

Approaching the intrinsic exciton physics limit in two-dimensional semiconductor diodes

<https://doi.org/10.1038/s41586-021-03949-7>

Received: 3 February 2021

Accepted: 24 August 2021

Published online: 17 November 2021

 Check for updates

Peng Chen^{1,6}, Timothy L. Atallah^{1,6}, Zhaoyang Lin¹, Peiqi Wang¹, Sung-Joon Lee², Junqing Xu³, Zhihong Huang², Xidong Duan⁴, Yuan Ping³, Yu Huang^{2,5}, Justin R. Caram^{1,5}✉ & Xiangfeng Duan^{1,5}✉

Two-dimensional (2D) semiconductors have attracted intense interest for their unique photophysical properties, including large exciton binding energies and strong gate tunability, which arise from their reduced dimensionality^{1–5}. Despite considerable efforts, a disconnect persists between the fundamental photophysics in pristine 2D semiconductors and the practical device performances, which are often plagued by many extrinsic factors, including chemical disorder at the semiconductor–contact interface. Here, by using van der Waals contacts with minimal interfacial disorder, we suppress contact-induced Shockley–Read–Hall recombination and realize nearly intrinsic photophysics-dictated device performance in 2D semiconductor diodes. Using an electrostatic field in a split-gate geometry to independently modulate electron and hole doping in tungsten diselenide diodes, we discover an unusual peak in the short-circuit photocurrent at low charge densities. Time-resolved photoluminescence reveals a substantial decrease of the exciton lifetime from around 800 picoseconds in the charge-neutral regime to around 50 picoseconds at high doping densities owing to increased exciton–charge Auger recombination. Taken together, we show that an exciton–diffusion-limited model well explains the charge-density-dependent short-circuit photocurrent, a result further confirmed by scanning photocurrent microscopy. We thus demonstrate the fundamental role of exciton diffusion and two-body exciton–charge Auger recombination in 2D devices and highlight that the intrinsic photophysics of 2D semiconductors can be used to create more efficient optoelectronic devices.

Nano-semiconductors such as two-dimensional (2D) materials and colloidal quantum dots^{6–16} feature weak dielectric screening and a strong confinement effect that leads to unusually large exciton (electron–hole quasiparticle) binding energies ($E_B > 100$ meV)^{1–5,17–19}. Thus, excitons often govern the nanoscale optoelectronic device performance even at room temperature^{18,20}. This is in contrast to conventional semiconductors (for example, gallium arsenide and silicon), where free carriers typically dictate the device properties at room temperature owing to much smaller exciton binding energies ($E_B < k_B T \approx 26$ meV, where k_B is the Boltzmann constant and T is temperature). The dominant role of excitons could lead to unique nanoscale device characteristics. For example, electron–electron/hole and hole–electron/hole Auger recombination is now a two-body (charge–exciton) process that linearly depends on the electron or hole density. Furthermore, the intrinsic lower dimensionality naturally leads to an enhanced many-body scattering probability of diffusive species^{21,22}, resulting in shorter diffusion

lengths and lower device efficiencies. However, despite the prevalence of excitons in nano-semiconductors, directly relating optoelectronic device properties to excitonic behaviour remains challenging owing to severe extrinsic losses from the non-idealities at contact or dielectric interfaces.

Here we show nearly intrinsic excitonic optoelectronic device performance in van der Waals (vdW)-contacted 2D diodes with nearly ideal metal–semiconductor interfaces. By combining optoelectronic studies with time-resolved photoluminescence (TRPL) measurements to probe the charge-density-dependent photovoltaic performance, we reveal the fundamental role of exciton diffusion and two-body exciton–charge Auger recombination in determining the 2D diode performance, and demonstrate a greatly boosted photocurrent approaching the intrinsic excitonic photophysics limit.

The 2D tungsten diselenide (WSe_2) p–n diodes were constructed with a split-gate (about 250 nm apart) structure (Fig. 1a, b), with multilayer

¹Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, USA. ²Department of Materials Science and Engineering, University of California, Los Angeles, Los Angeles, CA, USA. ³Department of Chemistry and Biochemistry, University of California Santa Cruz, Santa Cruz, CA, USA. ⁴State Key Laboratory for Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha, China. ⁵California NanoSystems Institute, University of California, Los Angeles, Los Angeles, CA, USA. ⁶These authors contributed equally: Peng Chen, Timothy L. Atallah. ✉e-mail: xduan@chem.ucla.edu; jcaram@chem.ucla.edu

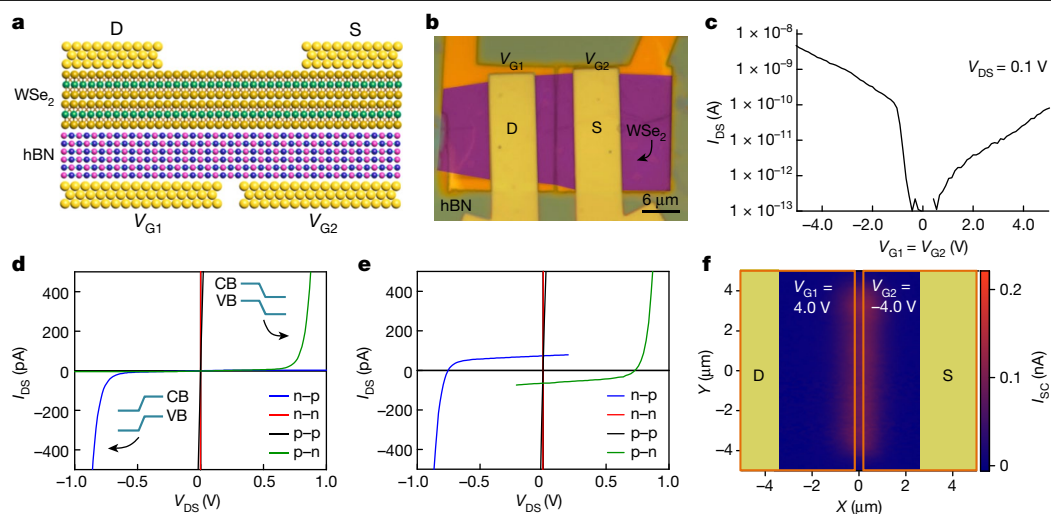


Fig. 1 | Atomically thin WSe₂ p-n diode with atomically clean vdW contacts.

a, A schematic of the device structure: V_{G1} and V_{G2} are the local left and right back-gates 250 nm apart. An exfoliated multilayer hBN flake is used as the gate dielectric and few-layer (one to four layers) WSe₂ is used as the semiconductor channel. Two Au electrodes are transferred onto the WSe₂ as source (S) and drain (D) electrodes. The green, dark yellow, blue, purple and light yellow spheres denote the tungsten, selenium, boron, nitrogen, and gold atoms, respectively. **b**, An optical microscope image of the 2D WSe₂ p-n diode. The trilayer WSe₂ area is artificially coloured. **c**, The transfer curve of the trilayer WSe₂ device measured at $V_{DS} = 0.1$ V. The thickness of hBN is 22 nm in

this device. **d**, The I_{DS} - V_{DS} curves for different device configurations under dark. The insets show the band diagram of the p-n and n-p configurations. CB, conduction band; VB, valence band. **e**, The I_{DS} - V_{DS} curves for different device configurations under laser illumination. The laser excitation wavelength $\lambda = 473$ nm and the power density $P_d = 24$ mW cm⁻². **f**, Spatially resolved photocurrent mapping of a 2D diode using a focused laser (405 nm, roughly 1 μm spot size and about 2.2 nW total power), highlighting photocurrent generation in the split-gate region. The $X = 0$ μm denotes the centre position of the split-gate region, the yellow squares denote the position of the electrodes and the orange line denotes the position of the split gates V_{G1} and V_{G2} .

hexagonal boron nitride (hBN) as a pin-free gate dielectric^{23,24} and a few-layer (one to four layers) WSe₂ flake as the semiconductor channel, which was contacted with a pair of transferred gold (Au) electrodes as the vdW contacts to avoid the metal-deposition-induced interfacial disorder²⁵.

The split gates (V_{G1} , the gate voltage of the left split gate; V_{G2} , the gate voltage of the right split gate) allow independent electrostatic control of the carrier density in the left and right sides of the WSe₂ channel to access various electronic configurations. When $V_{G1} = V_{G2}$, the transfer curve shows a charge-neutral region near $V_{G1} = V_{G2} \approx 0$ V for a trilayer WSe₂ device (Fig. 1c). The output characteristics (source-drain current, I_{DS} , versus source-drain voltage, V_{DS}) in the dark show distinct behaviour under four different gate-bias conditions (Fig. 1d). With $V_{G1} = -5$ V and $V_{G2} = 5$ V, the left side is p-doped and the right side is n-doped, and we obtain a p-n diode with clear current rectification in which the device shows a high current at positive V_{DS} and a negligible current at negative V_{DS} . Conversely, when $V_{G1} = 5$ V and $V_{G2} = -5$ V, the diode is reversed to the n-p configuration with opposite rectification. When the device operates with $V_{G1} = V_{G2} = 5$ V (n-n) and $V_{G1} = V_{G2} = -5$ V (p-p) configurations, there is no rectification and $|I_{DS}|$ is far greater than that in the diode configurations.

Under illumination, both the p-n and n-p configurations show a clear photovoltaic response with a short-circuit photocurrent (I_{SC}) and open-circuit voltage (V_{OC}) (Fig. 1e). Scanning photocurrent mapping reveals that the maximum photocurrent arises at the p-n junction area near the centre of the split-gate region, demonstrating that the photocurrent originated from the p-n interface rather than the Au-WSe₂ contacts (Fig. 1f). Additional analysis of the band diagram, contact resistance and series resistance further supports that the p-n junction dominates the photocurrent-generation process (Methods, Extended Data Figs. 1, 2).

