Charge-depletion-enhanced WSe_2 quantum emitters on gold nanogap arrays with near-unity quantum efficiency

Hongbing Cai[1,2,†], Abdullah Rasmita[1,†], Ruihua He[1], Zhaowei Zhang[1], Qinghai Tan[1], Disheng Chen[1,2], Naizhou Wang[1], Zhao Mu[1], John J. H. Eng[1,3], Yongzhi She[4], Nan Pan[4], Qian Wang[3], Zhaogang Dong[3], Xiaoping Wang [4], Juan Wang[5], Yansong Miao[1], Ranjan Singh[1,2], Chengwei Qiu[6], Xiaogang Liu[7], Weibo Gao[1,2,8]*

[1]*Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore.*

[2]*The Photonics Institute and Centre for Disruptive Photonic Technologies, Nanyang Technological University, Singapore 637371, Singapore*

[3]*Institute of Materials Research and Engineering, Agency for Science, Technology and Research (A*STAR), 2 Fusionopolis Way, Innovis 08-03, Singapore,138634, Singapore*

[4]*Department of Physics, University of Science and Technology of China, Hefei Anhui 230026, China*

[5]*College of Environmental & Resource Sciences, Zhejiang University, Hangzhou 310058, China*

[6]*Department of Electrical and Computer Engineering, National University of Singapore, Singapore, 117583, Singapore*

[7]*Department of Chemistry, National University of Singapore, Singapore, 117543 Singapore*

[8]*Centre for Quantum Technologies, National University of Singapore, Singapore.*

[†]*These authors contributed equally*

Achieving unity quantum efficiency in single-photon emitters (SPEs) is a holy grail in quantum information science. Through plasmonic coupling, it is possible to increase SPE quantum efficiency by increasing the radiative decay rate. However, to approach unity quantum efficiency, nonradiative decay must be mitigated. Here, we show that nonradiative decay in 2D WSe_2 quantum emitters can be electrically suppressed through charge depletion by utilizing dual gate configurations under a large electric field. In this condition, for site-controlled SPEs in WSe_2 coupled to gold nanogaps, the SPE transition quantum efficiency after gating is increased to 76.4 ± 14.6% on average, with some SPEs reaching near-unity (more than 90%) quantum efficiency. This study provides a novel approach for tuning SPEs with applied gate voltage and motivates further theoretical and experimental studies of SPE enhancement on vertically aligned nanogaps.

An on-demand solid-state single-photon emitter (SPE) is a crucial component in quantum technology development[1]. The localized exciton in a transition metal dichalcogenide (TMD) monolayer, such as the tungsten diselenide (WSe_2) material, is a promising SPE platform[2-6], given the possibility of deterministic placement of SPE on a chip material[7-12]. In many applications, the brightness of the SPE is an important criterion. In this regard, high collection efficiency, large emission rate, and near-unity quantum efficiency are desired. Due to the absence of total internal reflection at the surface, SPEs in 2D materials have the unique advantage of high collection efficiency[13]. Moreover, by utilizing strain and plasmonic Purcell effect, the SPE can be site-specifically defined with an enhanced emission rate[13-18]. However, it is not possible to achieve a unity quantum efficiency with the above methods.

Nonradiative decay suppression is essential to enhance the brightness further and approach the unity quantum efficiency. In particular, the quantum efficiency can be expressed as $\eta = \alpha_P k_R / (\alpha_P k_R + k_{NR})$, where α_P is the Purcell enhancement factor and $k_{R(NR)}$ is the radiative (nonradiative) decay rate. From this expression, it is clear that $k_{NR} = 0$ is needed for $\eta = 1$. Benefitting from the atomic thickness of the 2D material, the nonradiative decay pathway of the free exciton in a TMD monolayer can be suppressed by using gate voltages[19-21]. In particular, the gate voltages can be utilized to suppress nonradiative trion generation[19,22] and exciton-charge Auger recombination[21]. However, this suppression has not been achieved for localized excitons relevant to SPEs.

In this work, we report the successful suppression of the nonradiative decay of plasmon-enhanced localized WSe_2 excitons. The strategy we employed is illustrated in Fig. 1a. Initially, the exciton has a low radiative decay rate due to the low local density of states (LDOS) and a high nonradiative decay rate due to high exciton-charge interactions. To increase the SPE

brightness, the intensity is enhanced in two steps. In the first step, to increase the LDOS (i.e., Purcell effect), the WSe_2 monolayer is placed near a metal structure with plasmon energy on resonance with the localized exciton emission energy. The Purcell effect results in a higher radiative decay rate, which increases the SPE brightness. In the second step, charges are driven away from the neutral exciton by applying an electric field in the active medium, controlled by a dual gate configuration of the top and bottom gates. This charge depletion[23,24] results in a reduced interaction between the exciton and charges. As a result, the nonradiative decay is suppressed, increasing the SPE's lifetime and brightness.

Compared to uncoupled SPE, an intensity enhancement factor of 21-45 times was observed with plasmon-coupled SPE. When electrical gating was applied, the intensity was further enhanced, resulting in an overall two orders of magnitude increase in intensity and an enhanced quantum efficiency that can reach near-unity (> 90%). The SPE brightness and lifetime exhibit the same gate dependence, indicating the gate modulation of the nonradiative decay pathway. The gate dependence of PL linewidth and the second-order photon correlation confirms the control of nonradiative decay by electrical gating. Unlike previous studies using a single-gate configuration[19-21], our dual-gate configuration enables the decoupling of electrical doping and electric field contributions. We unambiguously show that the gate voltages affect the SPE mainly through the electric field instead of electrical doping[19-21]. Our demonstration provides a viable method to achieve SPE with near-unity quantum efficiency in layered materials by nonradiative decay suppression.

The device structure used in the experiment is illustrated in Fig. 1b (see Methods for more details on dimensions). The main structure consists of a WSe_2 monolayer on a metallic nanogap array[25,26] with a thin dielectric spacer layer of hexagonal boron nitride (h-BN) in between to

avoid quenching. The WSe_2 monolayer is electrically grounded. This structure is encapsulated by the top h-BN and bottom silicon dioxide (SiO_2) dielectric layers, which are connected to the top gate (few layers graphene (FLG)) and bottom gate (heavily doped Si substrate), respectively. This dual-gate configuration is beneficial for separating the electric field effect from the doping effect. In addition to top, bottom, and ground (drain) electrodes, a source electrode was added to the WSe_2 layer to study the charge transport of the monolayer WSe_2, which we did not use in this study. The strain at nanogap edges localizes the excitons, resulting in the deterministic generation of SPE at these locations. Additionally, the confinement of the electromagnetic field at these locations enhances the coupling between the plasmonic mode and the WSe_2 SPE.

The optical microscope image of the fabricated device is shown in Fig. 1c, with the zoom-in showing the scanning electron microscope (SEM) image of the nanogap array with a highly uniform gap width of 10 nm (see Supplementary Information Fig. S1 for a typical SEM image at the edge of the nanogap array). The fabrication parameters of the nanogap array were chosen to match the plasmonic resonance with the SPE energy at around 765 nm[8,9] (see Supplementary Information Fig. S2 for the reflectance spectrum of the nanogap array), and the location of hot spots was aligned to the expected location of strain-induced SPE (see Supplementary Information Fig. S3 for electric field distribution simulation). The conformality of the fabrication method enables a significant plasmonic enhancement, self-aligned to the strain-induced SPE (see Methods and Extended Data Fig. 1).

