek, B. Colloquium: Excitons in atomically thin transition metal dichalcogenides. *Rev. Mod. Phys.* **2018**, *90*, 021001.

[11] Fang, H. H.; Han, B.; Robert, C.; Semina, M. A.; Lagarde, D.; Courtade, E.; Taniguchi, T.; Watanabe, K.; Amand, T.; Urbaszek, B.; Glazov, M. M.; Marie, X. Control of the Exciton Radiative Lifetime in van der Waals Heterostructures. *Phys. Rev. Lett.* **2019**, *123*, 067401.

[12] Robert, C.; Lagarde, D.; Cadiz, F.; Wang, G.; Lassagne, B.; Amand, T.; Balocchi, A.; Renucci, P.; Tongay, S.; Urbaszek, B.; Marie, X. Exciton radiative lifetime in transition metal dichalcogenide monolayers. *Phys. Rev. B* **2016**, *93*, 205423.

[13] Nie, Z. et al. Tailoring exciton dynamics of monolayer transition metal dichalcogenides by interfacial electron-phonon coupling. *Commun. Phys.* **2019**, *2*, 1–8.

[14] Volmer, F.; Pissinger, S.; Ersfeld, M.; Kuhlen, S.; Stampfer, C.; Beschoten, B. Intervalley dark trion states with spin lifetimes of 150 ns in WSe$_2$. *Phys. Rev. B* **2017**, *95*, 235408.

[15] 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* **2015**, *9*, 466–470.

[16] Pogna, E. A. A.; Marsili, M.; De Fazio, D.; Dal Conte, S.; Manzoni, C.; Sangalli, D.; Yoon, D.; Lombardo, A.; Ferrari, A. C.; Marini, A.; Cerullo, G.; Prezzi, D. Photo-Induced Bandgap Renormalization Governs the Ultrafast Response of Single-Layer MoS$_2$. *ACS Nano* **2016**, *10*, 1182–1188.

[17] Zhou, Y.; Scuri, G.; Wild, D. S.; High, A. A.; Dibos, A.; Jauregui, L. A.; Shu, C.; De Greve, K.; Pistunova, K.; Joe, A. Y.; Taniguchi, T.; Watanabe, K.; Kim, P.; Lukin, M. D.; Park, H. Probing dark excitons in atomically thin semiconductors via near-field coupling to surface plasmon polaritons. *Nat. Nanotechnol.* **2017**, *12*, 856.

[18] Borghardt, S.; Tu, J.-S.; Taniguchi, T.; Watanabe, K.; Kardynal, B. E. The Effect of Exciton Localization on Exciton Dynamics in *p*-Doped WSe$_2$ Monolayers. *arXiv* **2019**, 1911.01092.

[19] Goryca, M.; Wilson, N. P.; Dey, P.; Xu, X.; Crooker, S. A. Detection of thermodynamic "valley noise" in monolayer semiconductors: Access to intrinsic valley relaxation time scales. *Sci. Adv.* **2019**, *5*, eaau4899.

[20] Allain, A.; Kang, J.; Banerjee, K.; Kis, A. Electrical contacts to two-dimensional semiconductors. *Nat. Mater.* **2015**, *14*, 1195–1205.

[21] Liu, Y.; Guo, J.; Zhu, E.; Liao, L.; Lee, S.-J.; Ding, M.; Shakir, I.; Gambin, V.; Huang, Y.; Duan, X. Approaching the Schottky–Mott limit in van der Waals metal–semiconductor junctions. *Nature* **2018**, *557*, 696–700.

[22] Kim, C.; Moon, I.; Lee, D.; Choi, M. S.; Ahmed, F.; Nam, S.; Cho, Y.; Shin, H.-J.; Park, S.; Yoo, W. J. Fermi Level Pinning at Electrical Metal Contacts of Monolayer Molybdenum Dichalcogenides. *ACS Nano* **2017**, *11*, 1588–1596.

[23] Efimkin, D. K.; MacDonald, A. H. Many-body theory of trion absorption features in two-dimensional semiconductors. *Phys. Rev. B* **2017**, *95*, 035417.