We next studied the doping-dependent optoelectronic properties. First, fixing $V_{G1} = -5$ V, we p-doped the left side of the WSe₂ junction ($n \approx 4.4 \times 10^{12}$ cm⁻², n is the charge doping density) and investigated the photovoltaic response with varying doping on the right side. By gradually increasing the V_{G2} from -5 V (hole doped) to $+5$ V (electron doped), we observed a systematic evolution of the photovoltaic response (Fig. 2a). Specifically, at $V_{G2} = -5$ V, the device behaves as a p-p junction with essentially zero I_{SC} and zero V_{OC} . As the device transforms from the p-p to the

p-n configuration, we observe a noticeable photovoltaic response with the V_{OC} increasing rapidly and then saturating at around 0.75 V (black dots in Fig. 2c). Interestingly, the I_{SC} shows an unusual yet highly reproducible doping dependence (black dots in Fig. 2e): first increasing sharply to reach a peak value of around 260 pA as the diode moves from the p-p to the p-n configuration and then decreasing to 65 pA at higher electron densities. Similarly, by fixing $V_{G1} = 5$ V (left side n-doped), we observed a similar trend (Fig. 2b, d, f) when the doping of the right side is gradually changed from electrons to holes, with a rapid increase of I_{SC} to a peak value about 300 pA as the junction is transformed from n-n to n-p followed by a decrease to 73 pA at higher hole densities (black dots in Fig. 2f).

As the electrostatic doping does not introduce impurity scattering in the 2D semiconductors, the deterioration of I_{SC} with increasing charge density suggests a reduced carrier lifetime with increasing doping density, a trend typically attributed to Auger recombination. Our device shows an extraordinarily large charge-induced photocurrent deterioration of more than 75%, suggesting an unusual Auger effect in 2D semiconductors. Specifically, exciton-charge Auger processes instead of free-carrier Auger processes probably govern the interactions as the excitons dominate 2D semiconductors owing to the large exciton binding energy²⁶. A complete mapping of the V_{OC} and I_{SC} dependence on both electron and hole densities further confirms the distinct doping-dependent photovoltaic response (Fig. 2g, h), revealing a highly symmetric behaviour for both electron and hole doping regimes for either side of the junction, as expected for the exciton-charge Auger process.

Parallel studies on the trilayer WSe₂ p-n diodes with a similar device geometry but evaporated contacts (evap-diodes) show notably different behaviour when compared with the diodes with vdW contacts (vdW-diodes). In the case of the evap-diodes, the I_{SC} rapidly increases when the device is transformed from a p-p to a p-n junction but quickly saturates to a maximum value of 53 pA without a peak value at low doping concentration (red dot in Fig. 2e). Notably, the apparent maximum external quantum efficiency (EQE) observed in the vdW-diodes is about five- to six-times higher than that of the evap-diodes (Extended Data Fig. 3). The maximum V_{OC} observed in the vdW-diodes is about 150 mV higher than that of the evap-diodes (Fig. 2c, d). The inferior

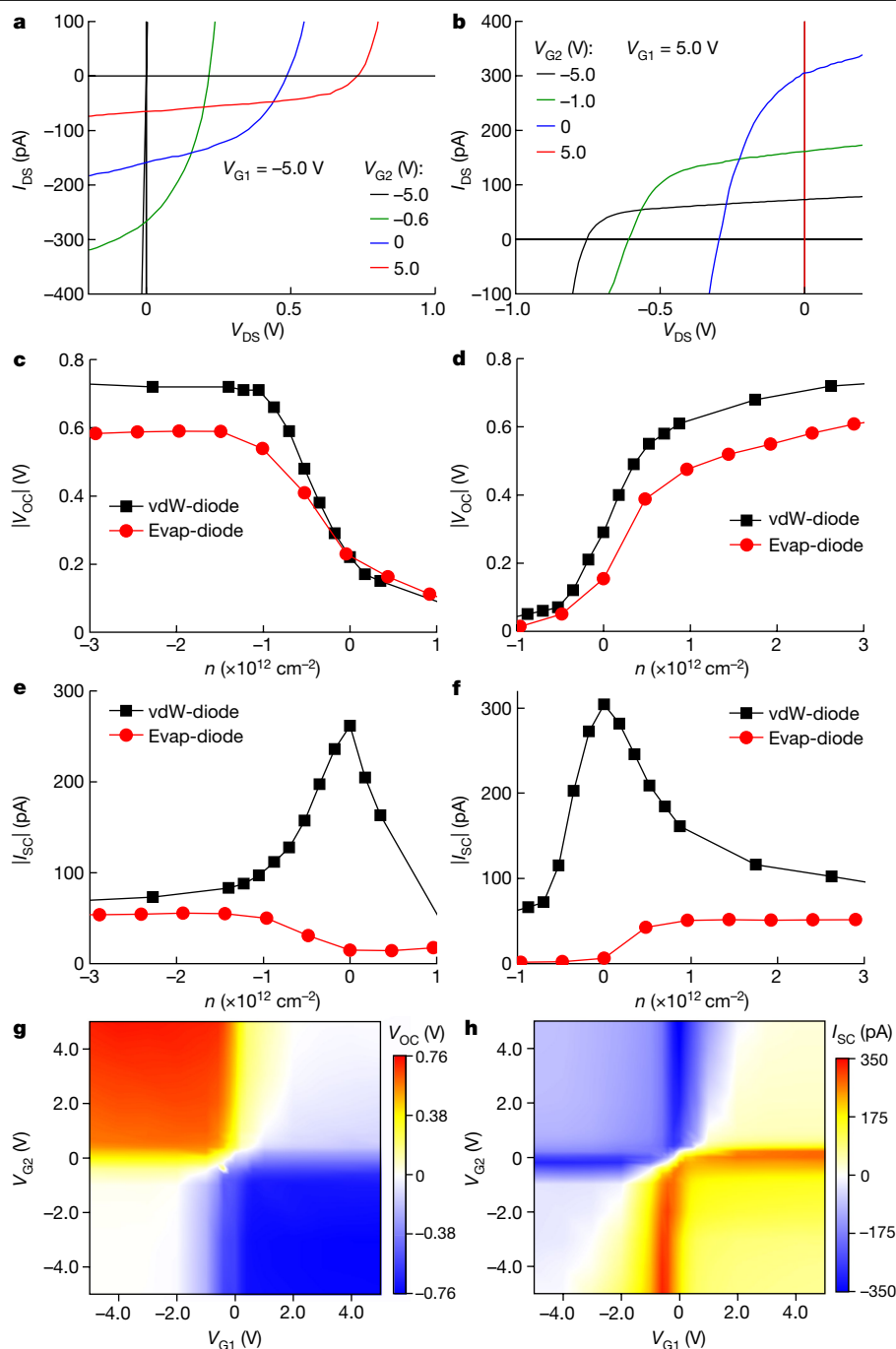


Fig. 2 | Doping-dependent optoelectronic performance of a 2D WSe₂ p-n diode. **a**, The $I_{\text{DS}}-V_{\text{DS}}$ curves measured at $V_{\text{G1}} = -5$ V and different V_{G2} under laser illumination. **b**, The $I_{\text{DS}}-V_{\text{DS}}$ curves measured at $V_{\text{G1}} = 5$ V and different V_{G2} under laser illumination. **c**, The $|V_{\text{OC}}|$ dependence on charge density, n , for diode with evaporated contacts (evap-diode, red dots) or vdW contacts (vdW-diode, black dots) (lines are a guide for the eyes) at $V_{\text{G1}} = -5$ V. **d**, The $|V_{\text{OC}}|$ dependence on charge density, n , for evap-diode (red dots) and vdW-diode (black dots) at

$V_{\text{G1}} = 5$ V. A positive carrier density indicates hole density and a negative value indicates electron density. **e**, The $|I_{\text{SC}}|$ dependence on charge density, n , for the evap-diode (red dots) and vdW-diode (black dots) at $V_{\text{G1}} = -5$ V. **f**, The $|I_{\text{SC}}|$ dependence on charge density, n , for the evap-diode (red dots) and with vdW-diode (black dots) at $V_{\text{G1}} = 5$ V. **g**, The V_{G1} and V_{G2} dependence of the V_{OC} of the vdW-diode. **h**, The V_{G1} and V_{G2} dependence of the I_{SC} of the vdW-diode. A broad-area laser excitation ($\lambda = 473$ nm, $P_d = 24$ mW cm⁻²) was used.

optoelectronic performance of the evap-diodes suggests severe carrier loss mechanisms, which may be attributed to the evaporation-induced defects at the Au–WSe₂ interface^{25,27} that act as the Shockley–Read–Hall (SRH) recombination centres to compromise charge-collection efficiency^{28,29}. In contrast, the vdW-diodes feature minimal interfacial disorder or trap states²⁵, thus offering an attractive platform for probing the intrinsic diode behaviour.

High exciton binding energies ($E_b \approx 350$ meV in trilayer WSe₂)³⁰ suggest that exciton photophysics govern photoexcited carriers at room

temperature¹⁸. We therefore hypothesized that charge Auger losses are dictated by exciton–charge Auger recombination in ideal 2D diodes. Exciton–charge Auger occurs when an exciton non-radiatively relaxes to promote a carrier to a higher excited state that rapidly cools, unable to contribute to photocurrent generation. This loss mechanism is more probable at higher charge densities. Consequently, fewer excitons can diffuse to the p–n junction interface to contribute to I_{SC} , consistent with what we observed in the vdW-diodes. By comparison, evap-diodes remain unaffected by charge modulation as their photocurrent is

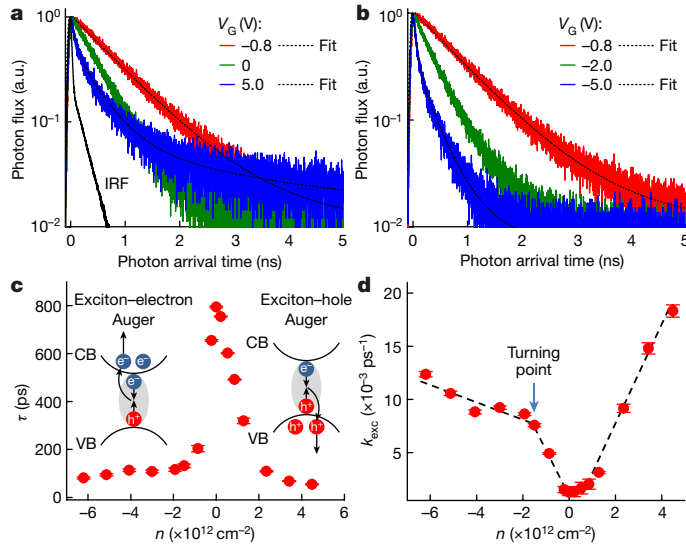


Fig. 3 | Doping-dependent TRPL and exciton-charge Auger. **a**, Normalized TRPL traces (solid lines) showing decreasing excited-state lifetimes with increasing electron doping in trilayer WSe₂, and the corresponding bi(tri)exponential fits (dashed lines) with instrument response function (IRF, black). **b**, Normalized TRPL traces (solid lines) showing decreasing excited-state lifetimes with increasing hole doping and the corresponding bi(tri)exponential fits (dashed lines). **c**, The doping dependence of the exciton lifetimes, τ . The inset shows the non-radiative decay of excitons through an exciton–electron Auger process and an exciton–hole Auger process, respectively (negative density refers to electrons and positive refers to holes). CB, conduction band; VB, valence band. **d**, Exciton recombination rates, k_{exc} , as a function of doping density. The corresponding linear fits give the exciton-charge Auger coefficient; at the electron side, the Auger coefficient shows a turning point (marked by the blue arrow) at high electron density beyond approximately $1.5 \times 10^{12} \text{ cm}^{-2}$. Error bars are given by the fit.

largely limited by interfacial-defect-induced SRH recombination at metal–semiconductor contacts.