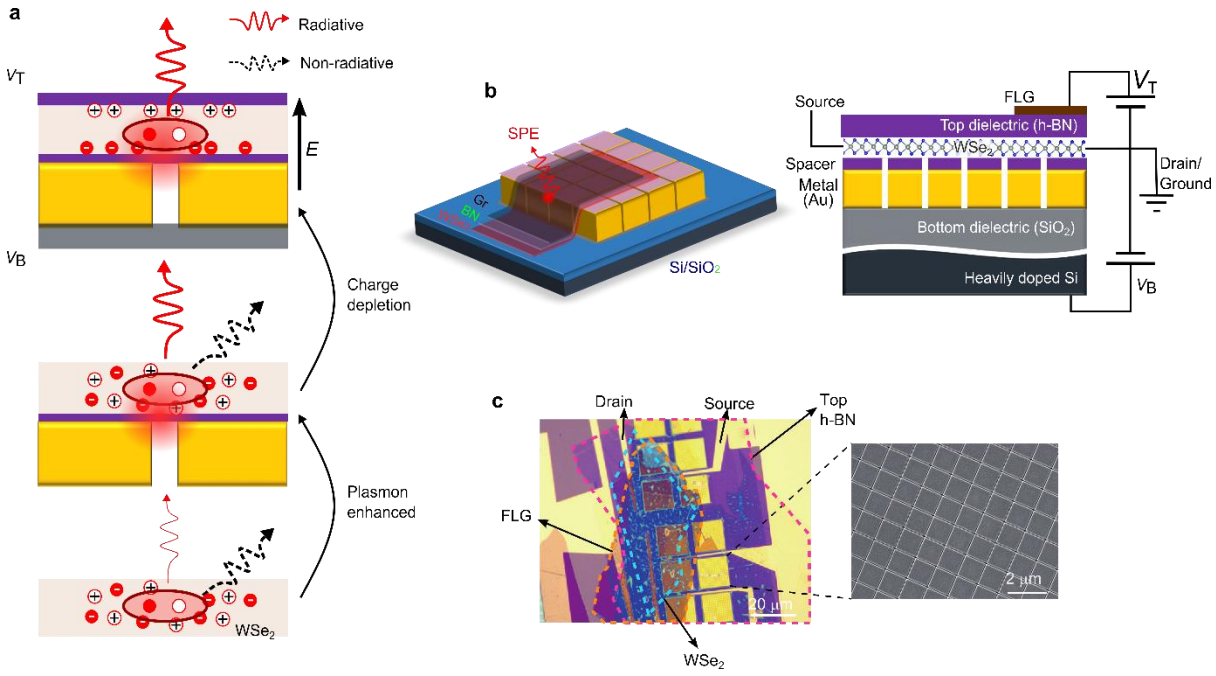


Figure 1 | Design of site-controlled WSe_2 emitters with plasmonic enhancement and gate tunability. a, Working principle. The red (black) wiggly arrows indicate the (non)radiative decay of the SPE. The radiative decay rate from the nanogap array increases due to plasmonic coupling (red blurred circle). The intensity is further enhanced by applying an electrical gating to suppress the nonradiative decay. **b,** Device structure. The layers, from top to bottom, are 3 nm FLG top electrode, 15 nm thick h-BN top dielectric, monolayer WSe_2, 5 nm thick h-BN spacer, 130 nm thick gold nanogap array, 300 nm thick SiO2 bottom dielectric, and a heavily doped Si substrate as the bottom electrode. **c,** Optical microscopy image of the device. The nanogap arrays have periods between 800 nm and 1000 nm. Zoom-in: scanning electron microscopy (SEM) image of the fabricated gold nanogap array with a 1000 nm period and a 10 nm width.

We first studied the continuous-wave (CW) PL at zero-gate voltages. Fig. 2a shows the PL intensity map under a 726 nm CW laser excitation. Bright emitters mainly reside at nanogaps, showing the deterministic placement of SPEs and plasmonic enhancement. The PL spectra of SPEs on the bare SiO_2 substrate and on the nanogap are given in Fig. 2b for comparison. On average, the intensity was enhanced by one order of magnitude by nanogaps. The power-dependent spectra (Supplementary Information Fig. S4) show that the dominant peak is the same at low and high power, indicating the emission is from exciton instead of biexciton.

Further evidence of the enhancement of this excitonic emission is provided by the power-dependent PL intensity of one filtered peak with a linewidth of less than 4 meV (Fig. 2c). Intensity enhancement was observed in all power levels with an enhancement factor of $\lambda_{int} = I_{gap}/I_0 \approx 21$ (with $I_{gap(0)}$ is the emission intensity in (no) gap region). The SPE power dependence obtained from multiple locations shows that the λ_{int} values range between 21 and 45 (see Supplementary Information Fig. S4).

To further demonstrate the quantum nature of the chosen emitters, second-order photon correlation ($g^{(2)}(\tau)$) measurements were performed for filtered emissions (containing only one peak) from SPEs on SiO_2 and on nanogap. The results obtained from two emitters are shown in Figs. 2d and 2e, with $g^{(2)}(0) = 0.38 \pm 0.06$ for SPE on SiO_2 and $g^{(2)}(0) = 0.18 \pm 0.06$ for SPE on nanogap, respectively. Both emitters show $g^{(2)}(0) < 0.5$, confirming their single-photon nature, with SPE on nanogap showing a shorter decay time. The shorter decay time of SPE on nanogap is further confirmed by measuring the time-dependent PL of filtered emission under 726 nm pulse excitation with an average power of 400 nW. The results are shown in Fig. 2f, together with the impulse response function (IRF). The decay times (see Methods for the fitting procedure) for SPEs on nanogaps and SiO_2 are 0.4 ns and 8.08 ns, respectively, giving a rate enhancement factor of $\eta_\tau = \tau_0/\tau_{gap} \approx 20$ (with $\tau_{gap(0)}$ is the emission lifetime in the (no) gap region), indicating the Purcell effect (see Supplementary Information Section 1 for the discussion on Purcell factor).

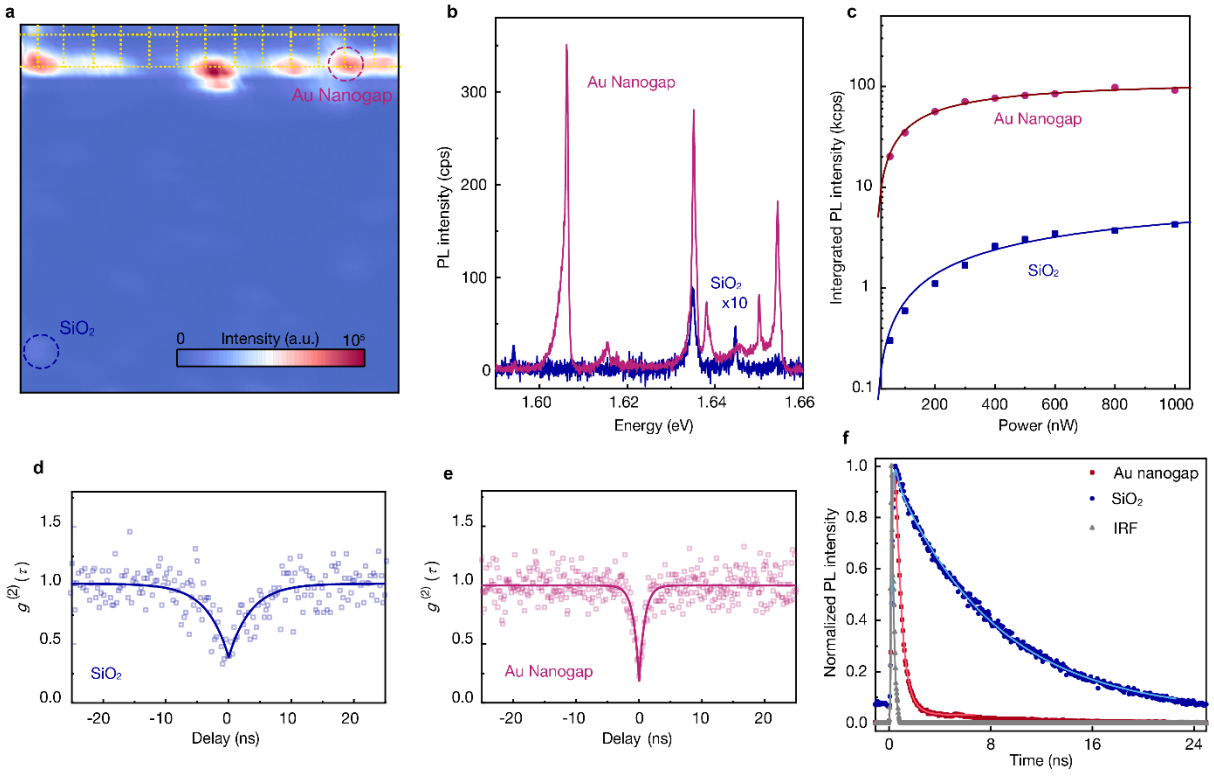


Figure 2 | Dynamic optical characterization of strain-induced SPEs coupled to metallic nanogap arrays. a, PL intensity mapping of the WSe_2 monolayer device. The dashed line denotes the nanogap array. Dashed circles show typical positions of SPEs on SiO_2 and nanogaps. **b** PL spectra of SPEs on the SiO_2 (CW excitation power: 300 nW) and the nanogap (CW excitation power: 200 nW) substrate. Intensity enhancement is of one order of magnitude. The characterization of the filtered peaks is shown in d-g. **c,** Power dependence, **d-e,** CW second-order photon correlation, and **f,** time-resolved PL (under 400 nW 726 nm pulse excitation) of the SPEs on the SiO_2 substrate and gold nanogaps. In d, the intensity is obtained from the total intensity of filtered spectra and converted to the photodetector-equivalent value considering the spectrometer and photodetector efficiency. The lines are fitting with the power saturation function (see Supplementary Information Section 1). In f, the grey triangles show the instrument response function (IRF) obtained from the back-reflected laser light, while the lines are fitting results. All figures are obtained without gate voltages applied. All normalization in this work is obtained by dividing by the maximum value. All data is taken from device #1.