[24] Sidler, M.; Back, P.; Cotlet, O.; Srivastava, A.; Fink, T.; Kroner, M.; Demler, E.; Imamoglu, A. Fermi polaron-polaritons in charge-tunable atomically thin semiconductors. *Nature Physics* **2016**, *13*, 255.

[25] Wang, Z.; Mak, K. F.; Shan, J. Strongly Interaction-Enhanced Valley Magnetic Response in Monolayer WSe_2. *Phys. Rev. Lett.* **2018**, *120*, 066402.

[26] Das Sarma, S.; Adam, S.; Hwang, E. H.; Rossi, E. Electronic transport in two-dimensional graphene. *Rev. Mod. Phys.* **2011**, *83*, 407–470.

[27] Zhang, Y.; Ye, J.; Matsuhashi, Y.; Iwasa, Y. Ambipolar MoS2 Thin Flake Transistors. *Nano Lett.* **2012**, *12*, 1136–1140.

[28] Ma, N.; Jena, D. Carrier statistics and quantum capacitance effects on mobility extraction in two-dimensional crystal semiconductor field-effect transistors. *2D Mater.* **2015**, *2*, 015003.

[29] Fang, N.; Nagashio, K. Band tail interface states and quantum capacitance in a monolayer molybdenum disulfide field-effect-transistor. *J. Phys. D: Appl. Phys.* **2018**, *51*, 065110.

[30] Kormányos, A.; Burkard, G.; Gmitra, M.; Fabian, J.; Zólyomi, V.; Drummond, N. D.; Fal'ko, V. k·ptheory for two-dimensional transition metal dichalcogenide semiconductors. *2D Materials* **2015**, *2*, 022001.

[31] Mak, K. F.; Shan, J. Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides. *Nature Photonic* **2016**, *10*, 216–226.

[32] Shin, B. G.; Han, G. H.; Yun, S. J.; Oh, H. M.; Bae, J. J.; Song, Y. J.; Park, C.-Y.; Lee, Y. H. Indirect Bandgap Puddles in Monolayer MoS_2 by Substrate-Induced Local Strain. *Advanced Materials* **2016**, *28*, 9378–9384.

[33] Luryi, S. Quantum capacitance devices. *Appl. Phys. Lett.* **1988**, *52*, 501–503.

[34] Lu, C.-P.; Li, G.; Mao, J.; Wang, L.-M.; Andrei, E. Y. Bandgap, Mid-Gap States, and Gating Effects in MoS_2. *Nano Lett.* **2014**, *14*, 4628–4633.

[35] Ochoa, H.; Roldán, R. Spin-orbit-mediated spin relaxation in monolayer MoS_2. *Phys. Rev. B* **2013**, *87*, 245421.

[36] Ochoa, H.; Finocchiaro, F.; Guinea, F.; Fal'ko, V. I. Spin-valley relaxation and quantum transport regimes in two-dimensional transition-metal dichalcogenides. *Phys. Rev. B* **2014**, *90*, 235429.

[37] Habe, T.; Koshino, M. Spin relaxation in a hole-doped transition-metal dichalcogenide monolayer and bilayer with crystal defects. *Phys. Rev. B* **2016**, *93*, 075415.

[38] Ge, Y.; Wan, W.; Feng, W.; Xiao, D.; Yao, Y. Effect of doping and strain modulations on electron transport in monolayer MoS_2. *Phys. Rev. B* **2014**, *90*, 035414.

[39] Gunst, T.; Markussen, T.; Stokbro, K.; Brandbyge, M. First-principles method for electron-phonon coupling and electron mobility: Applications to two-dimensional materials. *Phys. Rev. B* **2016**, *93*, 035414.

[40] Molina-Sánchez, A.; Sangalli, D.; Wirtz, L.; Marini, A. Ab Initio Calculations of Ultrashort Carrier Dynamics in Two-Dimensional Materials: Valley Depolarization in Single-Layer WSe_2. *Nano Lett.* **2017**, *17*, 4549.

[41] Crooker, S. A.; Awschalom, D. D.; Baumberg, J. J.; Flack, F.; Samarth, N. Optical spin resonance and transverse spin relaxation in magnetic semiconductor quantum wells. *Phys. Rev. B* **1997**, *56*, 7574–7588.