We next performed TRPL measurements on the trilayer WSe₂ at variable charge densities (see Methods for details; Fig. 3a, b). The recorded lifetime traces were fit to a convolution of an instrument response function and bi- or tri-exponential decays (dashed lines in Fig. 3a, b; Methods, Extended Data Fig. 4a, b, Extended Data Table 1). The dominant amplitude component is attributed to the exciton recombination³¹ and two other weaker components to the localized states (Methods, Extended Data Fig. 4c, d). The dominant component shows an exciton lifetime (τ) of 800 ps at the charge-neutrality point, which rapidly decreases to tens of picoseconds at high doping density ($>3 \times 10^{12} \text{ cm}^{-2}$) (Fig. 3c), supporting the above hypothesis that the exciton–charge Auger effect has a dictating role in 2D semiconductors.

We further investigated the doping-dependent exciton recombination rate ($k_{\text{exc}} = 1/\tau$). Overall, k_{exc} shows a linear dependence with respect to hole density (n) (Fig. 3d). The charge-density dependence of the excited-state decay rates in bulk semiconductors are typically described by the canonical ‘ABC recombination model’³², which assumes that all interactions are between independent free-charge carriers and predicts Auger processes to be a three-body phenomenon with a quadratic dependence on charge density (n^2). The linear dependence observed in our study implies a two-body interaction process. To understand this, we invoke exciton–charge interactions²⁶, describing exciton recombination as:

$$k_{\text{exc}} = k_{0,\text{exc}} + |n|B_{\text{Aug}} \quad (1)$$

where $k_{0,\text{exc}}$ is the intrinsic excitonic recombination rate, n is the charge density and B_{Aug} corresponds to the Auger decay constant for exciton–charge recombination. With this formula, we obtain an exciton–hole Auger coefficient $B_{\text{Aug,h}}$ of $4.7 \times 10^{-3} \pm 0.3 \times 10^{-3} \text{ cm}^2 \text{ s}^{-1}$, comparable

to the (two-body) biexciton annihilation rates for the trilayer WS₂ of $1.88 \times 10^{-3} \text{ cm}^2 \text{ s}^{-1}$ reported previously⁴.

The exciton recombination rate at the electron side shows a more complicated behaviour, with a similar exciton–electron Auger coefficient of $5.0 \times 10^{-3} \pm 0.3 \times 10^{-3} \text{ cm}^2 \text{ s}^{-1}$ at low electron density, which decreases to $0.8 \times 10^{-3} \pm 0.2 \times 10^{-3} \text{ cm}^2 \text{ s}^{-1}$ as the doping density exceeds $1.5 \times 10^{12} \text{ cm}^{-2}$ (Fig. 3d). This doping-dependent Auger recombination may be attributed to the unique band structure of the trilayer WSe₂. At low electron doping, the exciton–electron Auger recombination mainly involves electrons in the Q valley; whereas at higher doping, the process may start to involve electrons in the K valley (Methods, Extended Data Fig. 5). We hypothesize that the Auger recombination involving K-valley electrons is slower.

The excitons in 2D diodes only contribute to the photocurrent generation if they migrate to the p–n junction interface and undergo a charge separation. In this picture, only excitons within the collection length ($L_{\text{coll}} = L_{\text{exc}} + L_{\text{sg}} = L_{\text{p}} + L_{\text{n}} + L_{\text{sg}}$) contribute to I_{SC} (Fig. 4a, b), where L_{sg} is the length of split-gate region, and L_{exc} , L_{p} and L_{n} are the total exciton diffusion length and that in the p-region and n-region, respectively. The L_{exc} can be controlled by the charge doping density through the exciton–charge Auger recombination. At higher doping density, the exciton lifetime is shorter, leading to a smaller L_{exc} , and thus a smaller L_{coll} and a lower I_{SC} (Fig. 4a); conversely, a lower doping density leads to a higher L_{coll} and hence a higher I_{SC} (Fig. 4b).

To directly probe the doping-dependent L_{exc} , we conducted scanning photocurrent microscopy studies at different gate voltages. As expected, when both sides of the p–n junction are highly doped, we only observe photocurrent near the split-gate region, with little contribution from exciton diffusion from outside of the split-gate region (Fig. 1f). However, when the p-side is lightly doped to have a p⁺–n junction, the photocurrent generation region clearly extends towards the p⁺ region, suggesting a much longer exciton diffusion length in the low-doping region (Fig. 4c). The photocurrent line scans at different gate voltages more clearly show the gradual expansion of the photoactive area towards the p⁺ side with decreasing hole doping, suggesting an increased exciton diffusion length on the right side (Fig. 4d). By contrast, the spatial distribution on the left side remains essentially the same because of the fixed doping concentration (V_{G1} fixed at +4 V), which means that the exciton diffusion length is unchanged on the left side. Deconvolving the instrument response³³, we can estimate an exciton diffusion length of $720 \pm 100 \text{ nm}$ in the low-doping limit (Extended Data Fig. 6).

Our scanning photocurrent microscopy studies clearly indicate that both the photocurrent amplitude and the exciton diffusion length are highly dependent on the charge doping. To further correlate the device behaviour with the photophysics parameters determined from the TRPL measurements, we employed a uniform-generation, exciton-diffusion-limited model³⁴ to describe the I_{SC} for a p–n junction:

$$I_{\text{SC}} = I_{\text{diff}} + I_{\text{sg}} - I_{\text{los}} \quad (2)$$

where I_{diff} is the exciton diffusion limited photocurrent, I_{sg} accounts for the photocurrent generated by the excitations in the split-gate region and I_{los} accounts for the photocurrent losses such as SRH recombination. As SRH recombination is weakly sensitive to the charge doping^{35–38}, we assume it to be a constant value across different charge densities. The diffusion-limited photocurrent may be expressed as:

$$I_{\text{diff}} = eSG_0L_{\text{exc}} = eSG_0(D_{\text{exc}}\tau)^{1/2} \quad (3)$$

where e is the elementary charge, S is the p–n junction interface cross-section, G_0 is the exciton density generation rate accounting for the fraction of photon absorption (f_{abs}) at photon energy E_{ph} and is determined by $\frac{P_{\text{in}}}{E_{\text{ph}}} f_{\text{abs}}$ ($f_{\text{abs}} \approx 0.25$ for trilayer WSe₂, P_{in} is the input power) (see Methods for details), L_{exc} is the exciton diffusion length and D_{exc} is the exciton diffusion constant.

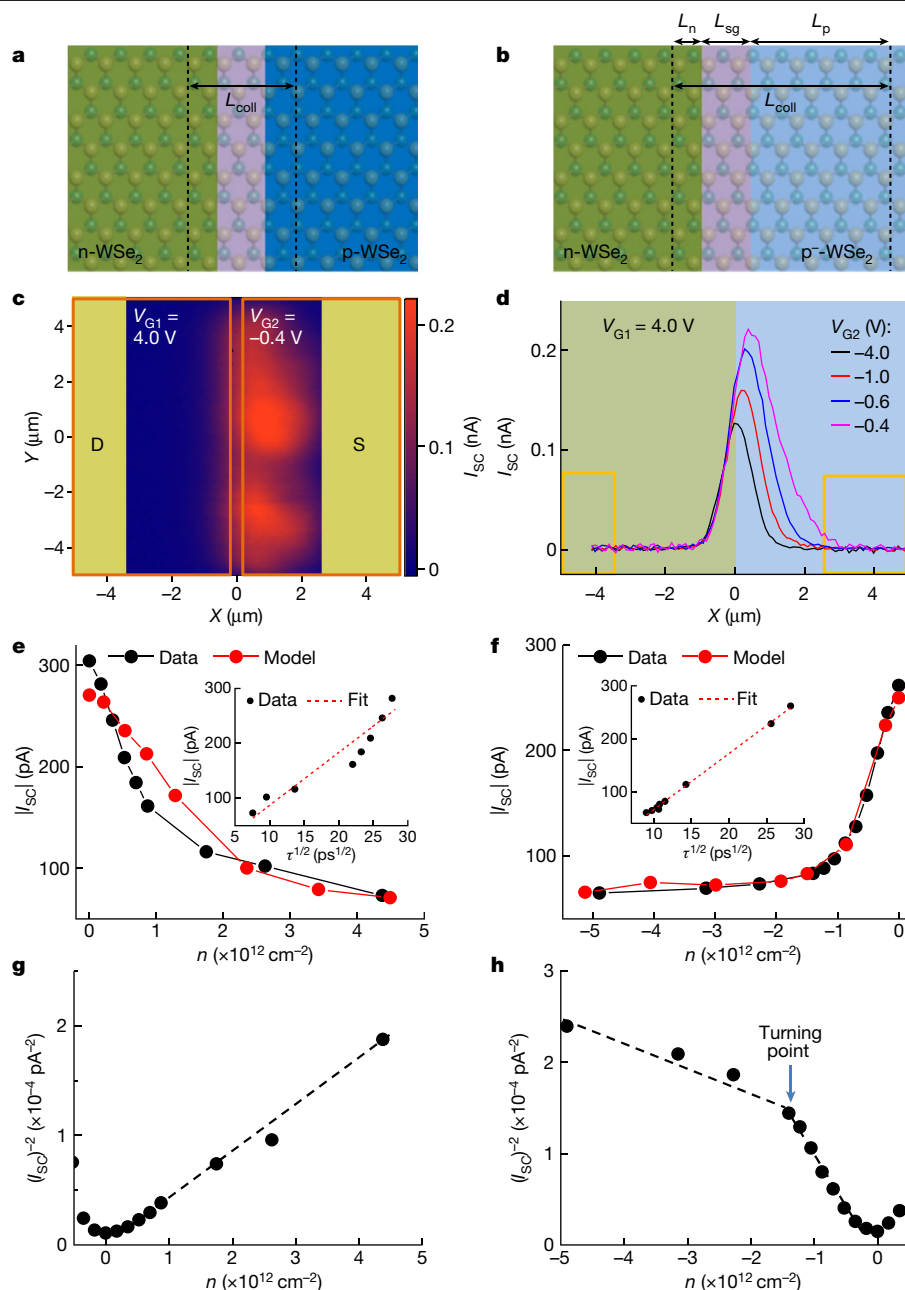


Fig. 4 | Correlation between photocurrent and exciton lifetime.