Next, we studied the gate tunability of the SPE. Fig. 3a shows the top gate voltage (V_T)-dependent PL spectrum of another SPE on the nanogap with the back gate grounded ($V_B = 0$ V) under the 726 nm CW excitation (see also Supplementary Information Fig. S5). The SPE

emission remains stable regardless of the gate voltages (see Extended Data Fig. 2). Additionally, the plasmonic resonance is unaffected by gate voltage (Extended Data Fig. 3). The total PL intensity is enhanced as V_T becomes more negative, indicating gate tunability. The single photon emission is shown by $g^{(2)}(0) < 0.5$ (see Fig. 3b). Moreover, when the gate voltage is applied, the SPE purity is improved, as indicated by the reduction of $g^{(2)}(0)$ from $g^{(2)}(0) = 0.29 \pm 0.02$ at $V_T = 0$ V to $g^{(2)}(0) = 0.04 \pm 0.01$ at V_T = -1.2 V . The shape of $g^{(2)}(\tau)$ also experiences a change with the applied gate voltage. Specifically, the shoulder peak at $\tau \approx 6$ ns observed in $g^{(2)}(\tau)$ at $V_T = 0$ V is absent at V_T = -1.2 V . Considering that such a shoulder peak comes from nonradiative decay pathways[27,28], the gate dependent $g^{(2)}(\tau)$ indicates that gate voltage can affect exciton dynamics through the nonradiative decay suppression.

To study the gate dependence of the exciton dynamics further, we conducted time-resolved PL measurements on this SPE. The corresponding integrated PL intensity is plotted together with the PL lifetime obtained from the time-resolved PL under 726 nm pulse excitation as a function of V_T in Fig. 3c (see Extended Data Fig. 4 for the example of PL fitting results). We observe that the PL lifetime and intensity exhibit a similar gate dependence, further confirming that the gate voltages modulate nonradiative decay[29-31], as illustrated in Fig. 3c. This concurrent change in intensity and lifetime is also observed in other SPEs from the same device (Extended Data Fig. 5 (a-d)) as well as in SPE from another device (Extended Data Fig. 6), indicating the reproducibility of this mechanism. This behaviour is observed in both low and high excitation power (Extended Data Fig. 7). We also observed that the PL spectrum gets narrower when V_T is reduced, i.e., becomes more negative (Extended Data Fig. 8), which can be understood as a consequence of the nonradiative decay suppression. In particular, the emission blinking due to

nonradiative decay pathways is suppressed, resulting in a narrower PL spectrum.

The observed modulation of the nonradiative decay rate should also change the SPE's quantum efficiency. To study the gate dependence of the quantum efficiency, we measured the power-dependent integrated PL intensity under pulse and CW excitation at different gate voltages (Fig. 3d). In both cases, the saturation count is enhanced after applying the gate voltage. Under an 80 MHz pulse excitation, the saturation count is increased from ~38 kilo counts per second (kcps) at $V_T = 0$ V to ~145 kcps at $V_T = -1.2$ V. From the calibration measurement, the collection efficiency is ~0.3%, which shows that the quantum efficiency is increased by ~4 times from 16% at $V_T = 0$ V to 60% at $V_T = -1.2$ V for this emitter. A similar conclusion is obtained from the analysis of the emission under CW excitation, where the saturation count is increased from 450 kcps (quantum efficiency of ~17%) at $V_T = 0$ V to 1280 kcps (quantum efficiency of ~66%) at $V_T = -1.2$ V. The gate-induced count increase is also observed in another device (Extended Data Fig. 6c), showing the reproducibility of this observation. The statistical analysis of multiple emitters shows that, by applying gate voltages, the average quantum efficiency reaches 76.4±14.6 %, with the highest quantum efficiency exceeding 90% (see Extended Data Fig. 9 and 10). More details on the collection efficiency calibration and quantum efficiency calculations can be found in Supplementary Information Section 1.

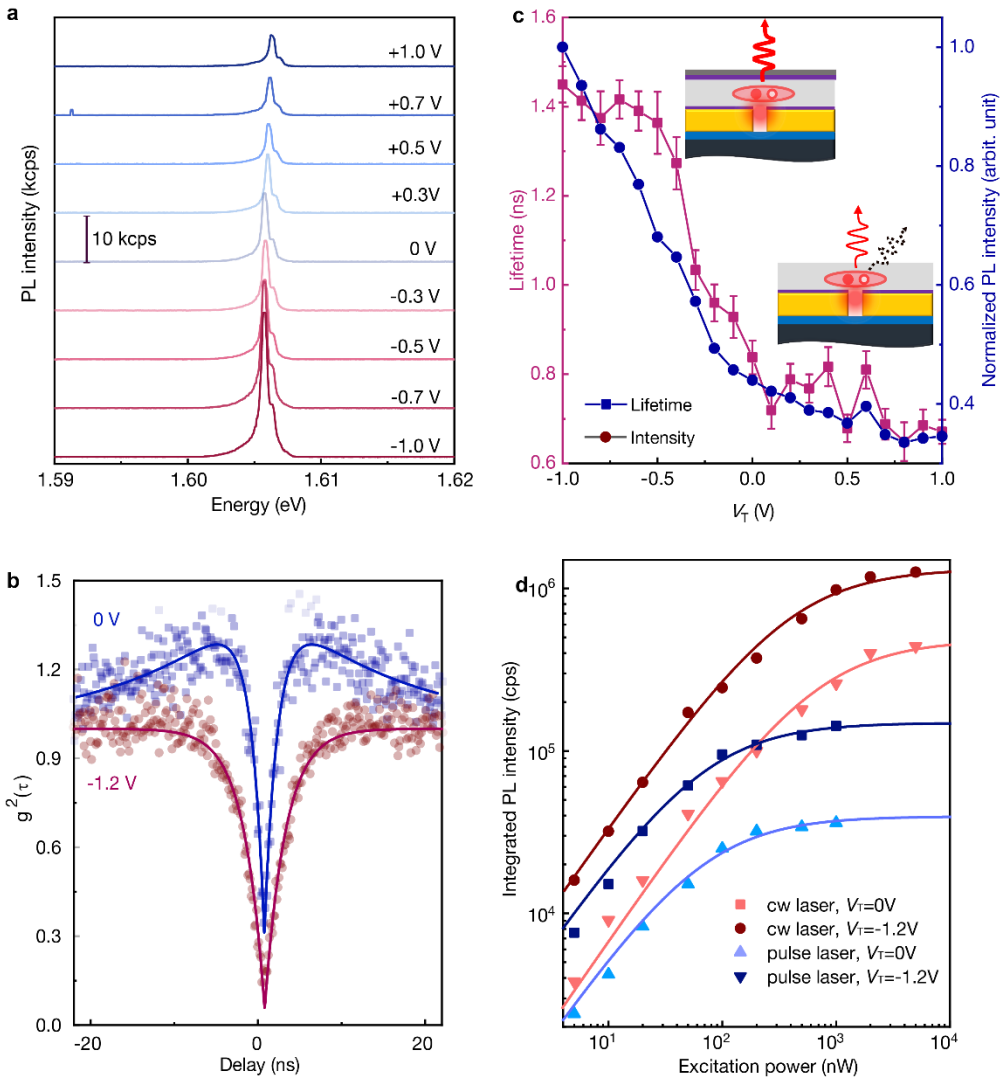


Figure 3 | Gate dependence of SPE on nanogaps. a, Gate-dependent PL spectra under 300 nW 726 nm CW excitation. **b,** The gate dependence of $g^{(2)}(\tau)$ for the same SPE under 1200 nW 726 nm CW excitation. The SPE purity is improved after gating. **c,** Gate-dependent PL lifetime and normalized integrated PL intensity of filtered emission under 500 nW 726 nm pulse excitation. The PL lifetime and intensity have a qualitatively similar shape, indicating a gate-modulated nonradiative decay process. **d,** Power dependence for the same SPE at different gate voltages under CW and pulse excitation. The lines are fitting with the power saturation function (see Supplementary Information Section 1). The repetition rate for the pulse laser is 80 MHz. The quantum efficiency is improved by ~4 times after gating. A 750 long pass filter and tilted 780±10nm band pass filter were used to collect the spectra and APD counts in a and d, while a 750 long pass filter and tilted 780±3nm band pass filter were used in b and c. In a-d, only the top gate voltage is modulated and the back gate is grounded. All the data is taken from emitter E3 in device #1.

