Hot carrier-enhanced interlayer electron-hole pair multiplication in 2D semiconductor heterostructure photocells

Fatemeh Barati, Max Grossnickle, Shanshan Su, Roger Lake, Vivek Aji, Nathaniel Gabor*

*Correspondence to: nathaniel.gabor@ucr.edu

Table of contents:

S1. MoSe₂/WSe₂ Heterostructure Synthesis (Figure S1)
S2. Device Fabrication, Raman, Photoluminescence, Photocurrent, and Differential Reflection Spectroscopy (Figure S2, Figure S3)
S3. MoSe₂/WSe₂ Electronic Band Structure (Figure S4)
S4. Interlayer Electron Transport Through an Atomic Layer Hetero-Junction (Figure S5)

Supplementary Materials:

The atomic layer semiconductor heterostructures studied in this work were assembled using exfoliated bulk transition metal dichalcogenide (TMD) crystals and novel annealing processes to explore intrinsic electronic transport behavior and electron-hole pair generation processes. The following supplement contains detailed characterization of two devices; each was assembled and annealed using novel techniques, described herein. Extensive optical spectroscopic characterization (including Raman, photoluminescence, differential reflection, and photocurrent spectroscopy) confirmed strong interlayer electronic coupling at the heterostructure interface. Band structure calculations of the MoSe₂/WSe₂ heterostructure show features indicative of an intrinsic type II hetero-junction, and exhibit strong qualitative and quantitative agreement with the band offset energies obtained from electronic transport measurements. We also present a phenomenological model of interlayer transport between two atomic layer semiconductors, which captures the key features of the interlayer current, including interlayer electron-hole pair production. A comparison of our model to experimental data is then followed by a brief discussion of photoresponse in the MoSe₂/WSe₂ photocell.
S1. MoSe$_2$/WSe$_2$ Heterostructure Synthesis

We fabricated MoSe$_2$/WSe$_2$ heterostructures by mechanical exfoliation of WSe$_2$ and MoSe$_2$ flakes from bulk crystals (2D Semiconductor) onto Si wafers coated with 290nm-thick SiO$_2$. The heterostructure devices were assembled using a highly customized, temperature-controlled transfer microscope that ensures that the interface between the two layers has no intentional contact to polymer films. The dry pick-up transfer process, described below, results in a heterostructure with minimal interfacial contamination, and is followed by two annealing processes (described in Section S2).

The van der Waals pick-up and transfer process is based on Andres Castellanos-Gomez, et al.$^1$, and is represented schematically in Figure S1 to highlight key distinctions of this work. We first use a stamp to pick up the target layer, (Fig. S1a, top). The stamp consists of a thin layer of polypropylene carbonate (PPC) and a layer of polydimethylsiloxane (PDMS) on top of a glass slide. The stamp is lowered until the PPC contacts the first target layer, WSe$_2$. We then heat the stage to 40° C. We next cool the stage to 34°C and very quickly lift the stamp, as shown in Figure S1b (top). To assemble the heterostructure, we lower the WSe$_2$ flake and place it on to the MoSe$_2$ flake on the stage as shown in Figure S1c (top). The stage is then heated to 80°C to melt the PPC and the stamp is lifted very slowly, causing the PPC to separate from the PDMS. This leaves the stack of two flakes with PPC on top. We remove the residual PPC using heated acetone. Figure S1 (a, and b), bottom, show optical images of exfoliated flakes of WSe$_2$ and MoSe$_2$. Dashed lines highlight the individual layers and overlapped region of the completed heterostructure.

S2. Device Fabrication, Raman, Photoluminescence, Photocurrent, and Differential Reflection Spectroscopy

In this section, we present comprehensive characterization of heterostructures composed of monolayer molybdenum diselenide (MoSe$_2$) and bilayer tungsten diselenide ($2L$-WSe$_2$). We include detailed descriptions of device fabrication, optical, and optoelectronic spectroscopy measurements on two devices that have been thoroughly characterized using all of the techniques described. All experimental results were typical of Devices 1 and 2.
Figure S1

