Electronic Supplementary Material (ESI) for Nanoscale Horizons.

Supplementary Information:

An ultrafast photodetector driven by interlayer exciton dissociation in a van der Waals heterostructure

Edoardo Lopriore[1,3*], Enrique G. Marin[2], and Gianluca Fiori[1]

[1]*Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy*

[2]*Departamento de Electrónica y Tecnología de Computadores, Universidad de Granada, Granada, Spain*

[3]*currently at Electrical Engineering Institute and Material Science Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

[*]*edoardo.lopriore@epfl.ch*

Supplementary Figures

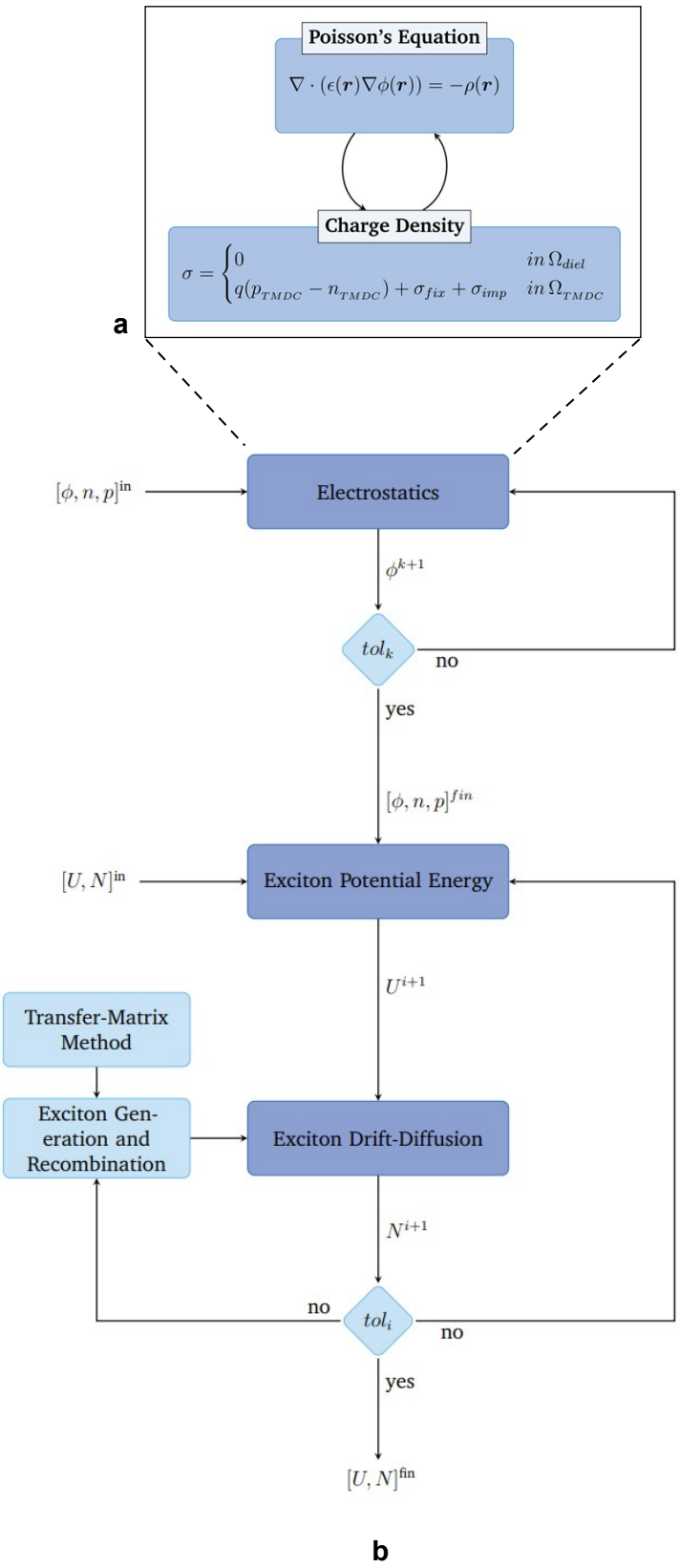
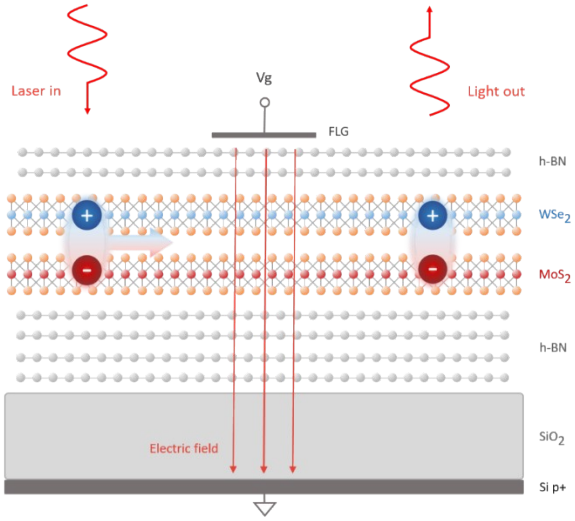
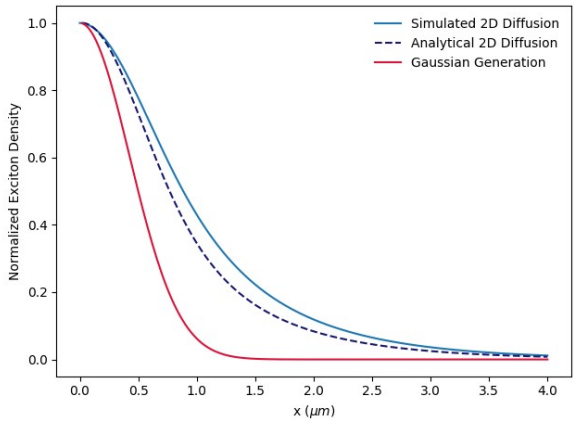


Figure S1 | Gummel map for interlayer exciton transport. a Schematic representation of the self-consistent loop included in the simulation platform. The charge density is null in the

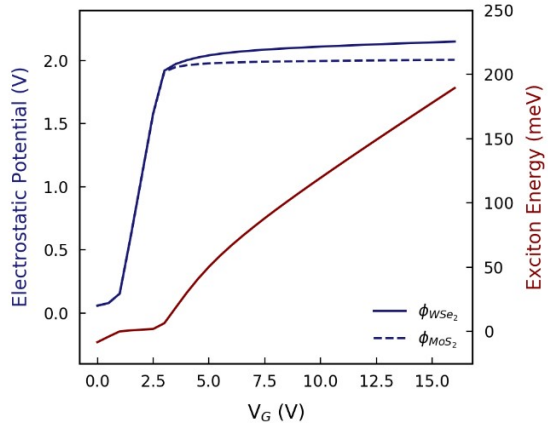
discretization domains belonging to dielectrics ($^{\Omega_{diel}}$). In TMDCs ($^{\Omega_{TMDC}}$), the 2D-confined electron and hole densities are considered. Fixed charge ($^{\sigma_{fix}}$) and ionized impurity ($^{\sigma_{imp}}$) densities can be included in the model when required. **b** The iterative solutions for electrostatics and interlayer exciton drift-diffusion are separated. The two iterative schemes end when the electrostatic potential and the exciton density fall within the respective tolerance interval. The optical transfer-matrix method described in Supplementary Note 1 provides the interlayer exciton generation term for all drift-diffusion calculations.