To uncover the mechanism behind the nonradiative decay suppression, we study the time-resolved PL emission under dual-gate modulation (top gate, V_T and back gate, V_B), which allows us to analyze the effect of carrier density and electric field independently. Both intensity and lifetime maps are similar (Fig. 4a and 4b), agreeable with the electrical gating modulation of nonradiative decay. The intensity can be increased with either gate applied. The PL intensity map can be divided into three different regions, labelled A, B, and C (Fig. 4a). The PL intensity is highest in region A and lowest in region B. The dividing lines between the regions have a slope, $s = dV_B/dV_T \approx 1$ (purple dash-dotted lines in Fig. 4a) when the WSe_2 is in the intrinsic doping regime, and the slope increases as the Fermi level approaches the valence band maximum (black dashed line in Fig. 4a, obtained using the charge density simulation described in Supplementary Information Section 3). This behaviour indicates that PL intensity is determined by the electric field rather than electrical doping (see Extended Data Fig. 11 and Supplementary Information Section 3). We note that all time-resolved PL data show exponential decay, indicating that exciton-exciton annihilation is not the dominant localized exciton decay process[32].

This dependence on the electric field can be explained by considering that the field shifts the charge cloud away from the neutral exciton and towards the dielectric layer (charge depletion), as illustrated in Fig. 4c (also see Fig. 1b). Due to the reduced overlap area between the exciton and charge, the nonradiative transition caused by Auger recombination and nonradiative trion generation is suppressed, resulting in an increase in both PL intensity and lifetime. The assignment of nonradiative decay to the Auger recombination is supported by the temperature dependence experiment showing that the decay rate increases with increasing temperature following the Auger process in 2D transition metal dichalcogenide (Extended Data Fig. 12 and Supplementary Information Section 5). The charge cloud can shift towards the top or bottom

dielectric, resulting in two transitions from low to high PL intensity regions, corresponding to B-to-A and B-to-C transitions, respectively, in Fig. 4a. The difference in PL intensity between regions A and C can be attributed to the asymmetry between top and bottom dielectric layers. Considering that the surrounding charge is positive, a higher intensity in Region A ($V_B > 0\,V @ V_T = 0\,V$) compared to Region C ($V_B < 0\,V @ V_T = 0\,V$) shows that it is easier to shift the charge cloud to the top dielectric layer than to the bottom dielectric layer. Finally, we note that the effect of charge depletion on the radiative decay rate also agrees with experimental observation (more discussion on this is presented in Supplementary Information Section 2).

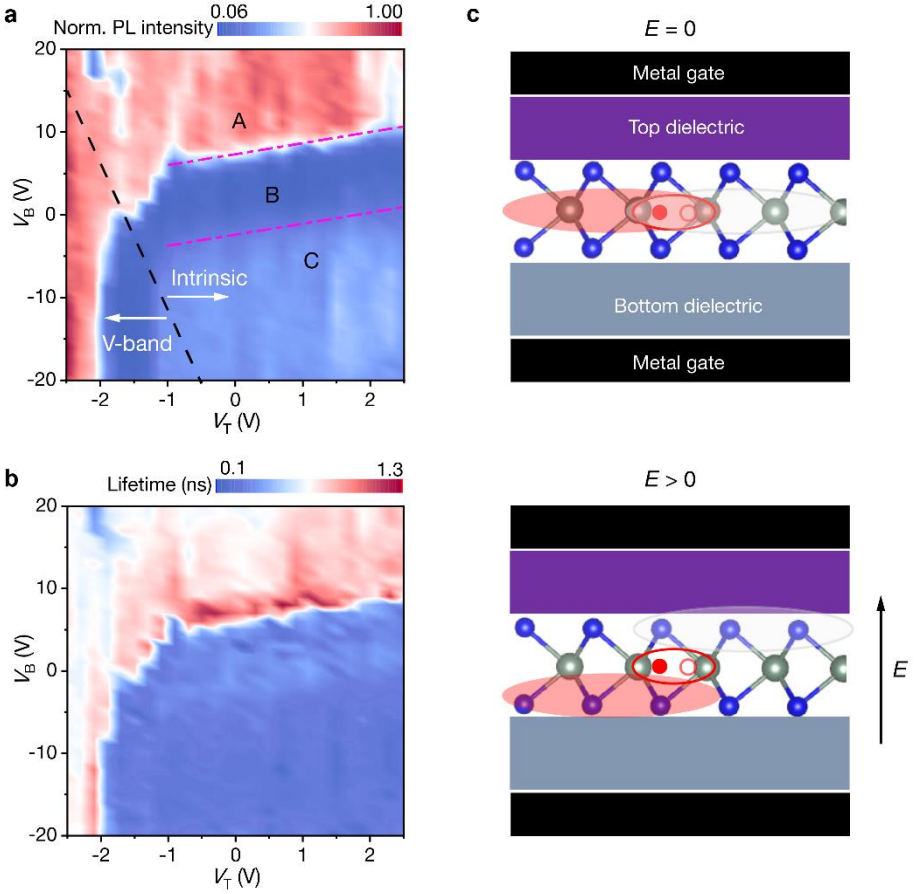


Figure 4 | Dual gate characterization and nonradiative decay suppression mechanism (device #1). a, Gate-dependent map of PL intensity. The black dashed line indicates the valence band (V-Band) maximum. The map is divided into three regions (A, B, and C) with different PL intensities. The dividing lines between the regime in the intrinsic doping regime are shown with two purple dash-dotted lines, both having slopes, $s = dV_B/dV_T = 1$. **b,**

Gate-dependent PL lifetime map. The similarity between the lifetime and the PL intensity maps indicates that gates modulate the nonradiative decay rate. **c,** Mechanism of gate tunability. The electric field reduces the overlap area between the charge cloud and exciton, reducing nonradiative decay and screening.

In conclusion, we have reported the enhancement of PL intensity of the site-controlled SPEs by combining plasmonic coupling and charge depletion. SPE site control and plasmonic enhancement of radiative decay are achieved using a gold nanogap array, while charge depletion-induced nonradiative decay suppression is achieved by utilizing a dual-gate configuration. Combined with the high collection efficiency of the 2D quantum emitter system, the WSe_2 SPE can be used to realize a bright, deterministic, on-demand quantum emitter desirable for many applications, including quantum communications and scalable optical quantum computation. The field-induced charge redistribution demonstrated in this work may also be applicable to other layered materials, potentially opening up new ways to control charge carrier transport in 2D heterostructures.

Methods

Sample fabrication

The detailed fabrication process for the nanogap array[33] is depicted in Extended Data Fig. 1. First, a 130-nm thick gold film was deposited onto a 300-nm thick SiO_2 layer on a heavily doped Si substrate by a standard thermal evaporation process (deposition rate of ~0.06 nm/s). A thin h-BN spacer layer of ~5 nm thick was transferred to the gold film and sent for the focused ion beam milling. A 10-nm wide nanogap array was fabricated with an ion current of 6 pA at 30 kV to achieve high lateral resolution. Then, the part of the continuous gold film without the pattern was peeled off with commercial adhesive tape, leaving behind the nanogap array on the SiO_2/Si substrate.

The source, drain, and top gate electrodes were fabricated by standard electron beam lithography (EBL), followed by metal deposition of a 5/50 nm-thick Cr/Au layer. A monolayer WSe_2 and a ~15 nm-thick h-BN top dielectric layer were then sequentially transferred to the nanogap array using a commercial dry method. Afterwards, the FLG top electrode layer (~3 nm thick) was placed on the h-BN and connected to one of the gold gate electrodes for electrical connection.

Measurement setup

The optical setup is illustrated in Supplementary Information Fig. S6. During the experiments, the samples were mounted in a cryostat (AttoDry 1000) at 4 K. An objective lens (numerical aperture = 0.9) focused the beam on the sample surface with a sub-μm excitation diameter. To obtain the spectrum, a tunable CW laser (M squared) with a wavelength of 726 nm was used. A 50/50 fibre coupler was used to split the signal into two. One signal was directed to a spectrometer (Princeton Instrument, PyLoN CCD), and another one to a superconducting single-photon detector (Gifford-McMahon cycle SSPD from SCONTEL) or avalanche photodiode (APD, Excelitas Technologies Inc.). The gratings used in the spectrometers were 300 g/mm and 1200 g/mm, both blazed at 750 nm. For $g^{(2)}(\tau)$ measurement, the emission count recorded by the SSPD or APD was sent to a time-correlated single-photon counter (TCSPC, Picoharp, PH300). Intensity mapping was obtained by scanning the sample using a pair of high-resolution stages (ANSx150/LT). For time-resolved PL, a 726-nm pulse excitation with an average power of 400 nW and 80 MHz repetition rate was used. The time-resolved PL was then measured using the same TCSPC system used for $g^{(2)}(\tau)$ measurements. During the gating dependence experiments, the applied voltages were controlled using a multiple-ports Keithley 2636B source meter.