a, b, Schematic illustration of the change of the exciton collection length (L_{coll}) from high doping (**a**) to low doping (**b**). **c**, I_{sc} mapping from low doping in a four-layer WSe_2 diode with a focused laser (405 nm, roughly $1\ \mu\text{m}$ spot size and about 2.2 nW total power) excitation. The $X = 0\ \mu\text{m}$ denotes the centre position of the split-gate region; the yellow squares denote the position of the electrodes and the orange line denotes the position of the split gates V_{G1} and V_{G2} . We note that there is some spatial heterogeneity in the photocurrent mapping, which might be attributed to non-uniform doping⁴² from charge impurities, strain or bubble formation in the local environment. **d**, The I_{sc} line scan at $V_{\text{G1}} = 4\ \text{V}$ and different V_{G2} in a four-layer WSe_2 diode with a focused laser (405 nm, roughly $1\ \mu\text{m}$ spot size and about 2.2 nW total power) excitation; the $X = 0\ \mu\text{m}$ denotes the centre position of the split-gate region and the yellow

squares denote the position of the electrodes. **e**, The $|I_{\text{sc}}|$ as a function of hole doping (black dots) compared with the reconstructed photocurrent from the exciton-diffusion-limited photocurrent model (red dots) (lines are a guide for the eyes). The inset shows the plot of $|I_{\text{sc}}|$ as a function of $\tau^{1/2}$, for hole-doped WSe_2 , with a linear fit (dashed red line). **f**, The $|I_{\text{sc}}|$ as a function of electron doping (black dots) compared with the reconstructed photocurrent from the exciton-diffusion-limited photocurrent model (red dots). The inset shows the plot of $|I_{\text{sc}}|$ as a function of $\tau^{1/2}$, for electron-doped WSe_2 , with a linear fit (dashed red line). The reconstructed photocurrent is calculated from the exciton lifetime shown in Fig. 3c. **g**, $(I_{\text{sc}})^{-2}$ versus hole doping. **h**, $(I_{\text{sc}})^{-2}$ versus electron doping, showing a turning point (marked by the blue arrow) at an electron density beyond about $1.5 \times 10^{12}\ \text{cm}^{-2}$, suggesting a reduced B_{Aug} .

By using the exciton diffusion model (equations (2) and (3)) and the exciton lifetime determined from TRPL measurements, we can reconstruct the doping-dependent I_{sc} and compare it against the experimental data. Notably, the reconstructed data (red dots) fit well with the measured photocurrent (black dots) (Fig. 4e, f). From the fit, we extract an exciton diffusion constant $D_{\text{exc}} = 7 \pm 2\ \text{cm}^2\ \text{s}^{-1}$,

similar to previous studies^{39,40}. From the exciton diffusion constant and the exciton lifetime, we calculated an exciton diffusion length of around $830 \pm 110\ \text{nm}$ near the charge-neutrality regime and of around $190 \pm 30\ \text{nm}$ in the high-doping regime. The exciton diffusion length in the intrinsic regime is about 4.4-times longer than that in the high-doping regime, which matches well with the recent photophysics

measurements where the exciton diffusion length at low doping is about 5-times larger than that at high doping⁴¹. Importantly, the extracted exciton diffusion length at low doping agrees well with our scanning photocurrent microscopy studies. The high-fidelity reconstruction of the photocurrent response with such a straightforward exciton-diffusion-limited model, along with the highly consistent exciton diffusion length extracted from our device measurements, robustly demonstrates that our 2D diodes are approaching the fundamental excitonic diffusion limit.

Within this model, I_{sc} is proportional to L_{exc} and $(\tau)^{1/2}$. We compared the $(\tau)^{1/2}$ obtained from the TRPL measurements to the doping-dependent I_{sc} (inset of Fig. 4e, f). The I_{sc} versus $(\tau)^{1/2}$ plots shows an almost linear relationship, further confirming the exciton-diffusion-limited behaviour. Moreover, combining equations (1) and (2), we expect that $(I_{sc})^{-2}$ is proportional to $B_{Aug}n$. Indeed, a plot of $(I_{sc})^{-2}$ versus doping density gives a nearly linear relationship with the hole density in the entire doping range (Fig. 4g). In contrast, the same plot of the electron-doped regime shows a turning point with a reduced slope and thus a reduced B_{Aug} when the electron doping exceeds around $1.5 \times 10^{12} \text{ cm}^{-2}$ (Fig. 4h), which is highly consistent with the optical measurements (Fig. 3d). Taken together, these results further demonstrate that our device performance is indeed dictated by intrinsic material properties rather than extrinsic non-idealities, allowing the direct correlation of photophysics parameters with practical device performance.

Similar doping-dependent photocurrent generation is observed in monolayer, bilayer and four-layer WSe₂ vdW-diodes (Methods, Extended Data Fig. 7), demonstrating the critical role of exciton–charge Auger effects in few-layer 2D diodes. We have further analysed the EQE of these diodes. The EQE is calculated as

$$\text{EQE} = I_{sc} E_{ph} / (e P_{in}) \quad (4)$$

where I_{sc} is the short-circuit photocurrent, E_{ph} is the energy per photon, e is the elementary charge and P_{in} is the input power (P_{in} = power density $P_d \times$ illuminated exciton-collection area A). However, there is considerable ambiguity in defining A . Previous studies⁸ have used the split-gate area for A , which does not consider exciton diffusion from outside the split-gate region and thus often underestimates P_{in} and overestimates EQE. Alternatively, A may be approximated by the device area between the source and drain electrodes, giving a lower-bound EQE as the exciton-collection area may be considerably smaller than the device area.

To more accurately determine P_{in} without the uncertainty associated with the illuminated charge-collection area A , we also used the total incident power of a focused laser excitation (as used in scanning photocurrent microscopy measurements) to determine the local EQE. On the basis of the photocurrent line scan across the device (Fig. 4d), we achieved a maximum EQE of 30% in a four-layer WSe₂ diode; and considering an absorbance of about 30% for the four-layer sample, we obtain an estimated unity internal quantum efficiency, further highlighting the intrinsic performance of our vdW-diodes. Similar studies with focused laser illumination previously gave considerably lower EQE values of 0.10–0.75% in monolayers (compared to 2.2% observed in monolayer vdW-diode) (Extended Data Table 2).

We further note that our vdW-diodes show, to our knowledge, the highest V_{oc} among all 2D semiconductor diodes (at the same layer thickness) reported so far (Extended Data Table 2). The V_{oc} reaches 0.90 V in the bilayer device and up to 1.05 V in the monolayer device, considerably larger than values reported previously (<0.83 V). We note that V_{oc} does not significantly depend on the charge collection area. Therefore, a higher V_{oc} provides robust evidence supporting a decrease in parasitic loss pathways in vdW-diodes.

In summary, we have shown that 2D WSe₂ diodes with defect-free vdW contacts allows the minimization of the extrinsic interfacial

disorder-dominated recombination and access to the intrinsic excitonic behaviour in 2D semiconductor devices. Our results highlight the direct impact of exciton–charge Auger effects on 2D p–n diode performance, and provide important guidance for pushing the limit of 2D diodes. First, it is vital to suppress contact-interface-state-induced SRH recombination that limits charge collection. Second, consideration of the unique 2D semiconductor physics, such as diffusion in a 2D plane and high exciton binding energy in 2D materials, is essential for exploiting these properties for improved device performance. Long exciton lifetime (that is, low Auger recombination rate) and large exciton diffusion constant are beneficial for the excitons to reach the p–n interface and undergo charge separation for more efficient photocurrent generation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-03949-7>.