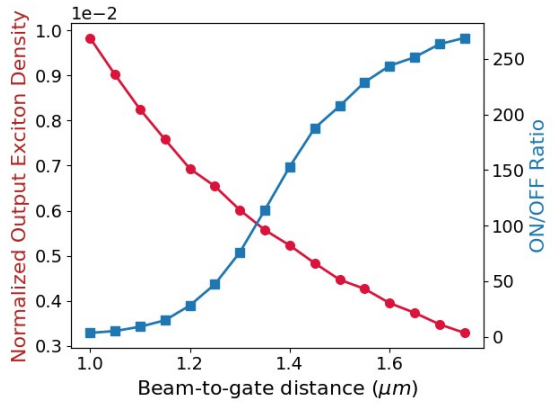
a



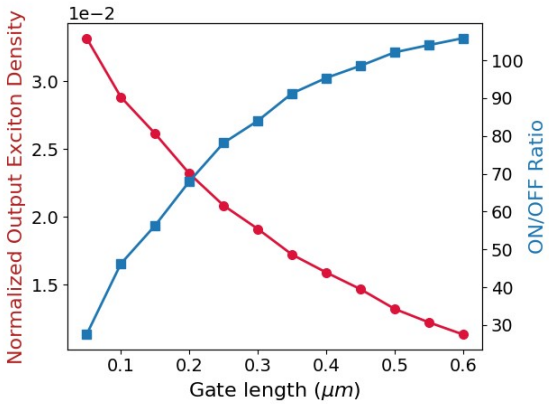
b



c



d



e

Figure S2 | Excitonic transistor engineering. a Holes and electrons are spatially separated in the corresponding layers on ultrafast time scales. The generated exciton population diffuses across the gate region. A voltage applied to the top gate induces a vertical electric field and raises a exciton potential barrier, thus modulating the output interlayer exciton density. **b** Comparison between the numerical (solid blue line) and analytical (dashed blue

line) solutions to the two-dimensional low-density free diffusion of interlayer excitons in the excitonic transistor. A gaussian generation factor is considered (red line). **c** Electrostatic potential on the TMDs and excitonic energy in the excitonic transistor as a function of the applied gate voltage. **d** ON/OFF ratio (blue) and normalized output interlayer exciton density in the ON state (red) for beam-to-gate distances between $1.0\ \mu m$ and $1.75\ \mu m$. The gate length is fixed ($1.1\ \mu m$). **e** Normalized N (red) and ON/OFF ratio for gate lengths between $50\ nm$ and $600\ nm$. The beam-to-gate distance is fixed ($1.35\ \mu m$).

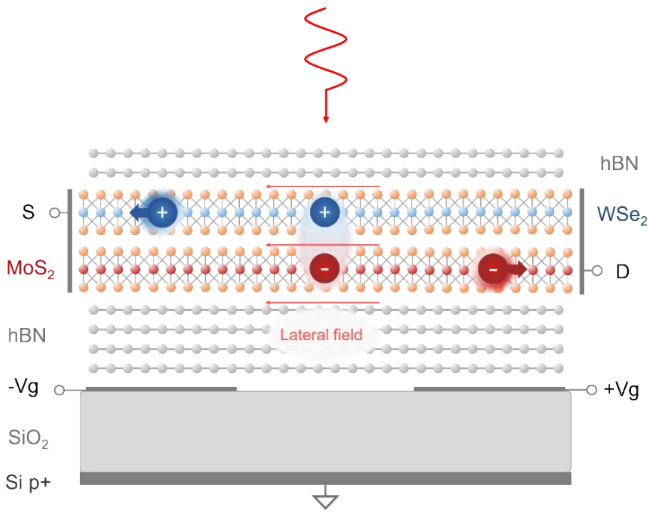
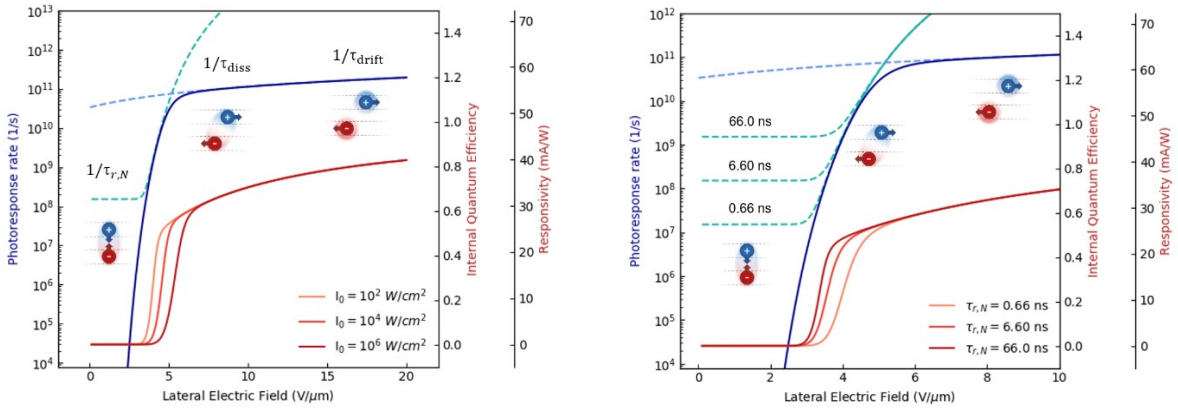


Figure S3 | Bottom-split-gate detector. Few-layer graphene electrodes are located on the bottom of the van der Waals heterostructure. All material properties and dimensions are identical to those of the device presented in the main text (Fig. 1a).



a **b**

Figure S4 | Analytical detector performance. Both the single-exciton decay rate (dashed green line) and the free-charge lifetime (dashed blue line) contribute to the photoresponse rate of the split-gate detector (solid blue line). The photoresponse rate is calculated relatively to the

low-density regime. The graphical representations of electrons and holes refer to the three main parts of the plot: low, low-to-moderate and high lateral fields, where the dominant physical phenomena are exciton recombination, dissociation and free-charge drift respectively. The responsivity and internal quantum efficiency of a hBN-encapsulated WSe_2/MoS_2 heterobilayer are reported in red as a function of lateral electric field in the simplified condition of uniform exciton generation. **a** Higher optical intensity results in a shift in responsivity curves due to the nonlinear term in the recombination factor. **b** In the low density regime, higher lifetimes correspond to different evolutions of the exciton decay rates. Thus, the responsivity curves exhibit lower threshold electric fields for higher recombination factors.

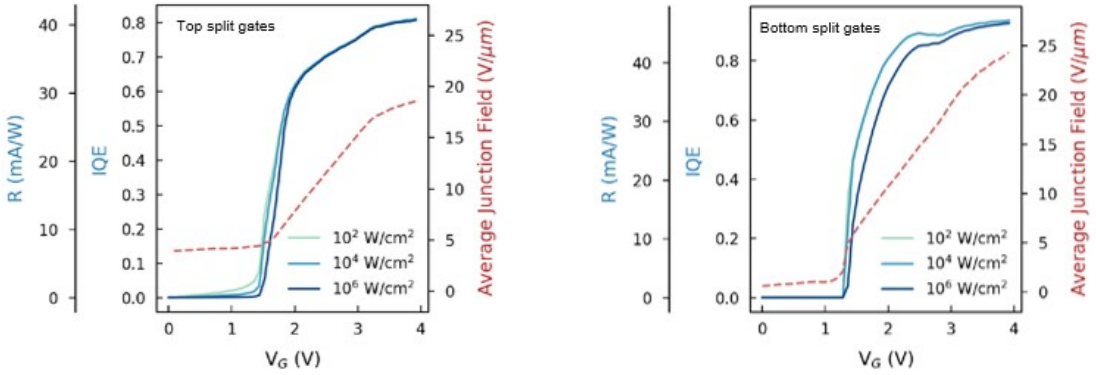


Figure S5 | Ultra-narrow laser beam. Responsivity and IQE in the top (**a**) and bottom (**b**) split-gate detectors as a function of the applied voltage for an unrealistic laser beam width equal to the nominal junction length. In particular, $2w_0 = L_{nom} = 100\ nm$, where w_0 is the laser beam waist radius.