Time-resolved PL fitting procedure

The time-resolved PL signal, $y_{\text{meas}}(t)$, is fitted using the following function,

$$y(t) = \left(I_1 / t_1 \exp\left(- t / t_1 \right) + I_2 / t_2 \exp\left(- t / t_2 \right) \right) * f_{\text{IRF}}(t) + y_0 ,$$

where $f_{\text{IRF}}(t)$ is an analytical function describing the IRF obtained from the IRF measurement and $I_1 \geq I_2$. The lifetime of a time-resolved PL signal is defined as t_1.

Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Acknowledgments

We thank the discussion with Indra Yudhistira and Javier Aizpurua. This work is supported by the Singapore National Research Foundation through QEP programme (No. NRF2021-QEP2-03-P01, No. NRF2021-QEP2-03-P10) and Singapore Ministry of Education (MOE2016-T3-1-006 (S)). W.G. and Z.D. would like to acknowledge the funding support of A*STAR MTC IRG grant (Project No. M21K2c0116).

Author contributions

All authors contributed extensively to the manuscripts.

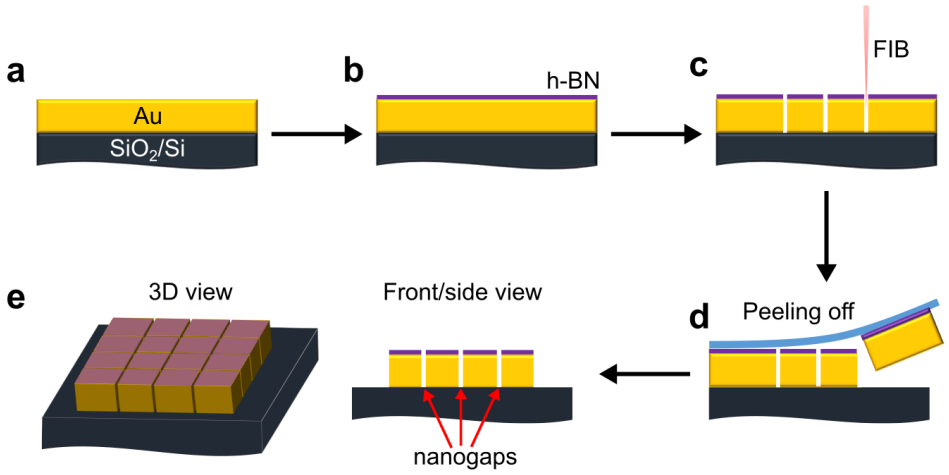
Competing interests

The authors declare no competing interests.

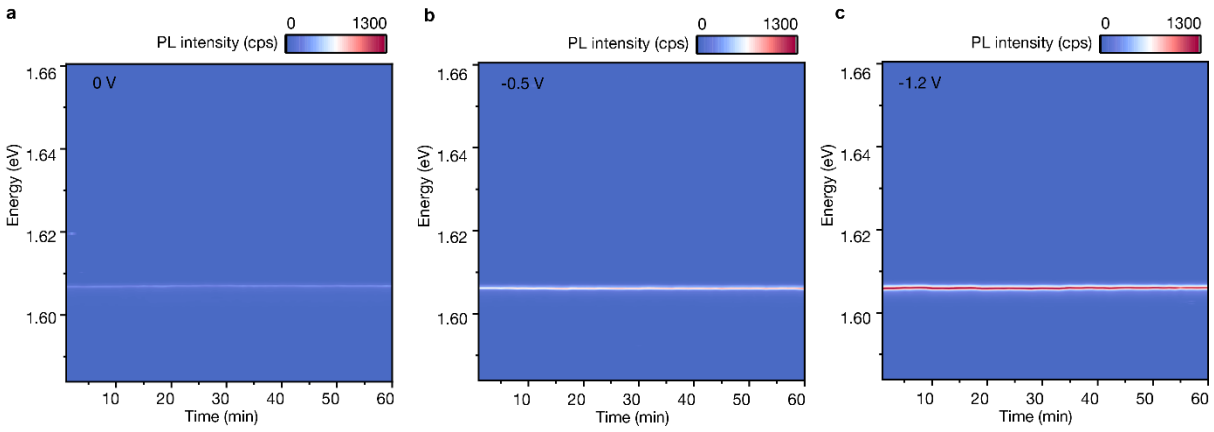
Additional information

Correspondence and requests for material should be addressed to Wei-bo Gao (wbgao@ntu.edu.sg).

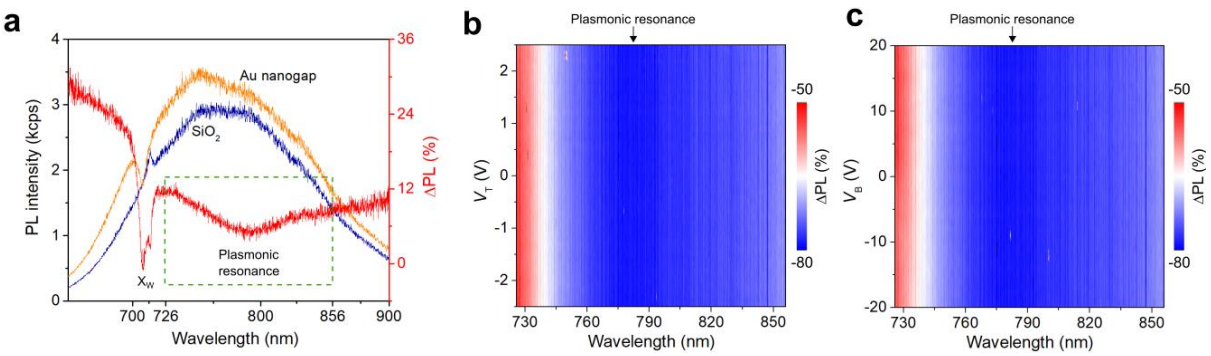
Extended Data Figures



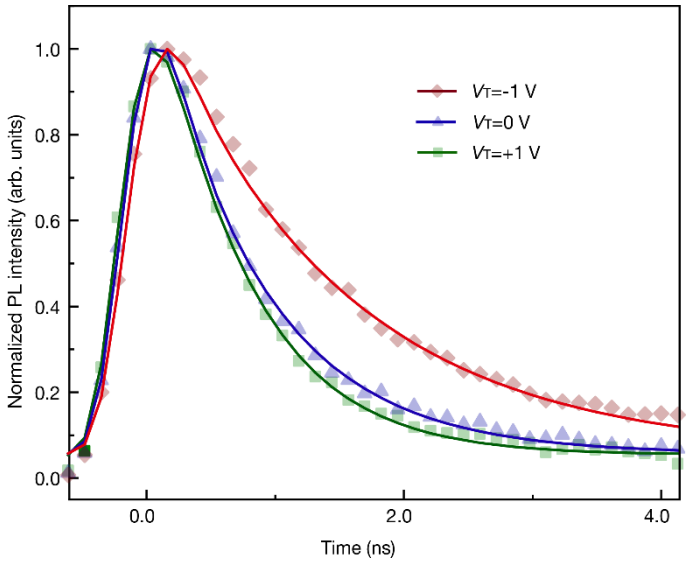
Extended Data Figure 1 | Fabrication of the nanogap array with sharp edges covered with h-BN by the peeling-off method. a, Deposition of a continuous gold film on the SiO_2 substrate by standard thermal evaporation with a gold deposition rate of ~0.06 nm/s. **b,** Transfer of a mechanically exfoliated h-BN thin layer on the gold film by the common dry method with PDMS. **c**, FIB milling of nanogaps with designed size and shapes, using an ion current of 6.0 pA at 30 kV to achieve high resolution. **d**, Fabrication of nanogap array with the peeling off method. The unpatterned gold film was peeled off while the as-prepared nanogap array was left behind. **e,** Schematic of the fabricated nanogap array with h-BN on top.



Extended Data Figure 2 | Time trace of CW PL mapping of the one emitter under different top gate voltages. The PL time traces at the top gate voltage of **a,** 0 V, **b,** -0.5 V, and **c,** -1.2 V, are shown. A bandpass filter was used to choose the emission peak with an energy of ~1.605 eV. While the PL intensity shows a gate dependence, the PL intensity remains stable in the time trace, indicating the absence of spectral jump and a stable photon emission.