Supplementary Fig. 1 | Schematic illustration and optical images of the dry-transfer process. a, b, c, top, Schematic illustration of the transfer microscope at different temperatures for different purposes. a, b, c, bottom, Optical images of exfoliated flakes of WSe$_2$, MoSe$_2$, and MoSe$_2$/2L-WSe$_2$ heterostructures respectively. Dashed lines indicate the area of the individual layers and the overlapped region, which are about 14µm$^2$, 8µm$^2$, and 4.3µm$^2$ respectively.
The devices, shown in Figure S2, were first characterized to identify layer thickness, and ensure interlayer coupling, as determined through Raman and photoluminescence (PL) spectroscopy. Many heterostructures were assembled, and only those exhibiting the PL signatures indicative of electron transfer and equilibrium charge redistribution between the layers (Neil R. Wilson, et al.)² were selected for comprehensive measurements. To electrically access individual layers of the heterostructures, we deposited 5/150nm thick Ti/Au contacts using standard electron beam lithography. We then mounted the devices into a customized Janis helium flow optical cryostat in order to measure the temperature-dependent current-voltage characteristics under high vacuum.

Figure S2 shows Raman and photoluminescence spectroscopy measurements that were carried out before and after constructing the heterostructure, as well as electronic transport measurements. No quantitative or qualitative changes were observed before/after transport measurements. Figure S2b shows the Raman spectrum for the MoSe₂ and WSe₂ flakes for Device 1. For MoSe₂, the A₁g peak at 241 cm⁻¹, E₂g¹ peak at 288 cm⁻¹, and lack of B₂g¹ peak between 350 and 360 cm⁻¹ are characteristic of monolayer thickness³². The Raman spectrum of the WSe₂ flake shows the A₁g mode at 250 cm⁻¹ and B₂g¹ mode at 309 cm⁻¹, which indicates bilayer thickness².

The PL spectra of the heterostructure devices show several key features that indicate interlayer electron transfer and equilibrium charge redistribution between the layers in the heterostructure (Pasqual Rivera, et al.⁴, Frank Ceballos, et al.⁵). Figure S2c shows the photoluminescence (PL) spectra of MoSe₂, 2L-WSe₂, and the heterostructure. The PL emission of MoSe₂ has a large peak at 1.57 eV and a shoulder at 1.75 eV, while the 2L-WSe₂ exhibits two peaks at 1.46 eV and 1.66 eV, in agreement with previously reported data². Figure S2 (d, e, and f) shows characterization data for a second 2L-WSe₂/MoSe₂ heterostructure, Device 2. Raman spectra and PL are again in agreement with the results of Tonndorf et al.³.

Annealing Process 1. After devices were placed into high vacuum (10⁻⁶ torr), we first annealed the samples at 420°C for 48 hours to improve device contact quality and remove surface contaminants (Marco M. Furchi, et al.⁶). Following this procedure, the devices were kept under a 10⁻⁶ torr vacuum at all times.
Supplementary Fig. 2 | Optical images, Raman spectroscopy, and PL spectroscopy of Devices 1 and 2. a, d, Optical images of the two MoSe₂/2L-WSe₂ heterostructures with contacts. b, e, Raman spectra (excitation wavelength λ=532nm) for MoSe₂ (green), and 2L-WSe₂ (blue). c, f, PL spectra of MoSe₂ (green), 2L-WSe₂ (blue), and the stack (black).
**Annealing Process 2.** We carried out scanning photocurrent spectroscopy measurements using a tunable Ti:sapph laser with 200 fs pulse duration and a 76 MHz rep rate combined with a custom-built scanning microscope. The photocurrent was measured with a diffraction-limited beam spot at various wavelengths over the Ti:sapph laser range (wavelengths 700-1000 nm) and at average powers below 1 mW. After initial photocurrent measurements to assess device contacts, the device resistance and DC current noise was significantly decreased when measured at low and high $V_{SD}$ voltage. An order of magnitude decrease in the low-bias ($V_{SD} < 0.5$ V) resistance was typical after 12-24 hours of sustained scanning photocurrent measurements under vacuum, after which time the device remained stable. Each of the devices described here were measured for more than 60 days, and demonstrated stable operation after annealing.

After annealing, photocurrent and differential reflection spectra were taken to confirm interlayer coupling at the interface, as determined by the strong correlation with PL spectral features. We measured photocurrent and reflection spectra with a 0.7 mW beam focused to a diffraction-limited 500 nm diameter (FWHM) spot. Figure S3a shows photocurrent (red), $\Delta R/R$ (blue), and PL (gray) as a function of incident photon energy. The quantity $\Delta R/R$ is the difference between the reflection at any point and the reflection of the silicon substrate, normalized by the background reflection from the substrate. In Figure S3b we show spatial $\Delta R/R$ maps taken at different wavelengths. We observed a clear spatial dependence of $\Delta R/R$, and a wavelength-dependent transition from positive, through zero, and then to negative $\Delta R/R$. We took the data from the green dot in Figure S3b to plot the $\Delta R/R$ spectrum in Figure S3a. The zero crossing of the $\Delta R/R$ spectrum occurs near 1.45 eV, and a minimum in $\Delta R/R$ occurs at 1.66 eV, slightly higher in energy than the peak in PL emission. The photocurrent shows a peak near 1.66 eV and continues to increase at higher photon energies.