- Ross, J. S. et al. Electrical control of neutral and charged excitons in a monolayer semiconductor. *Nat. Commun.* **4**, 1474 (2013).
- He, K. et al. Tightly bound excitons in monolayer WSe₂. *Phys. Rev. Lett.* **113**, 026803 (2014).
- Chernikov, A. et al. Exciton binding energy and nonhydrogenic Rydberg series in monolayer WS₂. *Phys. Rev. Lett.* **113**, 076802 (2014).
- Yuan, L. & Huang, L. Exciton dynamics and annihilation in WS₂ 2D semiconductors. *Nanoscale* **7**, 7402–7408 (2015).
- Paul, K. K., Kim, J.-H. & Lee, Y. H. Hot carrier photovoltaics in van der Waals heterostructures. *Nat. Rev. Phys.* **3**, 178–192 (2021).
- Ross, J. S. et al. Electrically tunable excitonic light-emitting diodes based on monolayer WSe₂ p–n junctions. *Nat. Nanotechnol.* **9**, 268–272 (2014).
- Baugher, B. W., Churchill, H. O., Yang, Y. & Jarillo-Herrero, P. Optoelectronic devices based on electrically tunable p–n diodes in a monolayer dichalcogenide. *Nat. Nanotechnol.* **9**, 262–267 (2014).
- Pospischil, A., Furchi, M. M. & Mueller, T. Solar-energy conversion and light emission in an atomic monolayer p–n diode. *Nat. Nanotechnol.* **9**, 257–261 (2014).
- Cheng, R. et al. Electroluminescence and photocurrent generation from atomically sharp WSe₂/MoS₂ heterojunction p–n diodes. *Nano Lett.* **14**, 5590–5597 (2014).
- Lee, C.-H. et al. Atomically thin p–n junctions with van der Waals heterointerfaces. *Nat. Nanotechnol.* **9**, 676–681 (2014).
- Li, M.-Y. et al. Epitaxial growth of a monolayer WSe₂–MoS₂ lateral p–n junction with an atomically sharp interface. *Science* **349**, 524–528 (2015).
- Zhang, Z. et al. Robust epitaxial growth of two-dimensional heterostructures, multiheterostructures, and superlattices. *Science* **357**, 788–792 (2017).
- Massicotte, M. et al. Dissociation of two-dimensional excitons in monolayer WSe₂. *Nat. Commun.* **9**, 1633 (2018).
- Wen, X., Xu, W., Zhao, W., Khurgin, J. B. & Xiong, Q. Plasmonic hot carriers-controlled second harmonic generation in WSe₂ bilayers. *Nano Lett.* **18**, 1686–1692 (2018).
- Livache, C. et al. A colloidal quantum dot infrared photodetector and its use for intraband detection. *Nat. Commun.* **10**, 2125 (2019).
- Khan, Q. et al. Overcoming the electroluminescence efficiency limitations in quantum-dot light-emitting diodes. *Adv. Opt. Mater.* **7**, 1900695 (2019).
- Hong, X. et al. Ultrafast charge transfer in atomically thin MoS₂/WS₂ heterostructures. *Nat. Nanotechnol.* **9**, 682–686 (2014).
- Yuan, L., Wang, T., Zhu, T., Zhou, M. & Huang, L. Exciton dynamics, transport, and annihilation in atomically thin two-dimensional semiconductors. *J. Phys. Chem. Lett.* **8**, 3371–3379 (2017).
- Steinhoff, A. et al. Biexciton fine structure in monolayer transition metal dichalcogenides. *Nat. Phys.* **14**, 1199–1204 (2018).
- Zhu, X.-Y. How to draw energy level diagrams in excitonic solar cells. *J. Phys. Chem. Lett.* **5**, 2283–2288 (2014).
- Pólya, G. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.* **84**, 149–160 (1921).
- Doyle, P. G. *Application of Rayleigh's Short-cut Method to Polya's Recurrence Problem*. PhD thesis, Dartmouth College (1982).
- Dean, C. R. et al. Boron nitride substrates for high-quality graphene electronics. *Nat. Nanotechnol.* **5**, 722–726 (2010).
- Chen, P. et al. Band evolution of two-dimensional transition metal dichalcogenides under electric fields. *Appl. Phys. Lett.* **115**, 083104 (2019).
- Liu, Y. et al. Approaching the Schottky–Mott limit in van der Waals metal–semiconductor junctions. *Nature* **557**, 696–700 (2018).
- Chow, C. M. E. et al. Monolayer semiconductor Auger detector. *Nano Lett.* **20**, 5538–5543 (2020).
- Wang, Y. et al. Van der Waals contacts between three-dimensional metals and two-dimensional semiconductors. *Nature* **568**, 70–74 (2019).

28. Luque, A. & Hegedus, S. *Handbook of Photovoltaic Science and Engineering* 2nd edn (John Wiley, 2011).
29. Allen, T. G., Bullock, J., Yang, X., Javey, A. & De Wolf, S. Passivating contacts for crystalline silicon solar cells. *Nat. Energy* **4**, 914–928 (2019).
30. Das, S., Gupta, G. & Majumdar, K. Layer degree of freedom for excitons in transition metal dichalcogenides. *Phys. Rev. B* **99**, 165411 (2019).
31. Lien, D.-H. et al. Electrical suppression of all nonradiative recombination pathways in monolayer semiconductors. *Science* **364**, 468–471 (2019).
32. David, A., Young, N. G., Lund, C. & Craven, M. D. The physics of recombinations in III-nitride emitters. *ECS J. Solid State Sci. Technol.* **9**, 016021 (2019).
33. Elbaz, G. A. et al. Unbalanced hole and electron diffusion in lead bromide perovskites. *Nano Lett.* **17**, 1727–1732 (2017).
34. Chuang, S. L. *Physics of Optoelectronic Devices* (John Wiley, 1995).
35. Passari, L. & Susi, E. Recombination mechanisms and doping density in silicon. *J. Appl. Phys.* **54**, 3935–3937 (1983).
36. Altermatt, P. P., Schmidt, J., Heiser, G. & Aberle, A. G. Assessment and parameterisation of Coulomb-enhanced Auger recombination coefficients in lowly injected crystalline silicon. *J. Appl. Phys.* **82**, 4938–4944 (1997).
37. Richter, A., Glunz, S. W., Werner, F., Schmidt, J. & Cuevas, A. Improved quantitative description of Auger recombination in crystalline silicon. *Phys. Rev. B* **86**, 165202 (2012).
38. Kadlec, E. et al. Effects of electron doping level on minority carrier lifetimes in n-type mid-wave infrared InAs/InAs_{1-x}Sb_x type-II superlattices. *Appl. Phys. Lett.* **109**, 261105 (2016).
39. Cadiz, F. et al. Exciton diffusion in WSe₂ monolayers embedded in a van der Waals heterostructure. *Appl. Phys. Lett.* **112**, 152106 (2018).
40. Mouri, S. et al. Nonlinear photoluminescence in atomically thin layered WSe₂ arising from diffusion-assisted exciton–exciton annihilation. *Phys. Rev. B* **90**, 155449 (2014).
41. Uddin, S. Z. et al. Neutral exciton diffusion in monolayer MoS₂. *ACS Nano* **14**, 13433–13440 (2020).
42. Martin, J. et al. Observation of electron–hole puddles in graphene using a scanning single-electron transistor. *Nat. Phys.* **4**, 144–148 (2008).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

Device fabrication and measurement

We first prepared titanium/gold (5/20 nm) split-gate electrodes with 250-nm gap on a silicon/silicon dioxide substrate by using standard electron-beam lithography and high-vacuum electron-beam evaporation. Next, we transferred a multilayer hBN flake onto the split-gate electrodes, onto which the few-layer WSe₂ was exfoliated. Finally, using a previously developed method²⁵, a pair of Au electrodes were transferred onto the WSe₂ as source and drain vdW contacts. For the diode with evaporated Au contacts, the device fabrication procedures are the same except for the final contact integration step, in which 40-nm Au contact electrodes were fabricated using the standard electron beam lithography and electron beam evaporation process. The electrical and optoelectronic properties of the devices were characterized with a precision source/measure unit (Agilent, B2902A) under vacuum at room temperature. The carrier concentration can be estimated based on the gate voltage and gate capacitance.

Time-resolved photoluminescence measurements

The device was placed in a home-built inverted epifluorescence microscope (Nikon, CFI Plan Apochromat Lambda ×40 objective) and operated in a single capacitor mode. We applied a gate voltage, V_G , to electrostatically electron (hole) dope the WSe₂. Using a 532-nm pulsed diode laser (PicoQuant, LDH-P-FA-530B), we excited the sample and recorded the photoluminescence lifetime using time-correlated single-photon counting with a silicon avalanche photodiode (Micro Photon Devices, PD-050-CTD) connected to a synchronized photon counter (Picoquant, HydraHarp 400).

Scanning photocurrent microscopy measurements

The device was placed in the above home-built inverted epifluorescence microscope (Nikon, CFI Plan Apochromat Lambda ×40 objective) and operated in photovoltaic mode as a split-gated WSe₂ p-n junction. We recorded the short-circuit photocurrent while scanning the device on a piezo stage (Mad City Labs, Nano-View/M200-3) using a 405-nm continuous-wave laser excitation (Thorlabs, CPS405).

Absorbance measurements

To estimate the sample absorbance, we implemented a reflectance measurement using the above home-built inverted epifluorescence microscope (Nikon, CFI Plan Apochromat Lambda ×40 objective). We illuminated the few-layer WSe₂ with Au under it with the corresponding laser (473 nm or 405 nm) and used reflectance from the area off the sample as the background (containing all the other components) with dark counts being subtracted from both sample and background intensity values (Ocean Insight, Flame UV-VIS Spectrometer and Andor, Marana 4.2B-11 detectors used).

Estimation of the impact of contact barrier and series resistance on the short-circuit photocurrent

The scanning photocurrent microscopy (Fig. 1f, Fig. 4c, d) clearly shows that there is no apparent photocurrent generation from the contact region, indicating that the contacts have a negligible contribution to the overall photocurrent. Although the Schottky barrier formation between Au and n-WSe₂ could in principle produce photocurrent, the formation of the p-n junction in the split-gate region suppresses the photocurrent generated at the contact interface, as also observed in a previous study⁷. In such a configuration, the carriers generated at the contact are blocked by the depletion region at the p-n junction (Extended Data Fig. 1). The photocurrent is thus dictated by the p-n diode.

We can further estimate the short-circuit photocurrent (I_{SC}) loss due to the contact resistance, which may be included in the series resistance (R_S). The Shockley diode equation with R_S and shunt resistance R_{SH} under light illumination can be written as:

$$I_{DS} = I_S \left(e^{\frac{V_{DS} - I_{DS} R_S}{\eta V_T}} - 1 \right) - I_L + \frac{V_{DS} - I_{DS} R_S}{R_{SH}} \quad (5)$$

where I_S is the saturation current, I_L is the short-circuit current when there are no parasitic resistances, V_T is the thermal voltage and η is the ideality factor.

At short-circuit condition, we get

$$I_{SC} = I_S \left(e^{\frac{I_{SC} R_S}{\eta V_T}} - 1 \right) - I_L - \frac{I_{SC} R_S}{R_{SH}} \quad (6)$$

as $|e^{\frac{I_{SC} R_S}{\eta V_T}} - 1| \leq 1$, the absolute value of first term on the right side is

$$|I_S \left(e^{\frac{I_{SC} R_S}{\eta V_T}} - 1 \right)| \leq I_S \quad (7)$$

As the I_S is less than 1.2×10^{-17} A (see ‘Extraction of R_S and R_{SH} ’ below), which is much smaller than the I_{SC} , equation (6) can be written as

$$|I_L| = I_{SC} (1 + R_S/R_{SH}) \quad (8)$$

As we can see, only excessively large R_S can reduce the short-circuit photocurrent. In our device, the R_S/R_{SH} ratio is smaller than 1% (see ‘Extraction of R_S and R_{SH} ’ below). Therefore, the loss of short-circuit photocurrent due to R_S (including the contact resistance) can be neglected in our device.