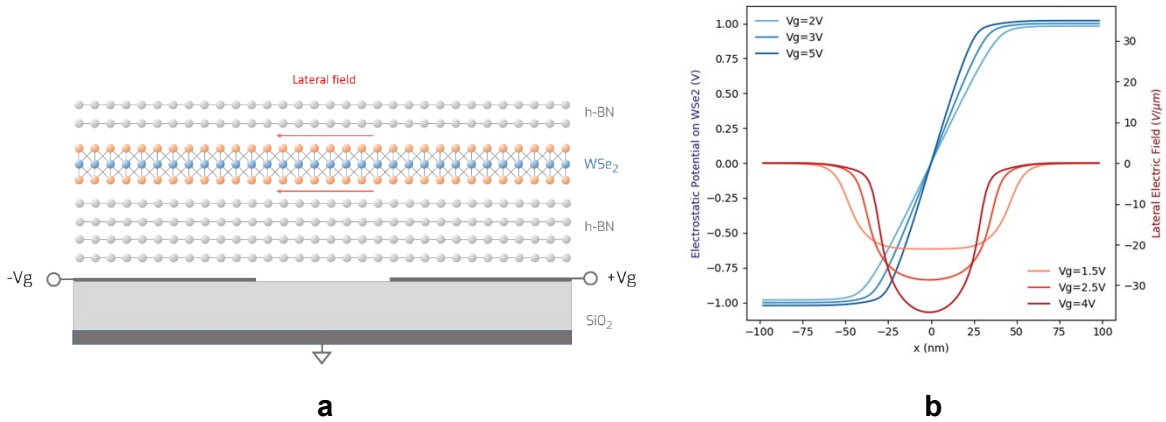


Figure S6 | Monolayer WSe_2 *p-n* junction. a Split-gate device based on an hBN-encapsulated monolayer WSe_2. Electrostatic *p-n* lateral doping of the TMDC layer induces a lateral junction electric field. The nominal junction length is $100\ nm$. **b** Electrostatic potential on WSe_2 (blue)

and lateral electric field (red) for different split-gate voltages in the structure involving a single monolayer TMDC.

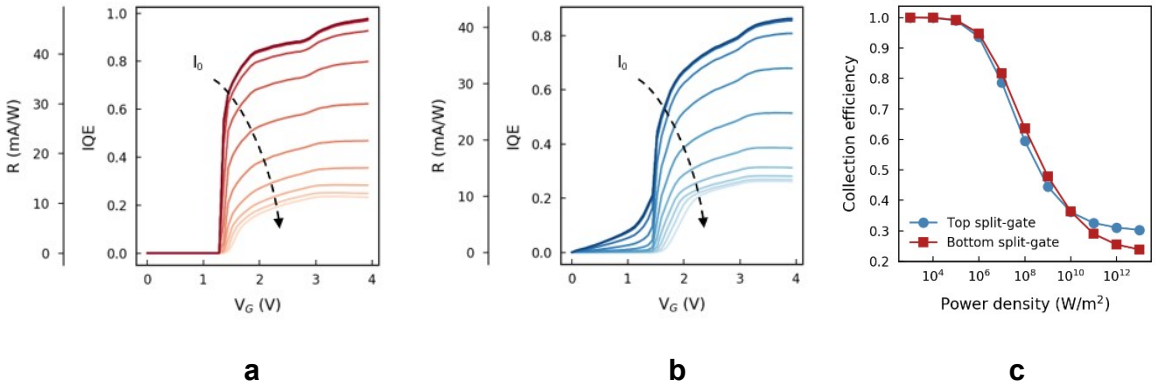


Figure S7 | Laser beam intensity and collection efficiency. For growing optical power densities, the non-linear components in the recombination factor become increasingly relevant and prevent interlayer excitons generated outside of the *p-n* junction to reach the high-field region and contribute to the photocurrent. **a,b** Responsivity curves as a function of split-gate voltages for optical beam intensities between $10^3\ W/m^2$ and $10^{13}\ W/m^2$ in the top (**a**) and bottom (**b**) split-gate photodetectors. **c** Conversion efficiencies of both geometries as a function of laser power density. In all simulations, the laser beam diameter and the nominal junction length are fixed at 820 nm and 100 nm respectively.

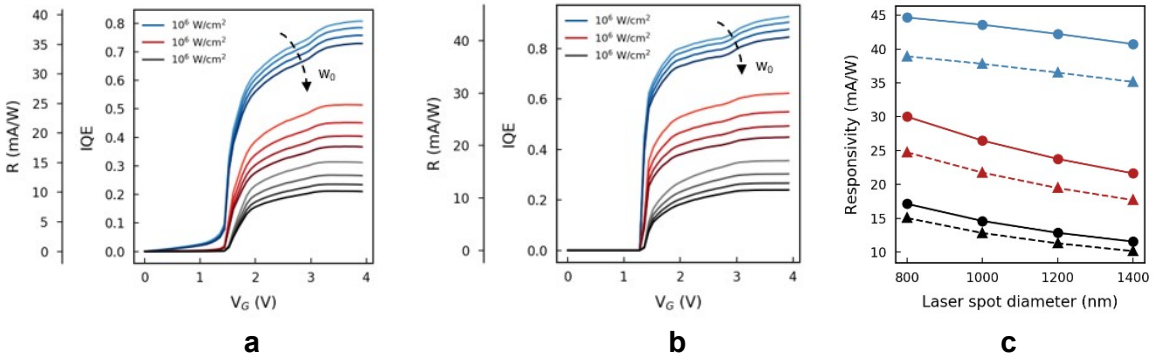


Figure S8 | Laser spot size. a,b Variation of the responsivity curves in the top (**a**) and bottom (**b**) split-gate detectors for laser beam diameters (^{2w_0}) between 800 nm to 1400 nm. We consider optical power intensities of 10^6 W/m^2 (blue), 10^8 W/m^2 (red) and 10^{10} W/m^2 (black). **c** Maximum responsivity as a function of laser spot size for the top (dashed lines) and bottom (solid lines) split-gate geometries.

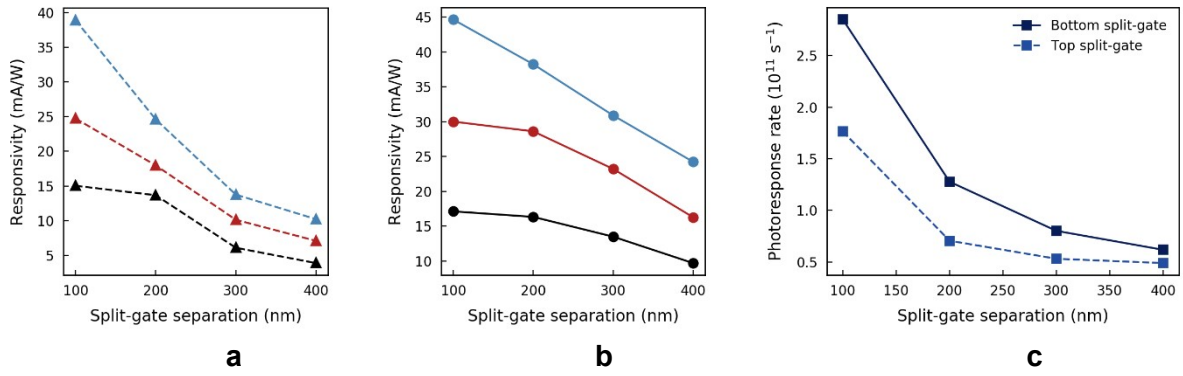


Figure S9 | Split-gate separation. a,b Maximum responsivity as a function of the split-gate distance in the top (**a**) and bottom (**b**) configurations. **c** Maximum photoresponse rate as a function of the nominal junction length for the two geometries.