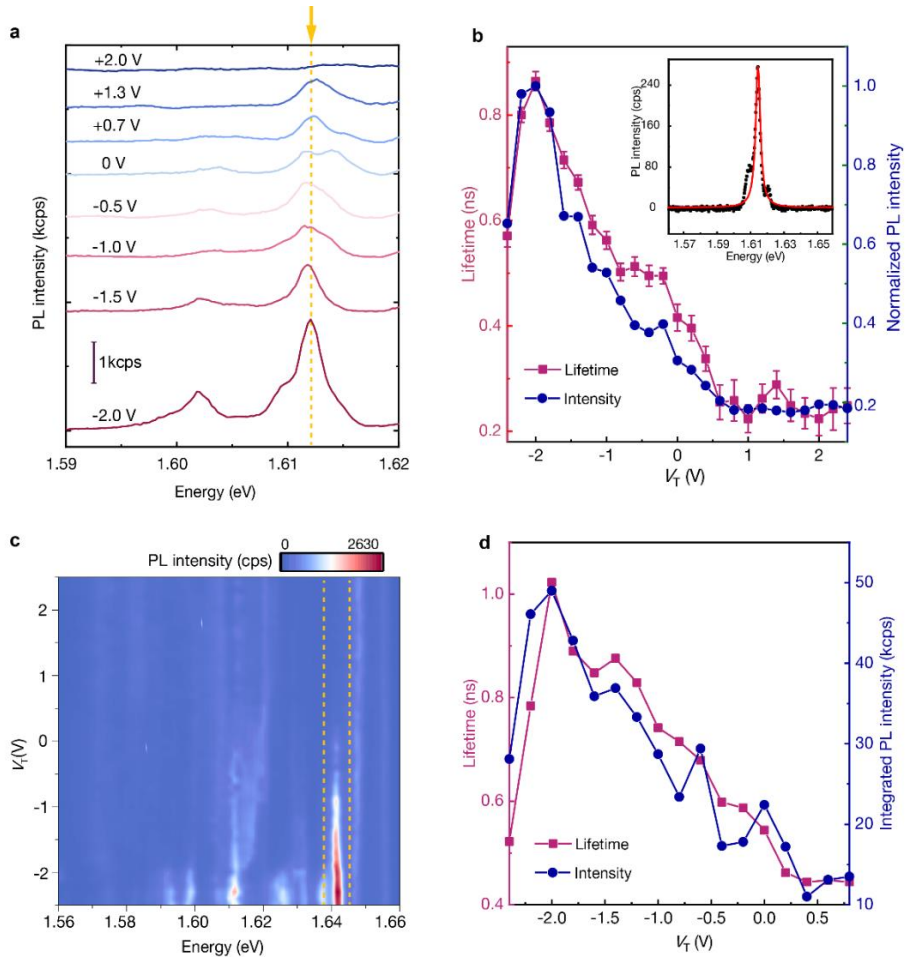


Extended Data Figure 3 | Gate dependence of plasmonic resonance. a, PL spectrum of the WSe_2 monolayer on a nanogap array under white light illumination. The ΔPL spectrum is defined as $\Delta PL = \left(I_{PL}^{nanogap} - I_{PL}^{SiO_2}\right) \Big/ \left(I_{PL}^{nanogap} + I_{PL}^{SiO_2}\right)$, where $I_{PL}^{nanogap(SiO_2)}$ is the PL spectrum at the nanogap (SiO_2) area. **b** and **c,** ΔPL spectrum as a function of the top gate and back gate, respectively.

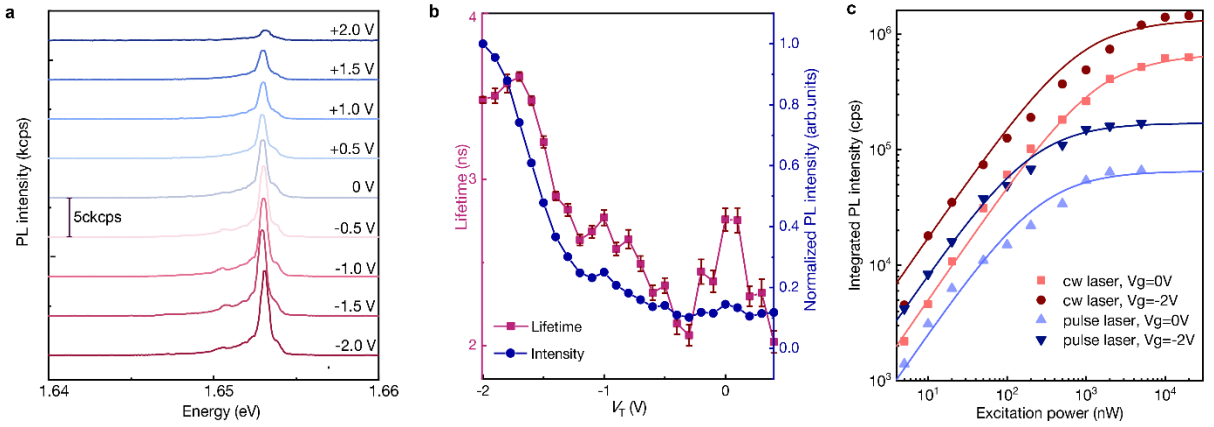


Extended Data Figure 4 | Gate-dependent time-resolved of SPEs on the nanogap array. The time-resolved PL at $V_B = 0\,V$ and $V_T = -1\,V$ (red), 0 V(blue), and 1 V (green) are shown. The symbols represent the measured

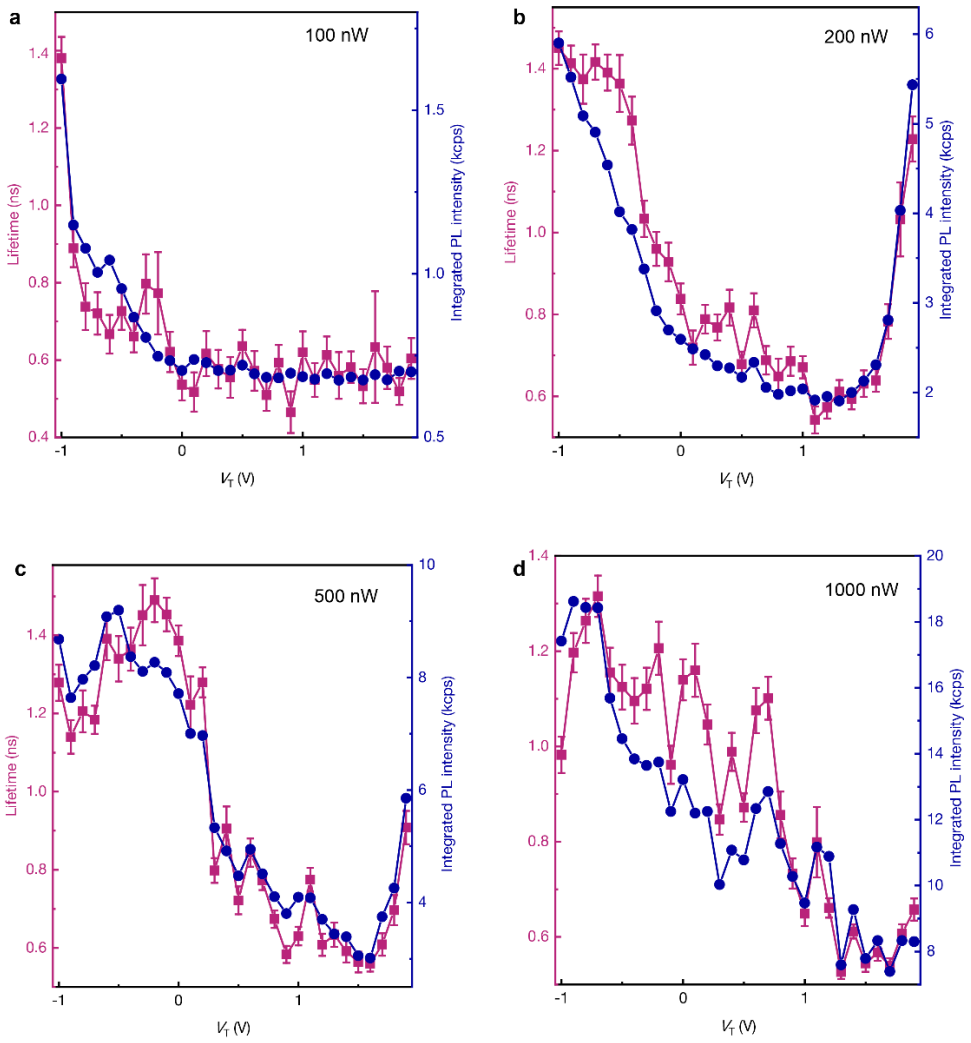
data, and the lines represent the fitting results. The normalization is with respect to the maximum of each time-resolved PL data.