Extraction of R_S and R_{SH}

The Shockley diode equation with a series resistance R_S and shunt resistance R_{SH} can be described as:

$I_{DS} = I_S \left(e^{\frac{V_{DS} - I_{DS} R_S}{\eta V_T}} - 1 \right) + \frac{V_{DS} - I_{DS} R_S}{R_{SH}}$, where I_S is the saturation current, V_T is the thermal voltage and η is the ideality factor. An explicit equation for the diode current can then be obtained by using the Lambert W function⁸.

$$I_{DS} = \frac{\eta V_T}{R_S} W \left(\frac{I_S R_S R_{SH}}{\eta V_T (R_S + R_{SH})} e^{\frac{V_{DS} + I_S R_S}{\eta V_T (R_S + R_{SH})}} \right) + \frac{V_{DS} - I_S R_{SH}}{R_S + R_{SH}} \quad (9)$$

By fitting the experimental data with this equation (Extended Data Fig. 2), we can extract the R_S , R_{SH} , I_S and η . By fitting the data at different gate voltages, we found the I_S in our device is in the range of 6×10^{-30} A to 1.2×10^{-17} A; the R_S in our device is in the range of 28 MΩ to 280 MΩ; the R_{SH} in our device is in the range of 35 GΩ to 67 GΩ; therefore, the overall R_S/R_{SH} is always smaller than 1%. Therefore, the loss of short-circuit photocurrent due to R_S (including the contact resistance) can be neglected in our device.

Lifetime fitting process

To determine the component decay rates (k_i , i is the component of fit) for each TRPL trace, we fit each curve to a convolution of a Gaussian instrument response function with a sum of two or three exponential decay functions: nonlinear least squares weighted fit (Extended Data Fig. 4a, b). We use the error for Poissonian process: $\delta C(t) = (C(t))^{0.5}$

$$C(t) = \sum_i \frac{C_{0,i}}{2} e^{-k_i t} e^{\left(\frac{t - \sigma^2 k_i}{2} \right)} \left(1 + \operatorname{erf} \left(\frac{t - \sigma^2 k_i}{\sqrt{2} \sigma} \right) \right) + C_0 \quad (10)$$

Where t is the time of photon arrival on the TRPL trace or equivalently the time after the laser excitation; the initial peak amplitudes, $C_{0,i}$, and the rates, k_i , Gaussian width, σ , and baseline offset, C_0 , were free fitting parameters. We determined whether a fit was bi- or tri-exponential when the fit error gave a value greater than the fit value itself. For example, if we use the tri-exponential fit for the TRPL data at $V_G = -0.8$ V, the fit error is greater than the fit value (Extended Data Table 1).

Assignment of different lifetime components

The relative intensity and lifetime of different lifetime components is shown in Extended Data Fig. 4c, d (t_1 , t_2 and t_3 are defined as $t_i = 1/k_i$, i.e. decay time constant; in the main text $t_i = \tau$). We assign t_1 to be the excitons for two reasons. (1) The relative photoluminescence intensity of t_1 is most prominent, nearly 96% at low doping and 70% at high doping, which is consistent with the fact that the excitons are prevalent in 2D semiconductors. (2) The lifetime of t_1 decreased significantly as the doping increased, which is expected for excitons, as the excitons will interact with the charge and recombine through an Auger process. The t_2 might be attributed to the trap bounded state for two reasons. (1) The t_2 emerges at higher doping concentration (about $1 \times 10^{12} \text{ cm}^{-2}$) and the relative photoluminescence intensity of t_2 decreases as the doping further increases. This behaviour suggests a trap-state-filling process⁴³. Near the charge intrinsic region, the emission from the trap states is negligible, which might be attributed to the strong electron–hole Coulomb interaction in the nano-semiconductors³¹. With increasing doping, electrons or holes start to fill the localized states. The trapped electrons/holes will attract holes/electrons and form localized excitons due to the strong Coulomb interaction in the 2D semiconductors. With further increasing doping density, the localized states are fully filled and the relative photoluminescence intensity of the localized states starts to drop. (2) The lifetime of t_2 shows weak doping dependence, which is expected for the trap recombination. One may argue that the t_2 might be a trion, as a trion is also shown for higher doping concentrations. However, the trion lifetime is about 10 ps in WSe_2 as previously measured⁴⁴, which is far shorter than the lifetime of t_2 , which is about 400 ps. The t_3 can be attributed to the trap state with even slower lifetime. The lowest relative photoluminescence intensity suggests that the t_3 component is negligible.

Band structure and the possible explanation for the turning point of trilayer WSe_2

We plot a highly simplified band structure with dispersion in close agreement with micro-angle-resolved photoemission spectroscopy measurements⁴⁵ (Extended Data Fig. 5). The calculated band structure is also largely consistent with photoluminescence spectra, where we observe direct emissive recombination from K–K at 1.65 eV and indirect emission at 1.45 eV assigned to Q– Γ (these gaps also include the exciton binding energy). The energy differences between the Q and K valleys in the conduction band are approximately 118 meV. Upon doping, we anticipate two major effects that would influence the Auger recombination rate. The first is a change in the Fermi level, which would change the valley occupancy of the electrons involved in Auger recombination. The second would be bandgap renormalization, which changes both the absolute and relative energies of the bands. We hypothesize that the turning point reflects a change in the intervalley scattering rates owing to a relative increase in the population of excess electrons in the K valley of the conduction band. Our calculations indicate that the Fermi level sits approximately 70 meV below the conduction band at the turning point (about $2 \times 10^{12} \text{ cm}^{-2}$) where we observe a slowed increase in Auger recombination. This may not result in a substantial population of the K band if one neglects bandgap renormalization. However, angle-resolved photoemission spectroscopy measurements⁴⁵ show notable electron population in the K band at a charge carrier density of approximately $1 \times 10^{13} \text{ cm}^{-2}$ (with a corresponding Fermi level at 30 meV below the band edge) in bilayer and trilayer WSe_2 . A more precise understanding of the exact turning point requires a revised calculation of the band structure at moderate to high doping. Previous theoretical studies of doping-dependent exciton–charge Auger in phosphorene also show a similar nonlinear dependence on doping density⁴⁶.

Doping-dependent I_{sc} at mono-, bi- and four-layer WSe_2 diodes

In addition to trilayer devices, we have also measured the monolayer, bilayer and four-layer WSe_2 diodes and observed similar behaviour, where the I_{sc} decreases as the doping increases (Extended Data Fig. 7). We note that in a monolayer device, the photocurrent at low doping

is only about 1.3 times that at high doping. This smaller photocurrent modulation by doping might be attributed to the increased exciton binding energy and faster exciton radiative recombination rate in the monolayer owing to its direct bandgap nature. In this case, the slow exciton dissociation process and radiative recombination limit the overall photocurrent in the low-doping region^{13,47}. The bilayer and four-layer devices show similar behaviour to the trilayer devices.

Diode performance as a function of excitation power

We observed a small decrease of EQE at about ten-times-higher laser power (Extended Data Fig. 8), which may suggest an enhanced exciton–exciton annihilation begins to have a role at above 1 W cm^{-2} . The exciton–exciton annihilation is a process where two excitons interact with each other and recombine non-radiatively^{31,48}.

Band structure calculation

The band structure is calculated based on density functional theory (DFT) using the open-source plane-wave code JDFTx (ref. 49) with ONCV norm-conserving pseudopotentials⁵⁰ and a kinetic energy cutoff of 850 eV for the wavefunctions. We used PBE exchange correlation functional⁵¹, including self-consistent spin–orbit coupling and van der Waals interactions⁵², in all simulations. The structure was fully relaxed. We used 20 Å of vacuum to separate periodic images. Convergence tests were performed for all of the forementioned settings.

Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding authors upon reasonable request.

43. Chen, K. et al. Experimental evidence of exciton capture by mid-gap defects in CVD grown monolayer MoSe_2 . *npj 2D Mater. Appl.* **1**, 15 (2017).
44. Liu, E. et al. Gate tunable dark trions in monolayer WSe_2 . *Phys. Rev. Lett.* **123**, 027401 (2019).
45. Nguyen, P. V. et al. Visualizing electrostatic gating effects in two-dimensional heterostructures. *Nature* **572**, 220–223 (2019).
46. Tea, E. & Hin, C. Charge carrier transport and lifetimes in n-type and p-type phosphorene as 2D device active materials: an ab initio study. *Phys. Chem. Chem. Phys.* **18**, 22706–22711 (2016).
47. Palummo, M., Bernardi, M. & Grossman, J. C. Exciton radiative lifetimes in two-dimensional transition metal dichalcogenides. *Nano Lett.* **15**, 2794–2800 (2015).
48. Ye, Z. et al. Efficient generation of neutral and charged biexcitons in encapsulated WSe_2 monolayers. *Nat. Commun.* **9**, 3718 (2018).
49. Sundararaman, R. et al. JDFTx: Software for joint density-functional theory. *SoftwareX* **6**, 278–284 (2017).
50. Schlipf, M. & Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **196**, 36–44 (2015).
51. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
52. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).
53. Groenendijk, D. J. et al. Photovoltaic and photothermoelectric effect in a double-gated WSe_2 device. *Nano Lett.* **14**, 5846–5852 (2014).

Acknowledgements Xiangfeng Duan acknowledges the support from the Office of Naval Research through Award N00014-18-1-2707. J.R.C. acknowledges NSF Career grant number 1945572. Y.H. acknowledges the financial support from the Office of Naval Research through award N00014-18-1-2491. Y.P. acknowledges the support from Air Force Office of Scientific Research under AFOSR award no. FA9550-YR-1-XYZQ.

Author contributions Xiangfeng Duan and P.C. conceived the research. P.C., T.L.A., J.R.C. and Xiangfeng Duan designed the experiment. P.C. fabricated the devices and performed optoelectrical measurements. Z.L., P.W., S.-J.L., Z.H., Xidong Duan and Y.H. contributed to materials, device fabrications, measurements and discussions. J.X. and Y.P. conducted band structure calculations. T.L.A. and P.C. conducted the time-resolved photoluminescence and photocurrent scanning measurements. P.C., T.L.A., J.R.C. and Xiangfeng Duan performed the data analysis. P.C., T.L.A., J.R.C. and Xiangfeng Duan co-wrote the manuscript. All authors discussed the results and commented on the manuscripts.

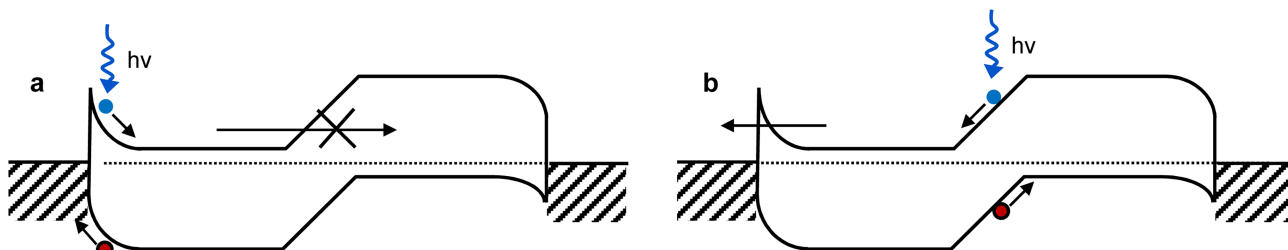
Competing interests The authors declare no competing interests.