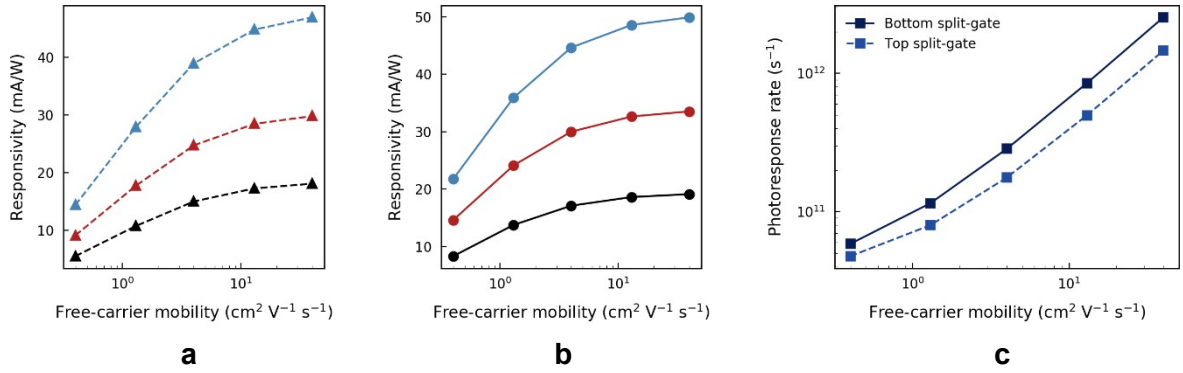


Figure S10 | Free-carrier mobility. a,b Maximum responsivity as a function of the free-carrier mobility in the top (**a**) and bottom (**b**) split-gate photodetectors. **c** Maximum photoresponse rate as a function of the free-carrier mobility for the two geometries.

Supplementary Note 1: Optical Input

We simulate an optical input source focused on the device in order to generate a population of interlayer excitons in the van der Waals heterostructure. We use a transfer-matrix method to study the propagation of electromagnetic waves through the layered media, assuming isotropic and homogeneous parallel layers and an electromagnetic field propagating perpendicularly to the structure itself. Different refractive indexes as a function of wavelength are considered for all the materials involved, namely graphene[1], hBN[2], WSe_2, MoS_2[3], SiO_2[4] and Si[5]. The time averaged energy dissipated along the z axis inside the j-th layer (Q_j) is obtained from the electric field distribution E_j as[8]:

$$Q_j(z,\lambda) = \frac{1}{2} c\epsilon_0 \alpha_j \eta_j |E_j(z)|^2 \tag{1}$$

and it is a function of both the position inside the layer (z) and the electric field wavelength (λ). The absorption coefficients are defined as $\alpha_j = 4\pi\kappa_j/\lambda$, where κ_j is the extinction coefficient characteristic of the layer. The total exciton generation rate at a certain point of the j-th layer is calculated by integrating the dissipated energy over the wavelength spectrum of the incident light beam:

$$G_j(z) = \eta_{cnv} \int_{\lambda_{min}}^{\lambda_{max}} Q_j(z,\lambda) d \tag{2}$$

with η_{cnv} being the exciton conversion efficiency, representing the probability that an absorbed photon is converted into an excited state. $G_j(z)$ is a volumetric factor, but for both intralayer and interlayer excitons in TMDC monolayers and heterostructures, the generation factor is better defined as the number of excited states induced per second and per unit area. For a single

layer, we can estimate the 2D exciton density generation by integrating the dissipation factor over the thickness of the layer:

$$G_{2D_j} = \eta_{cnv} \int_0^{d_j} \int_{\lambda_{min}}^{\lambda_{max}} Q_j(z,\lambda)\, d\lambda\, dz \tag{3}$$

where we are assuming that the energy absorbed in the whole layer thickness is relevant in determining the induced 2D exciton population.

In the proposed devices, interlayer excitons are generated in a type-II heterojunction of two stacked monolayer TMDCs and excited states from both layers contribute to the interlayer exciton population in the heterobilayer. Thus, the interlayer exciton generation factor G_{IX} is calculated as sum of two terms:

$$G_{IX} = \eta_e\, G_{2D_e} + \eta_h\, G_{2D_h} \tag{4}$$

where η_e (η_h) and G_{2D_e} (G_{2D_h}) are the transfer efficiency and the intralayer generation term for the electron-donor (hole-donor) layer respectively. Since high charge-transfer efficiencies up to 99% have been previously reported in WSe_2/MoS_2 heterobilayers[6] as well as in other structures[7,7–9], we assume η_e and η_h to be approximately equal to one.

The maximum light intensity focused on the device is given by $I_0 = 2P_0/(\pi w_0^2)$, where P_0 is the optical power and w_0 the light beam radius. Considering the relationship between power absorption and electric field in Eq. 1, the interlayer exciton generation factor depends linearly on the focused light intensity. If the light beam is monochromatic, we must solve the transfer-matrix equations only at a given λ_0, and a gaussian distribution of the interlayer exciton generation factor is obtained:

$$G_{IX}(r) = G_0\, exp\left(-\frac{2r^2}{w_0}\right) \tag{5}$$

In the experimental excitonic transistor[10], the input laser optical power is $P_0 = 270\,\mu W$ at a wavelength $\lambda_0 = 647\,nm$, with a gaussian beam of waist radius $w_0 = 0.6\,nm$. The maximum generation achieved with these optical parameters is $G_0 \simeq 1.26 \cdot 10^{22}\, s^{-1} cm^{-2}$, and the resulting interlayer exciton generation factor is the one included in Fig. 2a in the main text.

Supplementary Note 2: Electrostatics

Self-consistent solution. We apply a box-integration method to the device structure in order to obtain a discretization of the Poisson equation. Since all the devices of interest are translationally invariant in one direction (y axis), the two-dimensional Poisson equation is solved. Neumann boundary conditions on the electrostatic potential are set at the borders of the simulation region, except for those points located at the biased gates. The 2D-confined charge distributions in the TMDCs are included so to determine the electrostatic doping induced by the device terminals. Isotropic effective masses for electrons and holes are used for tangential transport with respect to the TMDC layer:

$$n_{TMDC}(x,y) \simeq g_v \frac{m_e^* kT}{\pi\hbar^2} \ln\left[1 + \exp\left(-\frac{E_C(x,y) - E_F}{kT}\right)\right] \tag{6}$$

$$p_{TMDC}(x,y) \simeq g_v \frac{m_h^* kT}{\pi\hbar^2} \ln\left[1 + \exp\left(+\frac{E_V(x,y) - E_F}{kT}\right)\right] \tag{7}$$

where g_v is the valley degeneracy factor of the material of interest, E_F is the Fermi energy of the layer. The energy difference in Eq. 7 is related to the electrostatic potential on the layer $\phi(x,y)$ as:

$$E_C(x,y) - E_F = q\phi(x,y) - \frac{E_g}{2} - (E_F - E_i) \tag{8}$$

where E_g is the band gap energy of the material, E_i is the intrinsic level in the semiconductor, and $\delta E = E_F - E_i$ is an energy variation term. A similar expression can be derived for $E_V(x,y)$. The main term contributing to δE is the charge transfer between stacked TMDC layers. The band alignment of WSe_2/MoS_2 heterobilayers has been investigated through ab-initio calculations by Latini *et al.*[11], from which we extract an energy variation of $0.15\ eV$ for each TMDC. The back gate voltage is kept at 0 V and taken as the reference level in all simulations. The bandgap values for the two TMDCs have been extracted from Latini et al. as well[11], with energy gaps of 2.48 eV and 2.08 eV for MoS_2 and WSe_2, respectively.

The charge distributions in Eq. 7 and Eq. 8 depend on the electrostatic potential along the layer. Therefore, the Poisson equation must be solved self-consistently with the surface charge density in the semiconducting monolayers in order to achieve the exact solution to the electrostatic problem (Figure S1).

Quantum-confined Stark effect in a WSe_2/MoS_2 heterobilayer. TMDC monolayers extend through the whole (x,y) plane and are assumed to be perfectly planar (i.e. limited to a single-point thickness in the vertical direction z). The band alignment between the two TMDCs gives a work function of $WF_{HS} \simeq 4.26\ eV$ for the heterojunction[11]. The heavily p-doped silicon work function is calculated as $WF_{Si\,(p+)} \simeq WF_{Si} + E_g^{Si}/2 = 5.17\ eV$, where $WF_{Si} \simeq 4.6\ eV$ and $E_g^{Si} = 1.14\ eV$. As for few-layer graphene, we use a value of $WF_{FLG} \simeq 4.3\ eV$ (refs. [12,13]).