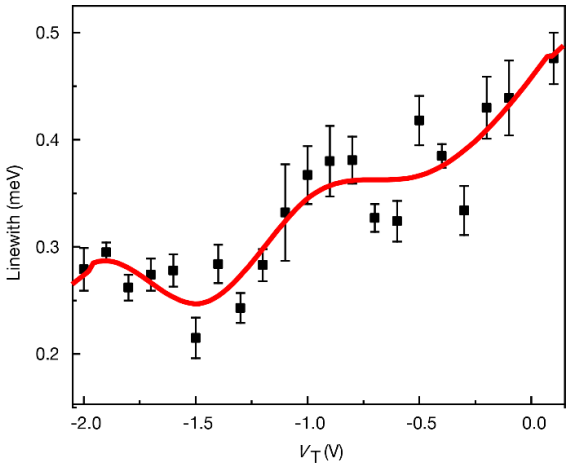
Extended Data Figure 5 | Gate tunable lifetime and PL intensity of two other SPEs in WSe_2 monolayer (device #1). a, Gate-dependent PL spectra under 726 nm CW excitation (power: 300 nW). The orange arrow and the dashed line indicate the filtered peak studied in b. **b,** Gate-dependent PL lifetime and normalized integrated PL intensity of filtered emission under 726 nm pulse excitation (power 300 nW). Inset: filtered PL spectrum at $V_T = V_B = 0$ V under 726 nm CW excitation (power 300 nW). The red line shows the Lorentzian fitting of the dominant peak. **c-d,** Similar to a and b but from another SPE under 726 nm CW excitation (power: 400 nW) and 726 nm pulse excitation (power: 400 nW). The dashed orange box in c indicates the filtered peak studied in d. In all studied SPEs, the PL lifetime and intensity have a qualitatively similar shape, indicating a gate-modulated nonradiative decay process.



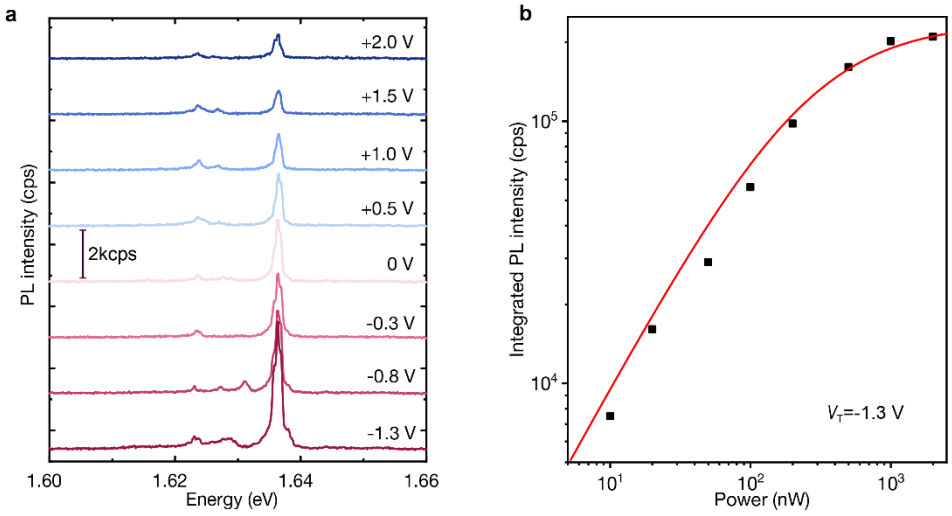
Extended Data Figure 6 | Gate-dependent PL and dynamic characterization of SPEs on nanogaps in another WSe_2 monolayer device (device #2). a, Gate-dependent PL spectra under 726 nm CW excitation (power: 300 nW). **b,** Gate-dependent PL lifetime and normalized integrated PL intensity of filtered emission under 726 nm pulse excitation. The PL lifetime and intensity have a qualitatively similar shape, indicating a gate-modulated nonradiative decay process. **c,** Power dependence for the same SPE at different gate voltages. The lines are fitting with the power saturation function (see Supplementary Information Section 1). The emitter is excited by a 726 nm CW laser and a 726 nm pulse laser with 80 MHz repetition rate, respectively. The data is taken from emitter E1 in device #2. The optical microscopy image and the structure of device #2 are shown in Supplementary Information Fig. S7.



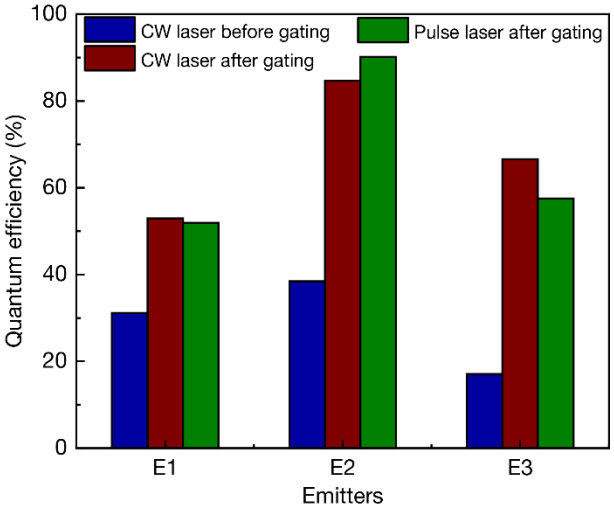
Extended Data Figure 7 | Power and gate dependence of SPE in WSe_2 monolayer (device #1). a-d, Integrated PL intensity and lifetime modulation by the gate voltages at **a,** 100 nW, **b,** 200 nW, **c,** 500 nW and **d,** 1000 nW excitation power. The proportionality between intensity and lifetime is observed in both low and high excitation power. The data is taken from emitter E2 in device #1. The gate-dependent PL spectra at various powers can be found in Supplementary Information Fig. S8.



Extended Data Figure 8 | Gate dependence of SPE linewidth. The full width at half maximum (FWHM) of the chosen peak in Extended Data Fig. 5a is plotted against the applied top gate voltage. The red line is a guide for the eye. The reduction of linewidth indicates suppression of blinking due to nonradiative decay pathways.



Extended Data Figure 9 | Electrically-gated SPE with near-unity quantum efficiency. a, Gate-dependent spectrum. A 726 nm 1000 nW pulse laser excitation is used. The emission at 1.636 eV is filtered to be studied further in b. **b,** Power saturation curve at $V_T = -1.3\,V$. The saturation count is ~216 kcps, corresponding to a quantum efficiency of about 90% under pulse excitation. The data is taken from emitter E2 in device #1.



410

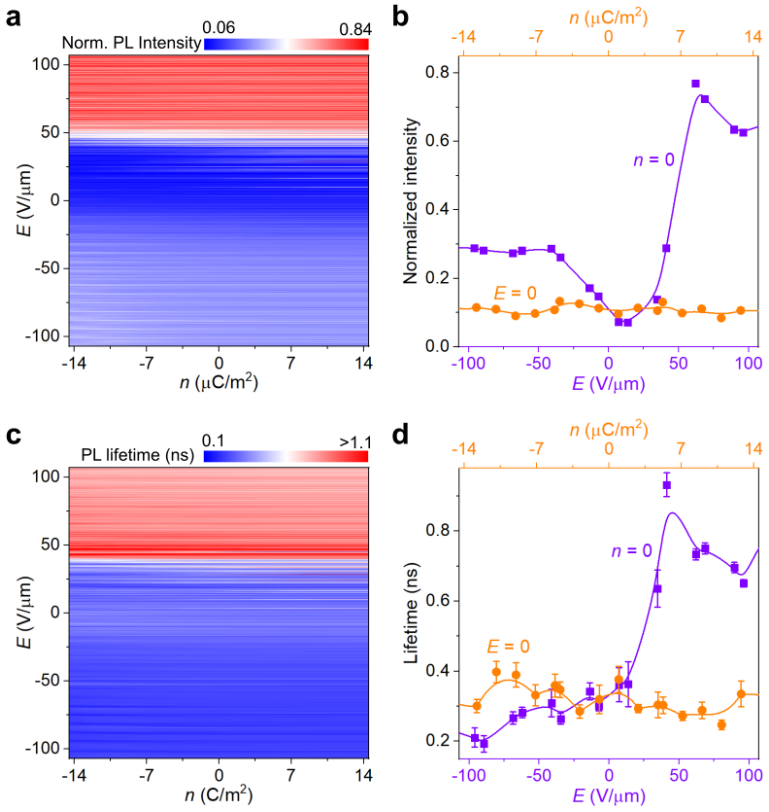
411 **Extended Data Figure 10 | Electrically-gated SPE quantum efficiency.** Representative results from three

412 different SPEs are shown. The "before gating" refers to $V_T = 0$ V case, and "after gating" refers to the case where

413 V_T is applied to reach the maximum count. The value of $V_B = 0$ V is used. Emitter E2 and E3 are in device #1,

414 while emitter E1 is in device #2. The data for emitter E3 is shown in the main text Fig. 3