Additional information

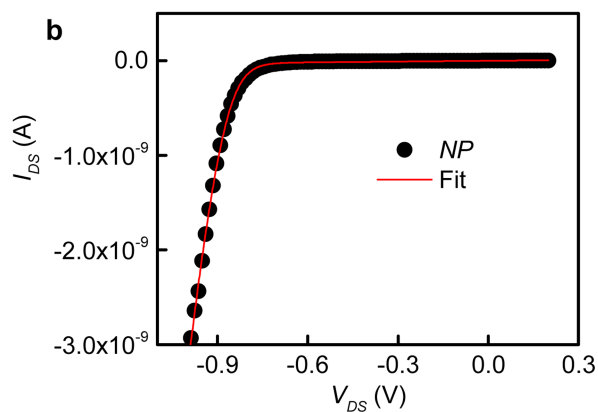
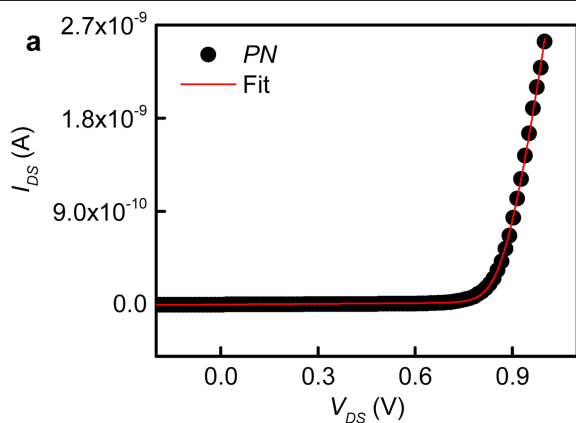
Correspondence and requests for materials should be addressed to Justin R. Caram or Xiangfeng Duan.

Peer review information Nature thanks Andrey Chaves, Lain-Jong Li and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

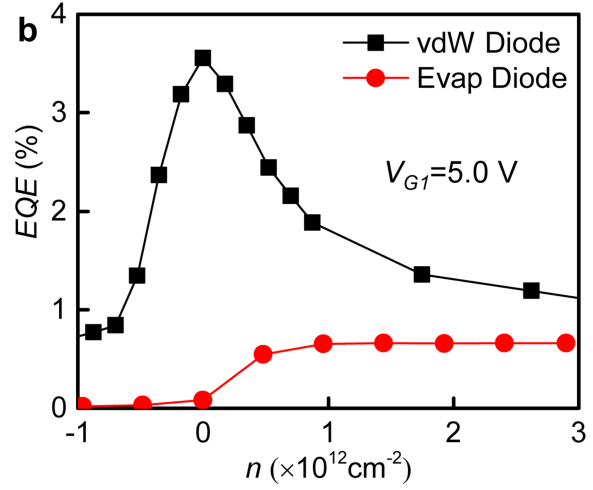
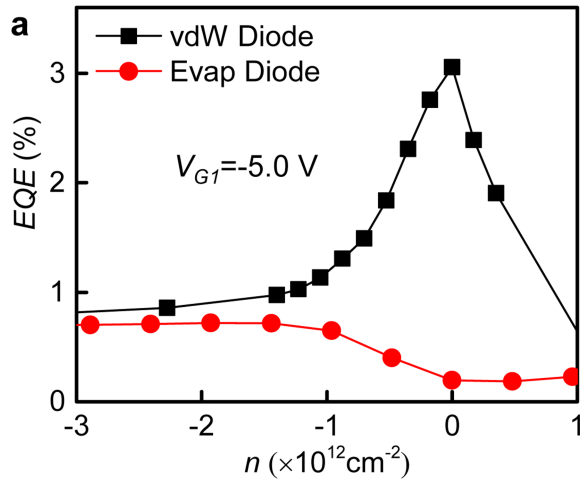


Extended Data Fig. 1 | Band diagram and photocurrent generation in diode. a, The carriers generated by the Schottky barrier are blocked by the barrier at the p-n interface. **b,** The carriers generated by the p-n junction may tunnel through the Schottky junction and contribute to the photocurrent.



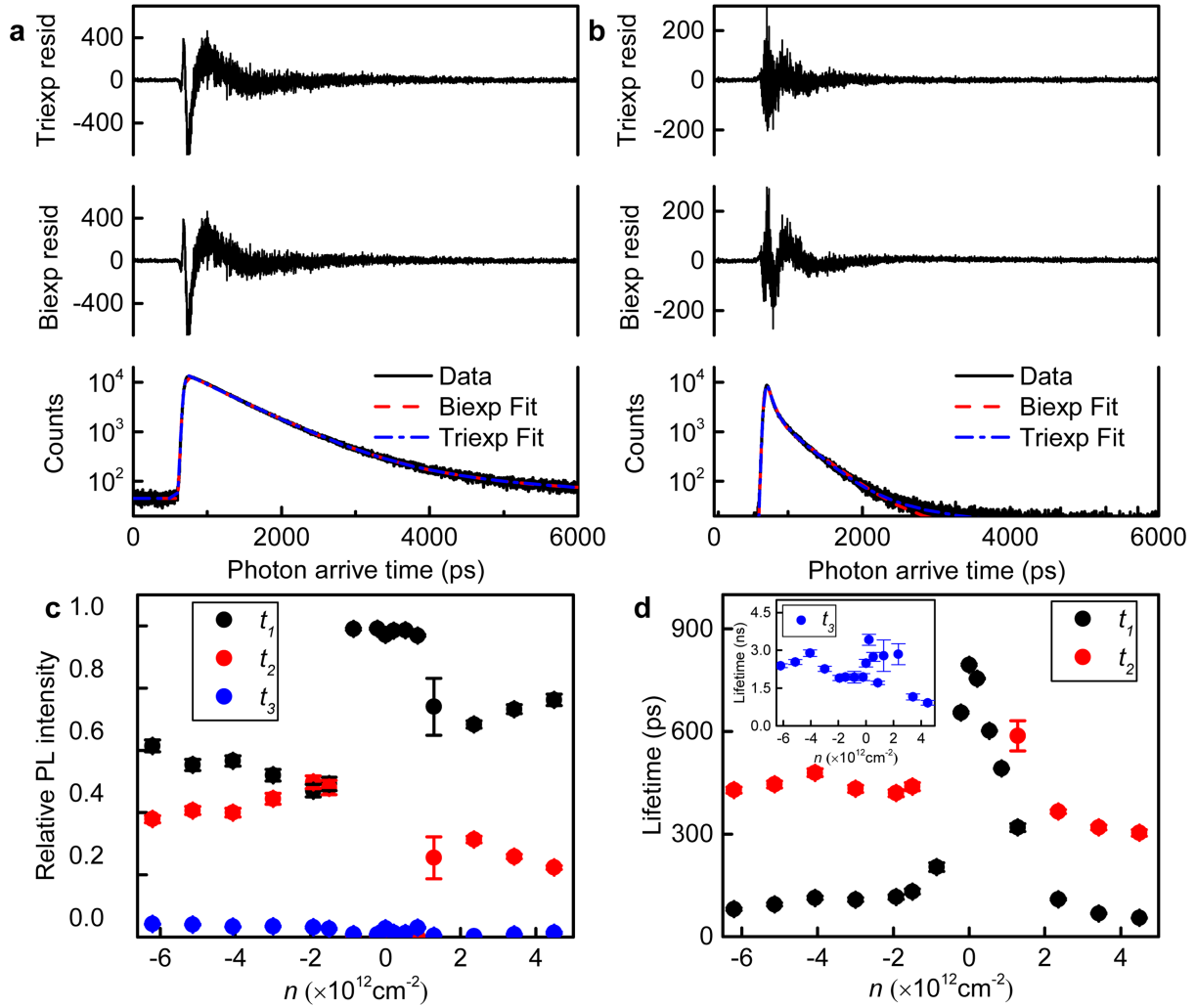
Extended Data Fig. 2 | Fitting the I_{DS} - V_{DS} characteristic of p-n junction diode. Black dot: experimental data; solid red line: fit of diode equation. **a**, The fit of p-n configuration; we extract the $R_s = 36 \text{ M}\Omega$, $R_{SH} = 47 \text{ G}\Omega$, $I_s = 4.6 \times 10^{-22}$

A and $\eta = 1.18$; **b**, The fit of NP configuration; we extract the $R_s = 28 \text{ M}\Omega$, $R_{SH} = 35 \text{ G}\Omega$, $I_s = 8.1 \times 10^{-21}$ A and $\eta = 1.3$.



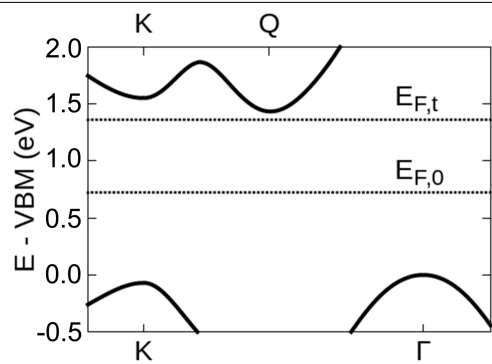
Extended Data Fig. 3 | Apparent external quantum efficiency (EQE) of 2D diodes by assuming the device area as the active area. a, EQE dependence on charge density for the evap-diode (red dots) and the vdW-diode (black dots) at $V_{G1} = -5$ V. The line serves as a guide for the eyes. **b,** EQE dependence on charge density for the evap-diode (red dots) and the vdW-diode (black dots) at $V_{G1} = 5$ V. The EQE is calculated as $\text{EQE} = I_{\text{sc}} E_{\text{ph}} / (e P_{\text{in}})$, where I_{sc} is the short circuit photocurrent, E_{ph} is the energy per photon, e is the elementary charge and P_{in} is

the input power. $P_{\text{in}} = \text{power density } (P_d) \times \text{illuminated exciton-collection area } (A)$. Note we estimated the apparent EQE by using the device area (the entire WSe_2 area between the source and drain electrodes) as A for simplicity, which may lead to a considerably underestimated EQE value as the device area is usually larger than the active area. If we consider the exciton diffusion model with a total exciton collection length of $\sim 1 \mu\text{m}$, the maximum EQE is estimated $\sim 21\%$.