Next, we extract the average vertical electric field in the heterobilayer $\bar{E}_{HB}$ and calculate the Stark effect component of the exciton potential energy. Figure 2c in the main text shows the exciton potential barrier height as a function of gate voltage in the excitonic transistor. Both layers of the heterostructure are lightly doped for $1\,V \leq V_G \leq 2.5\,V$, so we can develop a simple analytical capacitive model for the relation between $\bar{E}_{HB}$ and V_G:

$$\frac{\partial \bar{E}_{HB}}{\partial V_G} = \frac{\epsilon_d \;\; 1}{\epsilon_{HB} d_{tot}} \tag{9}$$

where $d_{tot} \simeq 301.2\,nm$ is the total distance between the top and bottom gates of the device, and ϵ_{HB} and ϵ_d are the dielectric constants of the heterobilayer and of the surrounding environment, respectively. From Eq. 14, the exciton potential energy U is calculated as:

$$\frac{\partial U}{\partial V_G} = \frac{\epsilon_d \; d_{IX}}{\epsilon_{HB} d_{tot}} \tag{10}$$

where $d_{IX} \simeq 0.6\,nm$ is the interlayer exciton dipole length. This linear relation is represented by a dashed red line in Fig. 2c in the main text, overlapping with our simulation results for gate voltages corresponding to lightly doped TMDCs. For values of V_G corresponding to strong charge accumulation, the simulated exciton potential barrier grows steeper with respect to the lightly doped approximation, evidencing that the exact solution to the Poisson equation is necessary to obtain reliable results in excitonic devices with electrostatic doping.

Supplementary Note 3: Interlayer Exciton Transport

Drift-Diffusion Equation. We assume exciton-exciton interactions to be dominant with respect to exciton − LA-phonon coupling so we can apply the theoretical work on indirect excitons in coupled quantum well systems to interlayer excitons in van der Waals

heterostructures[14]. In our work, we only consider transport in the classical limit at $T \gg T_0$, where T_0 is the degeneracy temperature of the exciton gas[15]. We assume the diffusion coefficient D_x to be constant in space and we suppose room temperature conditions, where the exciton mobility and the diffusion coefficient are linked by the Einstein relation $D_x = \mu_x kT$ (ref. [14]). The diffusion length is defined as $L_D = \sqrt{D_x \tau_{r,N}}$, where $\tau_{r,N}$ is the exciton decay time. Interlayer exciton transport is computed through the drift-diffusion equation:

$$\frac{\partial N}{\partial t} = \nabla \cdot \left[D_x N + \mu_x \nabla U \right] + G - R \tag{11}$$

where N is the interlayer exciton density, G and R are the generation and recombination factors, and U is the exciton potential energy. The recombination rate is calculated as $R = N/\tau_N + \gamma N^2$, where τ_N is the exciton lifetime and γ represents a second-order term comprising mechanisms such as exciton-exciton annihilation (EEA) or Auger-type processes, which were demonstrated to be relevant for both intralayer[16–18] and interlayer[19,20] exciton dynamics at high beam intensities. The layers in the experimental work of reference[10] were stacked with crystalline axes near 0° or 60°. Since the twist angle in WSe_2/MoS_2 heterobilayers has been proven not to be relevant for interlayer exciton properties[8,21], we neglect the effects of an unexpected misalignment between the two layers. The transport parameters for interlayer excitons in the excitonic transistor are determined from free diffusion simulations (Fig. 2a in the main text) as $\tau_{r,N} = 6.6$ ns, $L_D = 1.0$ μm and $\gamma = 0.08$ cm^2 s^{-1}, and are used in every device model in this work. We note that these parameters are comparable with other works on interlayer exciton transport[10,19,20,22]. Nonetheless, they are far from the best performance obtained for interlayer exciton transport in other structures and devices[23,24]. We also note that by considering experimentally extracted parameters, we are inherently working on a platform

which is taking into account the effects of defects and other non-idealities on the excitonic species, without making any assumptions on the nature and quality of the materials as well as their assembly.

The two main components in the exciton potential energy U are the quantum-confined Stark effect (QCSE) and the nonlinear term given by exciton-exciton interaction (EEI). To address exciton-exciton interaction of indirect states in van der Waals heterostructures, a mean-field approximation is used. Following previous works on coupled quantum wells[25,26], the position-dependent interaction energy between excitons is given by:

$$u(r) = \frac{e^2}{4\pi\epsilon_0\epsilon_{HS}}\left(\frac{2}{r} - \frac{2}{\sqrt{r^2 + d^2}}\right) \tag{12}$$

where ϵ_{HS} is the dielectric constant of the heterostructure and d is the exciton dipole length, which is approximately equal to the distance between the heterojunction layers (~ 0.6 nm)[23,24]. The interaction energy is approximated to $\left(e^2d^2\right)/\left(4\pi\epsilon_0\epsilon_{HS}r^3\right)$ at relatively large distances. Integrating over space, the parallel plate capacitor formula develops:

$$u_0 \simeq \frac{de^2}{\epsilon_0\epsilon_{HS}} \tag{13}$$

The interlayer exciton potential energy U is then calculated as:

$$U = ed\,\overline{E_{HB}} + u_0N$$

We extract the Stark shift component in the exciton potential from the average vertical field in the TMDC heterobilayer ($\overline{E_{HB}}$) obtained from the electrostatic solution.

Since the exciton transport problem is two-dimensional, the points of interest are (x,y) in $\Omega_{HS} \subset R^2$, where Ω_{HS} is the bounded, convex domain representing the heterobilayer. Since

interlayer excitons are forced to dissociate or recombine on the heterostructure edges, Dirichlet boundary conditions are applied to the interlayer exciton density on the points belonging to the border of the heterobilayer ($^{\partial\Omega_{HS}}$). As well as for electrostatics (Supplementary Note 2), we use a finite-difference approach based on a rectangular grid to discretize the interlayer exciton drift-diffusion equation. The implementation is first carried out on the x axis, in a one-dimensional problem, and then extended to the full (x,y) plane.

Modified Gummel map. All numerical methods are treated within the same software platform, following a Gummel-map scheme specifically adapted to interlayer excitons. The decoupled Gummel map has been introduced as an alternative to the fully coupled Newthon-Raphson method for semiclassical solutions to electron and hole transport in electronic devices[27]. In fact, it is more insensitive with respect to initial variable guesses and is preferable in terms of both computational burden and memory occupation. The main difference between electrical and interlayer exciton transport is that the exciton energy is not dependent on the electron and hole densities, but only on the vertical electric field and on the interlayer exciton density itself. Thus, the solution procedure for interlayer exciton transport is characterized by an inherently decoupled structure. As a result, the iterative solution scheme is mainly based on corrective terms in the exciton potential and in the recombination factor (Figure S1).

The structure of the decoupled Gummel map for interlayer excitons is as follows. First, we apply the Newton-Raphson method to the nonlinear Poisson equation, which is solved self-consistently with the electron and hole charge densities in the TMDC layers. From the electrostatic solution, we extract the electric field distribution in the device, and we calculate the Stark shift energy U^0. With the initial guess for interlayer exciton density N^{in}, we obtain the excitonic potential and recombination term along the heterobilayer region. The exciton

drift-diffusion equation is solved and the new exciton density term N^{i+1} is inserted in the next iteration. When convergence is achieved, we obtain the final interlayer exciton density distribution N^{fin} and potential energy U^{fin} in the heterobilayer.