415

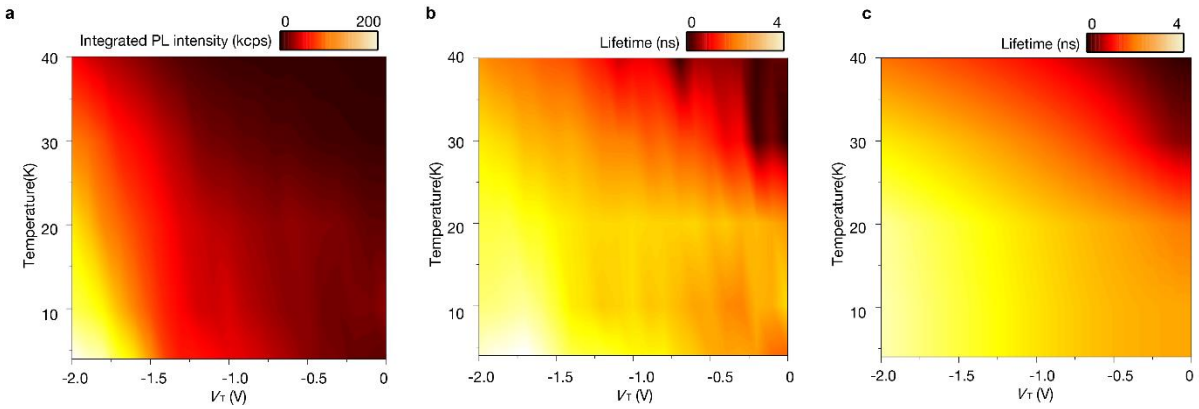


416

417 **Extended Data Figure 11 | Electric field and doping dependence of the PL intensity and lifetime. a,** Gate-

418 dependent map of PL intensity. **b,** Line cuts of the map along constant doping ($n = 0\,\mu C/m^2$, purple line and

symbols) and constant field ($E = 0\,\mathrm{V}/\mu\mathrm{m}$, orange line and symbols). Field-induced change can be clearly observed, while doping-induced change is negligible. **c, d** Like a and b but for gate-dependent PL lifetime. The similarity between the lifetime and the PL intensity maps indicates that gates modulate the nonradiative decay rate. The E and n represent the gate-induced electric field and doping, respectively (see Supplementary Information Section 3 for more details). The intensity and lifetime data are taken from the intrinsic region of Fig. 4(a, b) in the main text. In b and d, the lines are guides for the eye.



Extended Data Figure 12 | Temperature and gate dependence of SPE in WSe_2 monolayer (device #2). a, Integrated PL intensity and **b,** lifetime modulation by the gate voltages at different temperatures under 500 nW 726 nm pulse excitation. The proportionality between intensity and lifetime is maintained until 40 K. **c**, Fitting result with the theoretical model including Auger recombination, described in Supplementary Information Section 5. The gate-dependent PL spectra at various temperatures can be found in Supplementary Information Fig. S9.

References

1 Aharonovich, I., Englund, D. & Toth, M. Solid-state single-photon emitters. *Nat. Photonics* **10**, 631-641 (2016).

2 Chakraborty, C., Kinnischtzke, L., Goodfellow, K. M., Beams, R. & Vamivakas, A. N. Voltage-controlled quantum light from an atomically thin semiconductor. *Nat. Nanotechnol.* **10**, 507-511 (2015).

3 Tonndorf, P. *et al.* Single-photon emission from localized excitons in an atomically thin semiconductor. *Optica* **2**, 347-352 (2015).

4 Koperski, M. *et al.* Single photon emitters in exfoliated WSe_2 structures. *Nat. Nanotechnol.* **10**, 503-506 (2015).

5 Srivastava, A. *et al.* Optically active quantum dots in monolayer WSe_2. *Nat. Nanotechnol.* **10**, 491-496 (2015).

6 He, Y.-M. *et al.* Single quantum emitters in monolayer semiconductors. *Nat. Nanotechnol.* **10**, 497-502 (2015).

7 Parto, K., Azzam, S. I., Banerjee, K. & Moody, G. Defect and strain engineering of monolayer WSe_2 enables site-controlled single-photon emission up to 150 K. *Nat. Commun.* **12**, 3585 (2021).

8 Palacios-Berraquero, C. *et al.* Large-scale quantum-emitter arrays in atomically thin semiconductors. *Nat. Commun.* **8**, 15093 (2017).

9 Branny, A., Kumar, S., Proux, R. & Gerardot, B. D. Deterministic strain-induced arrays of quantum emitters in a two-dimensional semiconductor. *Nat. Commun.* **8**, 15053 (2017).

10 Kern, J. *et al.* Nanoscale positioning of single-photon emitters in atomically thin WSe_2. *Adv. Mater.* **28**, 7101-7105 (2016).

11 Shepard, G. D. *et al.* Nanobubble induced formation of quantum emitters in monolayer semiconductors. *2D Mater.* **4**, 021019 (2017).

12 Kumar, S., Kaczmarczyk, A. & Gerardot, B. D. Strain-induced spatial and spectral isolation of quantum emitters in mono- and bilayer WSe_2. *Nano Lett.* **15**, 7567-7573 (2015).

13 Cai, T. *et al.* Radiative enhancement of single quantum emitters in WSe_2 monolayers using site-controlled metallic nanopillars. *ACS Photonics* **5**, 3466-3471 (2018).

14 Luo, Y. *et al.* Deterministic coupling of site-controlled quantum emitters in monolayer WSe_2 to plasmonic nanocavities. *Nat. Nanotechnol.* **13**, 1137-1142 (2018).

15 Iff, O. *et al.* Deterministic coupling of quantum emitters in WSe_2 monolayers to plasmonic nanocavities. *Opt. Express* **26**, 25944-25951 (2018).

16 Luo, Y., Liu, N., Li, X., Hone, J. C. & Strauf, S. Single photon emission in WSe_2 up 160 K by quantum yield control. *2D Mater.* **6**, 035017 (2019).

17 Dutta, S. *et al.* Coupling quantum emitters in WSe_2 monolayers to a metal-insulator-metal waveguide. *Appl. Phys. Lett.* **113**, 191105 (2018).

18 Blauth, M. *et al.* Coupling single photons from discrete quantum emitters in WSe_2 to lithographically defined plasmonic slot waveguides. *Nano Lett.* **18**, 6812-6819 (2018).

19 Lien, D.-H. *et al.* Electrical suppression of all nonradiative recombination pathways in monolayer semiconductors. *Science* **364**, 468-471 (2019).

20 Li, M., Biswas, S., Hail, C. U. & Atwater, H. A. Refractive index modulation in monolayer molybdenum diselenide. *Nano Lett.* **21**, 7602-7608 (2021).

21 Chen, P. *et al.* Approaching the intrinsic exciton physics limit in two-dimensional semiconductor diodes. *Nature* **599**, 404-410 (2021).

22 Hötger, A. *et al.* Gate-switchable Arrays of quantum light emitters in contacted monolayer MoS_2 van der Waals heterodevices. *Nano Lett.* **21**, 1040-1046 (2021).

23 Anderson, C. P. *et al.* Electrical and optical control of single spins integrated in scalable semiconductor devices. *Science* **366**, 1225-1230 (2019).

24 Doi, Y. *et al.* Deterministic electrical charge-state initialization of single nitrogen-vacancy center in diamond. *Phys. Rev. X* **4**, 011057 (2014).

25 Wang, Z. *et al.* Selectively plasmon-enhanced second-harmonic generation from monolayer tungsten diselenide on flexible substrates. *ACS Nano* **12**, 1859-1867 (2018).

26 Wang, Z. *et al.* Giant photoluminescence enhancement in tungsten-diselenide–gold plasmonic hybrid structures. *Nat. Commun.* **7**, 11283 (2016).

27 Fuchs, F. *et al.* Engineering near-infrared single-photon emitters with optically active spins in ultrapure silicon carbide. *Nat. Commun.* **6**, 7578 (2015).

28 Wang, J. *et al.* Bright room temperature single photon source at telecom range in cubic silicon carbide. *Nat. Commun.* **9**, 4106 (2018).

29 Scavuzzo, A. *et al.* Electrically tunable quantum emitters in an ultrathin graphene–hexagonal boron nitride van der Waals heterostructure. *Appl. Phys. Lett.* **114**, 062104 (2019).

30 Fisher, B. R., Eisler, H.-J., Stott, N. E. & Bawendi, M. G. Emission intensity dependence and single-exponential behavior in single colloidal quantum dot fluorescence lifetimes. *J. Phys. Chem. B* **108**, 143-148 (2004).
31 Schlegel, G., Bohnenberger, J., Potapova, I. & Mews, A. Fluorescence decay time of single semiconductor nanocrystals. *Phys. Rev. Lett.* **88**, 137401 (2002).
32 Sun, D. *et al.* Observation of rapid exciton–exciton annihilation in monolayer molybdenum disulfide. *Nano Lett.* **14**, 5625-5629 (2014).
33 Chen, Y. *et al.* Rapid focused ion beam milling based fabrication of plasmonic nanoparticles and assemblies via "sketch and peel" strategy. *ACS Nano* **10**, 11228-11236 (2016).