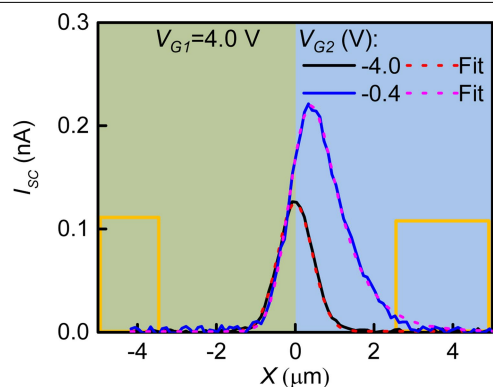


Extended Data Fig. 4 | Fitting lifetimes and doping dependence of relative PL intensity and lifetime for different components. **a**, An example of biexponential fit: ($V_G = -0.8 \text{ V}$, $P = 244 \text{ nW}$). The top panel is the residuals of tri-exponential fitting. The middle panel is the residuals of bi-exponential fitting. The bi-exponential residual is identical to the tri-exponential implying the tri-exponential is an over-fit confirmed by the error in k_3 (Extended Data Table 1); therefore, we used the bi-exponential fit. The bottom panel is the TRPL data and bi-exponential and tri-exponential fitting curve. **b**, An example of triexponential fit: ($V_G = -4 \text{ V}$, $P = 244 \text{ nW}$).

The top panel is the residuals of tri-exponential fitting. The middle panel is the residuals of bi-exponential fitting. The tri-exponential residual is better than the bi-exponential without fit errors larger than the fit values; there we used the tri-exponential fit. The bottom panel is the TRPL data and bi-exponential and tri-exponential fitting curve. **c**, Doping dependence of relative PL intensity for different components. There are three components, which are t_1 , t_2 and t_3 . **d**, Doping dependence of the PL lifetime for different components. The inset shows the lifetime of t_3 .

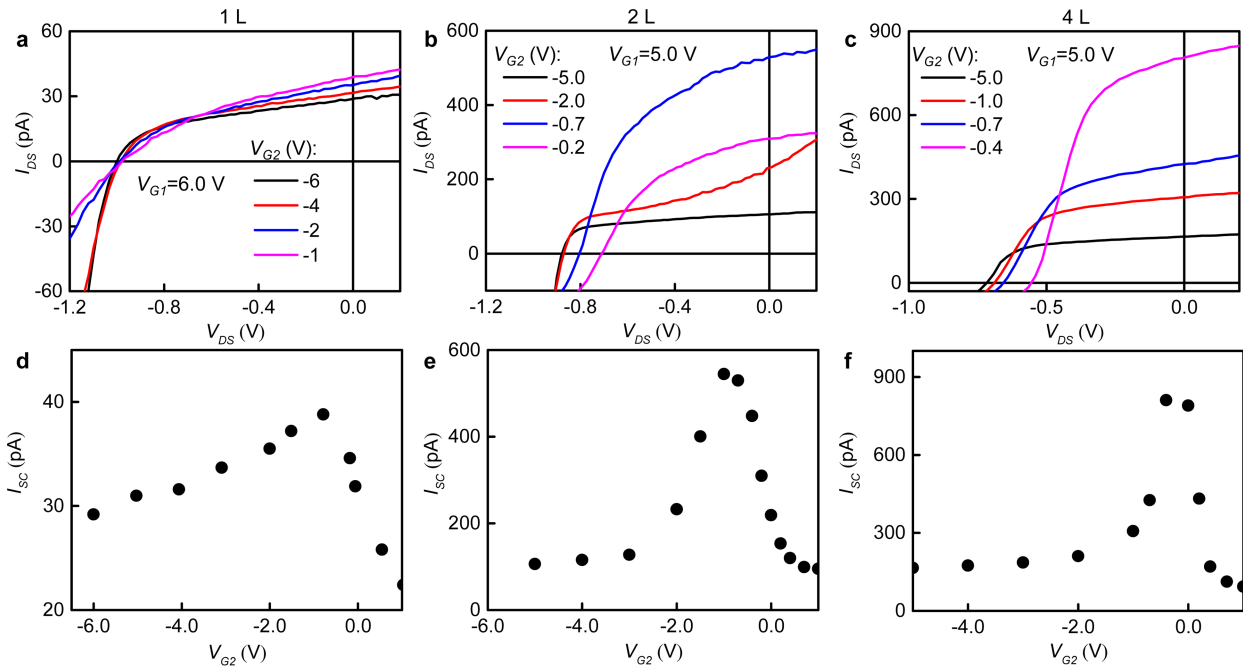


Extended Data Fig. 5 | A highly simplified band diagram showing the relevant states for band-edge carriers in WSe₂. $E_{F,0}$ denotes the Fermi level of undoped system; $E_{F,t}$ denotes the Fermi level at turning point.



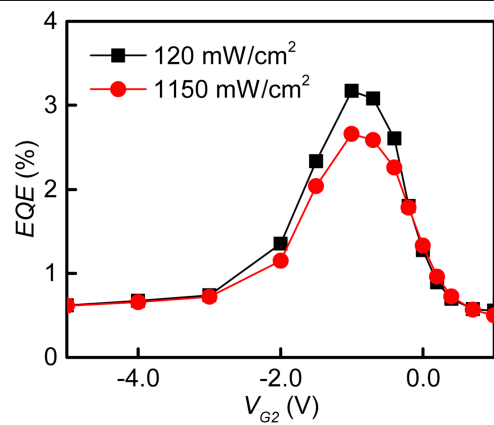
Extended Data Fig. 6 | Deconvolution of the exciton diffusion from

scanning photocurrent microscopy studies. Specifically, we used $V_{G1} = 4$ V and $V_{G2} = -4$ V (black line) as our measure of laser spot size since the photocurrent collection is exclusively from the diode interface, which is much smaller than our laser spot size (instrument response function, IRF) and fit (red dashed line) it to a single Gaussian function. We fit (pink dashed line) $V_{G1} = 4$ V and $V_{G2} = -0.4$ V (blue line) with a function being convolution of the IRF Gaussian with an exponential centred at the middle of the interface ($X = 0$ μm) for the low-doping limit. The decay constant for the fit corresponds to exciton diffusion length $L_{exc} = 0.72 \pm 0.10$ μm. The yellow square denotes the position of electrodes.



Extended Data Fig. 7 | Gate dependent I_{sc} in monolayer, bilayer and four-layer WSe₂ vdW-diodes. a, I_{DS} - V_{DS} curve of the monolayer WSe₂ diode under illumination. b, I_{DS} - V_{DS} curve of the bilayer WSe₂ diode under

illumination. c, I_{DS} - V_{DS} curve of the four-layer WSe₂ diode under illumination. d, Gate dependent I_{sc} in monolayer diode. e, Gate dependent I_{sc} in bilayer diode. f, Gate dependent I_{sc} in four-layer diode.



Extended Data Fig. 8 | Power dependent apparent EQE in a bilayer diode at $V_{G1} = 5$ V and different V_{G2} .

Extended Data Table 1 | Fitting parameters for Extended Data Fig. 4a, 4b

Fitting Extended Data Fig. 4a.	Biexponential Fit Values	Biexponential Fit Errors (δ)	Triexponential Fit Values	Triexponential Fit Errors (δ)
$C_{0,1}$	14829	9	14820	20
k_1 (ps ⁻¹)	0.0017180	2×10^{-6}	0.0017190	3×10^{-6}
$C_{0,2}$	376	8	380	10
k_2 (ps ⁻¹)	0.000486	5×10^{-6}	0.00049	2×10^{-5}
$C_{0,3}$	/	/	2	8
k_3 (ps ⁻¹)	/	/	0.0001	0.0003
σ (ps)	28.27	0.05	28.28	0.06
Fitting Extended Data Fig. 4b.	Biexponential Fit Values	Biexponential Fit Errors (δ)	Triexponential Fit Values	Triexponential Fit Errors (δ)
$C_{0,1}$	11440	40	11730	50
k_1 (ps ⁻¹)	0.014170	6×10^{-5}	0.01660	0.0001
$C_{0,2}$	2290	10	2850	20
k_2 (ps ⁻¹)	0.0025990	6×10^{-6}	0.0031330	2×10^{-6}
$C_{0,3}$	/	/	70	4
k_3 (ps ⁻¹)	/	/	0.0007780	2×10^{-6}
σ (ps)	23.96	0.06	24.64	0.07

Extended Data Table 2 | Summary of the diode parameters in different studies

Layer number	Wave-length (nm)	Power	Abs (%)	V _{OC} (V)	W (μm)	L _{CH} (μm)	L _{sg} (nm)	I _{sc} (pA)	EQE (%) (split gate area as active area)	EQE (%) (device area as active area)	EQE (%) (laser spot as active area)	IQE (%)	Source
1	473	1200 W/m ²		1.05	3.4	3.8	250	39	10	0.67	2.2		This work (vdW-diode)
2	473	1200 W/m ²		0.90	6.7	5.6	250	533	69	3.1			This work (vdW-diode)
3	473	240 W/m ²	25 ±10	0.75	16	6	250	304	83.6	3.5			This work (vdW-diode)
4	405	2.2 nW (2750 W/m ²)	30 ±10	0.73	8	6	250	220			30	100±33	This work (vdW-diode)
3	473	240 W/m ²		0.6	14	6	250	53		0.69			This work (evap-diode)
1	500-800	1500 W/m ²		0.83	3.1	1.3	460	22	1.9				Ref. 8
1	500-800	450 W/m ²		0.75	3.1	1.3	460	6	1.9				Ref. 8
1	522	2 μW		0.6	2	1.54	100				0.2		Ref. 7
1	660	10 μW			4	6.1	300	5 nA			0.1		Ref. 6
1	750	1 μW			4	6	200	4.5 nA			0.75		Ref. 13
1	532	8400 W/m ²		0.7	7.7	8.7	300	500		0.1			Ref. 53

*The EQE in ref. ⁸ is estimated by assuming the split-gate region as the active area, which may give an overestimated EQE. The EQE in ref. ⁵³ is estimated by assuming the entire device channel as the active area, which may give an underestimated EQE. The EQE in refs. ^{6,7,9} is estimated by assuming total incident laser power as P_{in}.