Low-density regime. We have exploited a simplified structure to only consider interlayer exciton diffusion and compare the simulation results with an analytical model, which is commonly used in the literature for both intralayer[28] and interlayer exciton diffusion[23,29]. We assume a low-density regime ($N \leq 1 \times 10^{10}\ cm^{-2}$) to neglect the nonlinear terms in both the recombination factor and the exciton potential energy, giving $R \simeq N/\tau_N$ and $U \simeq 0$. In these conditions, the steady-state exciton drift-diffusion equation simplifies to:

$$G(r) = \frac{N(r)}{\tau_N} + D_x \Delta N(r) \tag{14}$$

We solve the low-density diffusion in two dimensions considering a gaussian generation profile and using the transport parameters derived in the main text. An analytical solution to Eq. 23 is given by the convolution between the modified Bessel function of the second kind K_0 and the laser beam gaussian profile (Figure S2b)[28]:

$$N(r) \propto \int_{-\infty}^{+\infty} K_0\left(\frac{r}{L_x}\right) e^{-(r-r')^2/w_0^2}\, dr' \tag{15}$$

Supplementary Note 4: Excitonic transistor engineering

The structure of the excitonic transistor is reported in Figure S2a. The distance between the laser beam focus and the nearest border of the gate region is $d_{B-G} = 1.35\ \mu m$ for the results

in the main text (Fig. 2). We apply the same simulation procedure to beam-to-gate distances varying between $1.0\ \mu m$ and $1.75\ \mu m$. The maximum normalized ON output density and the ON/OFF ratio are reported for 16 d_{B-G} values in Figure S2d. For $d_{B-G} < 1.0\ \mu m$, the exciton density modulation is ineffective, and the ON/OFF ratio is negligible. In fact, for low distances the density in the gate region is too high to be blocked by the potential barrier and a negligible change in output population is detected. For $d_{B-G} > 1.75\ \mu m$, the ON/OFF ratio saturates to a value slightly under 270, which is more than twice what was obtained in the experimental excitonic transistor. Instead, the normalized output density in the ON state decreases monotonically through the whole d_{B-G} range due to the growing diffusion length that the excitons must travel before they reach the output region of the device. Thus, a trade-off between the two parameters of interest is present. We note that $d_{B-G} = 1.35\ \mu m$ maximizes both parameters if equal importance is given to both. This finding is consistent the experimental work by Unuchek *et al*[10], where the maximum output photoluminescence was obtained with this beam-to-gate distance.

A similar approach is pursued for the gate length l_G. In the simulations of Figure S2e, the beam-to-gate distance is kept at the experimental value (1.35 μm), while l_G is varied between 50 nm and 600 nm. We note that these values are significantly lower than the gate length of the results reported in the main text (1.1 μm). In the range of interest, the normalized density stays between 1×10^{-2} and 3.5×10^{-2}, which is more than 6 times greater than the value reported in the main text (Fig. 2a). For $l_G \geq 600$ nm, the ON/OFF ratio saturates to a value around 110, which is compatible with what was obtained for $l_g = 1.1$ μm (Fig. 2c in the main text). Lower gate lengths give a reduced ON/OFF ratio, since the barrier becomes thinner and less effective with respect to exciton density modulation. As an example, with $l_G = 100\ nm$

the ON/OFF ratio is approximately 45, while the output exciton density in the ON state is more than 5 times higher than in the experiment conditions, thus limiting modulation losses.

Therefore, geometrical parameters provide a viable way of tuning the performance of the excitonic transistor. Our results show that the excitonic transistor is highly scalable, with boundaries on geometrical parameters dictated by exciton diffusion itself. It is worth noting that the developed modeling approach can be applied to novel excitonic devices and circuits[30], as well as future optical interconnects based on interlayer excitons in van der Waals heterostructures[31].

Supplementary Note 5: Analytical photoresponse

In order to gain rapid insights into the operation of the heterobilayer photodetectors, we consider a simplified device operation to extract an analytical approximation of the responsivity and IQE. In this simplified model, the generation rate is assumed to be constant along x. Thus, the exciton density and potential energy are also uniform, which makes the exciton density flux J_x null, and only generation-recombination mechanisms are relevant for the excitonic response. In these conditions, a 1D rate equation for excitons develops:

$$\frac{dN}{dt} = G - \frac{N}{\tau_{r,N}} - \frac{N}{\tau_{diss}} - \gamma N^2 = G - \frac{N}{\tau_N} - \gamma N^2 \tag{16}$$

where γ is described in Supplementary Note 3. Kamban and Pedersen recently derived an approximating expression for interlayer exciton dissociation which is valid for low and moderate lateral fields[32]:

$$\Gamma_{inter} \simeq \Gamma_0 E_{Lat}^{1/2-2\sqrt{\mu_X}/(\kappa k)} \exp\left(-\frac{2\sqrt{\mu_{IX}}k^3}{3E_{Lat}}\right) \quad (17)$$

where μ_{IX} is the reduced interlayer exciton mass, k is related to the exciton binding energy ($k = \sqrt{2|E_B|}$) and Γ_0 is a field-independent material constant. The dissociation of interlayer excitons in a WSe_2/MoS_2 heterobilayer encapsulated in hBN is achieved in tens of picoseconds for electric fields higher than $4\ V\ \mu m^{-1}$ (ref. [20]). Instead, for low fields, exciton dynamics is dominated by the recombination rate, giving negligible photocurrent extraction.

Another rate equation must be considered for free-carrier transport. We assume that the same behavior applies to both electrons and holes, thus a simplified formulation is proposed with n equivalently representing both free-carrier densities[33]. A drift term (τ_{drift}) is included to account for the charges that are swept away by the lateral electric field. The free-carrier lifetime in the junction is calculated as $1/\tau_n = 1/\tau_{r,n} + 1/\tau_{drift}$, including both recombination ($\tau_{r,n}$) and drift (τ_{drift}):

$$\frac{dn}{dt} = \frac{N}{\tau_{diss}} - \frac{n}{\tau_{drift}} - \frac{n}{\tau_{r,n}} = \frac{N}{\tau_{diss}} - \frac{n}{\tau_n} \quad (18)$$

In the heterobilayer photodetectors, the split-gate electrodes induce an electrostatic *p-n* junction of length L in both TMDC layers, thus generating a lateral eletric field. The drift rate inside the junction can be approximated as $1/\tau_{drift} = 2\mu\overline{E_{Lat}}/L$, where $\overline{E_{Lat}}$ is the average electric field along the lateral direction in the junction and μ is the free-carrier mobility. An average value between electron and hole mobilities has been previously extracted for monolayer photodetecting devices ($\mu \sim 4\ cm^2\ V^{-1}\ s^{-1}$)[33], which we also adapt in our model as a reference.

In the low-density regime, the photoresponse rate can be estimated as $\Gamma_{ph} \simeq 1/(\tau_N + \tau_n)$[33] . In Figure S4a, three main zones are highlighted for Γ_{ph} corresponding to low, moderate and high

lateral fields. For $3\ V\ \mu m^{-1} \leq E_{Lat} \leq 6\ V\ \mu m^{-1}$, dissociation dominates the single-exciton decay rate, causing a steep rise in Γ_{ph} due to the exponential term in Eq. 27. With growing optical intensities, the second-order term in the recombination factor causes an increase in the electric field necessary to obtain the same output photocurrent, causing a shift in responsivity for moderate fields. For higher fields, free-carrier drift becomes the limiting factor for photocurrent generation, thus dictating the response times in the device. For E_{Lat} as high as $25\ V\ \mu m^{-1}$, response times of few picoseconds are possible, indicating photodetection speed comparable to the fastest 2D-material based devices ever reported in literature[34].

The steady-state solution to Eq. 27 gives the following interlayer exciton density:

$$N = \frac{1}{2\gamma\tau_N}\left[\sqrt{1+\gamma G(2\tau_N)^2} - 1\right] \tag{19}$$

With low generation factors ($\gamma G(2\tau_N)^2 \ll 1$), a linear relationship develops as $N \simeq G\tau_N$. By combining Eq. 29 with Eq. 27 we obtain:

$$n = \frac{1}{2\gamma}\frac{\tau_{r,n}\,\tau_{drift}}{\tau_{r,N}\,\tau_{diss}^2}\frac{\tau_{r,N}+\tau_{diss}}{\tau_{r,n}+\tau_{drift}}\left[\sqrt{1+\gamma G(2\tau_N)^2} - 1\right] \tag{20}$$

The extracted photocurrent is estimated by considering the number of free charges that are drifted out of the junction and then diffuse to the lateral contacts:

$$J_{ph} = q\frac{n}{\tau_{drift}} \tag{21}$$

The device responsivity is calculated as $R = J_{ph}/I_{opt}$, where I_{opt} is the input light intensity. In this analytical model, optical power is assumed to be uniformly distributed. The internal quantum efficiency is similarly extracted as the ratio between the photocurrent density and the generated electron-hole pairs per second in the device area ($IQE = J_{ph}/G$). Due to the relation

between the generation factor and the light beam intensity (Supplementary Note 1), R and IQE are linearly related whether they are calculated with the procedure and under the assumptions described above.

Dependency on optical intensity. Figure S4a shows R and IQE as a function of lateral electric field. The steep increase of R is shifted depending on light beam intensity, showing higher threshold values for growing input power. This behavior is an expression of the nonlinear component in exciton recombination. For $E_{Lat} > 6\ V\ \mu m^{-1}$, the response is limited by free-carrier dynamics and the separation between the three curves disappears. For lateral fields as high as $20\ V\ \mu m^{-1}$, responsivities of approximately $41\ mA\ W^{-1}$ and IQEs up to 82% are estimated. Similarly, Figure S4b shows the impact of different linear recombination times $\tau_{r,N}$ on the device response. Both higher optical intensities and lower first-order decay rates determine a decrease in R. In the case of the ideal p-n junction with uniform illumination, the main impact of these factors is a variation in the threshold lateral electric field, and thus in the threshold split-gate voltage.

These analytical results provide a good approximation to the performance of top-split-gate structure (Fig. 3c), where the assumption of a stepwise linear lateral junction with fixed length of $100\ nm$ is reasonable. Instead, for the bottom-split-gate geometry only the complete simulation procedure can fully encapsulate the photoresponse evolution with varying gate voltages. The maximum performance in the bottom split-gate detector can be analytically derived by considering an effective length of approximately $75\ nm$, based on the effective junction length (Fig. 3d-e).

Supplementary Note 6: Monolayer WSe_2 p-n junction

Here, we test the split-gate electrostatics in a simplified structure in order to compare it with other simulations reported in the literature. In the device of Figure S6a, the geometry is similar to the bottom-split-gate WSe_2/MoS_2 detector (Figure S3), but the MoS_2 layer is now absent. The *p-n* electrostatic potential is obtained solving the electrostatic problem for varying gate voltages with the methods described in Supplementary Note 2. In order to have a completely ideal case and focus on the electrostatically induced *p-n* junction, all work function differences are neglected. Therefore, we obtain a symmetrical lateral field distribution on WSe_2 (Figure S6b). With a linear approximation of the electrostatic potential in the junction, the maximum electric field can be estimated as the maximum potential difference (determined by the material bandgap) between the *p* and *n* regions divided by the nominal junction length (L):

$$E_{lin} = -\frac{E_g^{WSe_2}}{e\,L} \simeq -20\ V\ \mu m^{-1} \tag{22}$$

which is approximately equal to the simulated average junction field for $V_G = 1.5\ V$. However, for higher split-gate voltages, the effective junction length reduces due to a progressive decrease in the incompressible strip width[35], thus inducing higher lateral electric fields in the middle of the junction region. The simulated lateral field profile and its variations with split-gate voltages are compatible with what has been previously obtained in devices based on hBN-encapsulated monolayer WSe_2[33]. We note that a similar electric field distribution is also present in the results of the bottom-split-gate detector due to the similarities in the position and geometry of the electrodes (Fig. 3 in the main text).

Supplementary Note 7: Photocurrent extraction

Here, we highlight all the necessary steps for photocurrent extraction in our devices, where optical power, exciton densities and potentials are not uniform. First, the vertical and lateral electric fields in the heterostructure are both given by the electrostatic solution. Next, we extract the Stark shift component of the exciton potential energy from the vertical electric field as described in Supplementary Note 2. We solve exciton transport and obtain the interlayer exciton density distribution. Then, we calculate the lateral field in the heterobilayer as the average between the lateral field distributions in the two TMDC layers: $E_{Lat}^{HB} = \left(E_{Lat}^{WSe_2} + E_{Lat}^{MoS_2}\right)/2$. The effective junction length L_{eff} at a certain voltage V_G is defined as the portion of the x axis where the lateral field is sufficiently high to dissociate interlayer excitons. Next, we extract the average electric field $\overline{E_{Lat}}$ in the effective junction region and calculate the average drift time as $L_{eff}/\left(2\mu\overline{E_{Lat}}\right)$. Then, we extract the dissociated charge density from Eq. 28 and we estimate the generated photocurrent density as the average number of dissociated charge carriers that are effectively drifted out of the junction per second ($J_{ph} \simeq q\,\overline{n(x)}/\tau_{drift}$).

The detector responsivity is defined as the ratio between the extracted current and the optical power from the laser beam. The total input optical intensity is calculated as the beam intensity averaged over the total extension of the generation profile L_{tot}:

$$\overline{I(x)} = \frac{1}{L_{tot}} \int_{-L_{tot}/2}^{L_{tot}/2} I(x)dx \tag{23}$$

Since the current density is known, we calculate the responsivity as $R = J_{ph}/\overline{I(x)}$. Finally, the internal quantum efficiency is defined as the ratio between the photocurrent density and the generation term averaged over the active area $IQE = J_{ph}/\overline{G(x)}$, which is equivalent to the number of extracted carriers divided by the total generated electron-hole pairs in the structure.

Collection efficiency. With a diffraction-limited laser spot of diameter $2w_0 \sim 820\ nm$ and a split-gate separation of 100 nm, a significant part of the total interlayer exciton population is generated outside of the junction region. The collection efficiency in our devices can be defined as the ratio between the totally generated interlayer excitons and those that effectively reach the p-n junction[33]. In the low-density regime, the interlayer excitons generated in the *p* and *n* regions diffuse based on their diffusion length L_D within a decay time $\tau_{r,N}$. Thus, a near-unity collection efficiency is obtained for low optical power densities. For high generation factors, the second-order processes in R become dominant over the single-exciton decay term, and a growing number of interlayer excitons generated outside of the junction region recombine before they can diffuse towards the high-field region and contribute to the extracted photocurrent. Figure S7a-b show the responsivity curves obtained for optical power densities ranging from $10^3\ W/m^2$ to $10^{13}\ W/m^2$ in the top and bottom split-gate detectors. The decrease in collection efficiency for growing optical power is shown in Figure S7c.

In Figure S5, we have reported the responsivity curves for the two photodetector geometries with an unrealistic ultra-narrow beam spot diameter of 100 nm, equal to the split-gate separation adopted in the device configuration presented in the manuscript. These results show a change in the photodetector threshold voltage with no degradation of the maximum responsivity for higher optical power densities. Differently from what is shown in Figure 3 of the main text, in this case all the incident light falls within the p-n junction region. Thus, an increase in optical power does not influence the collection efficiency, which is inevitably equal to one. Therefore, the shape of the responsivity curves for varying power densities resembles closely that obtained in the analytical case of Figure S4, where a uniform distribution of light in the junction region was assumed.

Moreover, we have performed simulations for increasing spot size in the range between 800 nm and 1400 nm (Figure S8). For higher beam sizes, more interlayer excitons are generated further away from the junction region, thus impacting on the collection efficiency and causing a reduction in responsivity. This behavior is consistent between different power densities, taken in the range with the biggest variations in collection efficiency (10^6, 10^8 and $10^{10}\ W/m^2$).

Effect of electrostatic doping. Since $2w_0 > L_{eff}$ for both top and bottom split-gate devices, interlayer excitons are generated both inside and outside of the junctions. Thus, both neutral and charged interlayer excitons contribute to the response of the photodetector. A difference of less than one order of magnitude was previously reported between the lifetimes of neutral and charged interlayer excitons[24]. In Supplementary Note 5, we have considered the effects of the variability in the linear recombination terms on the detector response. From these investigations, we know that excitonic species with different lifetimes cause a shift in the threshold electric fields of the responsivity curves. Thus, we can include the effects of electrostatic doping on the detector performance by considering a variability in the threshold electric field of approximately $1\ V/\mu m$.

Split-gate separation and free-carrier mobility. An increase in L_{nom} is related to lower average junction fields due to the fact that the maximum *p-n* potential difference in the TMDs is equal to the bandgap of the material (Eq. 22). Thus, we report a decrease in responsivities and photoresponse rates for growing L_{nom} in both top and bottoms split-gate detectors (Figure S9).

In order to analyze the variability of electrical transport in the TMDs, for example due to different densities of defects, we have directly investigated the change in the photodetector performance in the case of field-effect mobility in a range between 0.4 and $40\ cm^2\ V^{-1}\ s^{-1}$. The results of these simulations are reported in Figure S7 for different optical intensities in both top and bottom split-gate geometries. Higher free-carrier mobilities correspond to higher maximum responsivities due to the increase in the drift rate. These results are compatible with Eq. 20 in Supplementary Information Note 5, from which we extract a behavior $R \propto \alpha\mu/(1+\beta\mu)$, where α and β are constants given by the corresponding combination of kinetic terms. This picture also indicates an increase in the photoresponse rate with respect to the mobility, as obtained in Figure S10.

Supplementary Information References

1. Weber, J. W., Calado, V. E. & van de Sanden, M. C. M. Optical constants of graphene measured by spectroscopic ellipsometry. *Appl. Phys. Lett.* **97**, 091904 (2010).
2. Lee, S.-Y., Jeong, T.-Y., Jung, S. & Yee, K.-J. Refractive Index Dispersion of Hexagonal Boron Nitride in the Visible and Near-Infrared. *physica status solidi (b)* **256**, 1800417 (2019).

3. Jung, G.-H., Yoo, S. & Park, Q.-H. Measuring the optical permittivity of two-dimensional materials without a priori knowledge of electronic transitions. *Nanophotonics* **8**, 263–270 (2018).
4. Malitson, I. H. Interspecimen Comparison of the Refractive Index of Fused Silica*,†. *J. Opt. Soc. Am., JOSA* **55**, 1205–1209 (1965).
5. Aspnes, D. E. & Studna, A. A. Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV. *Phys. Rev. B* **27**, 985–1009 (1983).
6. Peng, B. *et al.* Ultrafast charge transfer in MoS 2 /WSe 2 p–n Heterojunction. *2D Mater.* **3**, 025020 (2016).
7. Ceballos, F., Bellus, M. Z., Chiu, H.-Y. & Zhao, H. Ultrafast Charge Separation and Indirect Exciton Formation in a MoS2–MoSe2 van der Waals Heterostructure. *ACS Nano* **8**, 12717–12724 (2014).
8. Zhu, H. *et al.* Interfacial Charge Transfer Circumventing Momentum Mismatch at Two-Dimensional van der Waals Heterojunctions. *Nano Lett.* **17**, 3591–3598 (2017).
9. Yuan, L. *et al.* Twist-angle-dependent interlayer exciton diffusion in WS 2 –WSe 2 heterobilayers. *Nature Materials* **19**, 617–623 (2020).
10. Unuchek, D. *et al.* Room-temperature electrical control of exciton flux in a van der Waals heterostructure. *Nature* **560**, 340–344 (2018).
11. Latini, S., Winther, K. T., Olsen, T. & Thygesen, K. S. Interlayer Excitons and Band Alignment in MoS2/hBN/WSe2 van der Waals Heterostructures. *Nano Lett.* **17**, 938–945 (2017).
12. Ziegler, D. *et al.* Variations in the work function of doped single- and few-layer graphene assessed by Kelvin probe force microscopy and density functional theory. *Phys. Rev. B* **83**, 235434 (2011).

13. Leenaerts, O., Partoens, B., Peeters, F. M., Volodin, A. & Haesendonck, C. V. The work function of few-layer graphene. *J. Phys.: Condens. Matter* **29**, 035003 (2016).

14. Ivanov, A. L. Quantum diffusion of dipole-oriented indirect excitons in coupled quantum wells. *EPL* **59**, 586 (2002).

15. Ivanov, A. L., Littlewood, P. B. & Haug, H. Bose-Einstein statistics in thermalization and photoluminescence of quantum-well excitons. *Phys. Rev. B* **59**, 5032–5048 (1999).

16. Shi, H. *et al.* Exciton Dynamics in Suspended Monolayer and Few-Layer MoS2 2D Crystals. *ACS Nano* **7**, 1072–1080 (2013).

17. Sun, D. *et al.* Observation of Rapid Exciton–Exciton Annihilation in Monolayer Molybdenum Disulfide. *Nano Lett.* **14**, 5625–5629 (2014).

18. Mouri, S. *et al.* Nonlinear photoluminescence in atomically thin layered ${\mathrm{WSe}}_{2}$ arising from diffusion-assisted exciton-exciton annihilation. *Phys. Rev. B* **90**, 155449 (2014).

19. Wang, Z., Chiu, Y.-H., Honz, K., Mak, K. F. & Shan, J. Electrical Tuning of Interlayer Exciton Gases in WSe2 Bilayers. *Nano Lett.* **18**, 137–143 (2018).

20. Binder, J. *et al.* Upconverted electroluminescence via Auger scattering of interlayer excitons in van der Waals heterostructures. *Nature Communications* **10**, 1–7 (2019).

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

22. Rivera, P. *et al.* Observation of long-lived interlayer excitons in monolayer MoSe 2 – WSe 2 heterostructures. *Nature Communications* **6**, 6242 (2015).

23. Unuchek, D. *et al.* Valley-polarized exciton currents in a van der Waals heterostructure. *Nature Nanotechnology* **14**, 1104–1109 (2019).

24. Jauregui, L. A. *et al.* Electrical control of interlayer exciton dynamics in atomically thin heterostructures. *Science* **366**, 870–875 (2019).

25. Zimmermann, R. & Schindler, C. Exciton–exciton interaction in coupled quantum wells. *Solid State Communications* **144**, 395–398 (2007).

26. Schindler, C. & Zimmermann, R. Analysis of the exciton-exciton interaction in semiconductor quantum wells. *Phys. Rev. B* **78**, 045313 (2008).

27. Gummel, H. K. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *IEEE Transactions on Electron Devices* **11**, 455–465 (1964).

28. Cadiz, F. *et al.* Exciton diffusion in WSe2 monolayers embedded in a van der Waals heterostructure. *Appl. Phys. Lett.* **112**, 152106 (2018).

29. Choi, J. *et al.* Moiré potential impedes interlayer exciton diffusion in van der Waals heterostructures. *Science Advances* **6**, eaba8866 (2020).

30. Liu, Y. *et al.* Electrically controllable router of interlayer excitons. *Science Advances* **6**, eaba1830 (2020).

31. Baldo, M. & Stojanović, V. Excitonic interconnects. *Nature Photonics* **3**, 558–560 (2009).

32. Kamban, H. C. & Pedersen, T. G. Interlayer excitons in van der Waals heterostructures: Binding energy, Stark shift, and field-induced dissociation. *Sci Rep* **10**, 5537 (2020).

33. Massicotte, M. *et al.* Dissociation of two-dimensional excitons in monolayer WSe 2. *Nature Communications* **9**, 1633 (2018).

34. Koppens, F. H. L. *et al.* Photodetectors based on graphene, other two-dimensional materials and hybrid systems. *Nature Nanotech* **9**, 780–793 (2014).

35. Chklovskii, D. B., Shklovskii, B. I. & Glazman, L. I. Electrostatics of edge channels. *Phys. Rev. B* **46**, 4026–4034 (1992).