**S3. MoSe$_2$/WSe$_2$ Electronic Band Structure**

Band dispersions and band gap energies of the MoSe$_2$/2L-WSe$_2$ heterostructure including spin-orbit coupling (SOC) were calculated using the Vienna ab initio simulation package (VASP)$^{8-10}$ in the projected-augmented-wave method$^{11}$. We used the generalized
Figure S3

Supplementary Fig. 3 | Differential reflection and photocurrent spectroscopy measurements of Device 1. a, Photocurrent, ΔR/R, and PL spectra from MoSe₂/2L-WSe₂ heterostructure illustrated respectively by red, blue, and gray lines. b, Spatial ΔR/R maps at different excitation photon energies (labeled).
gradient approximation (GGA) of the Perdew-Burke-Ernzerhof (PBE) form\textsuperscript{12-14} for the exchange correlation energy. The van der Waals (vdW) interactions between the layers are accounted for by using the DFT-D2 method of Grimme\textsuperscript{15}. We set the kinetic energy cutoff for our calculation at 500 eV. For all structural relaxations, the convergence tolerance on the Hellmann-Feynman forces is less than 0.01 eV/Å. An $8 \times 8 \times 1$ $\Gamma$-centered Monkhorst-Pack k-point mesh is used for the 2D films. A vacuum layer of 20Å is included in the supercell in the z-direction to prevent interactions between the periodic repetitions of the two-dimensional structure. With these settings, the lattice constant of MoSe$_2$ is calculated as 3.3247 Å and the lattice constant of WSe$_2$ as 3.326 Å, so the lattice mismatch between the two materials is less than 0.1%.

Our simulation results are consistent with the published lattice constants of the two materials\textsuperscript{16,17}. For the trilayer heterostructure, the lattice constant is 3.3254 Å. The calculated vdW gap between MoSe$_2$ and bilayer WSe$_2$ is 3.15 Å (the vertical distance between the Se and Se atoms) which is consistent with the published results of Ref.\textsuperscript{18}. The thickness of the trilayer system is 12.94 Å (the vertical distance between Mo to the furthest W atom). The calculated PBE bandgaps of monolayer MoSe$_2$ and bilayer WSe$_2$ are 1.317 eV and 1.18 eV, respectively. Our calculated results are consistent with prior simulation results\textsuperscript{19,20}. For comparison, experimental measurements show that the bandgap of monolayer MoSe$_2$ is 1.55 eV\textsuperscript{21}, and the bandgap of bilayer WSe$_2$ has an indirect bandgap of 1.26 eV\textsuperscript{22}.

The band alignment and band structure of the combined trilayer TMD heterostructure under different electric fields are calculated by determining the orbital compositions of the bands at the high symmetry points. If, for example, the band is primarily composed of Mo $d_{z^2}$ orbitals, then it is assigned to the MoSe$_2$ layer. A small sample of band-alignments are shown in Figure S4 below.

Band structure calculations of the MoSe$_2$/2L-WSe$_2$ heterostructure give an intrinsic type II heterojunction and exhibit strong qualitative and quantitative agreement with the band offset energies obtained from electronic transport measurements. Figure S4 shows the calculated MoSe$_2$/2L-WSe$_2$ heterostructure band alignment under negative (Fig. S4a), zero (Fig. S4b), and positive (Fig. S4c) applied electric field. The electric field is oriented perpendicular to the 2D plane to simulate the electric field due to an interlayer applied $V_{SD}$. 
With no applied electric field, the band gap is indirect in k-space with a 0.94 eV gap, and it shows features indicative of a type II heterojunction, as illustrated in Fig. S4b. The CBM of the trilayer is at the K valley of the MoSe₂ composed of Mo \( d_{z^2} \) orbitals, while the VBM is at the \( \Gamma \) valley of the WSe₂ composed of W \( d_{z^2} \) orbitals. The bandgaps associated with the bilayer WSe₂ and the monolayer MoSe₂ are affected by the proximity of the other material and slightly shift to 1.14 eV and 1.25 eV, respectively. The WSe₂ valence K valley is 80 meV below the valence \( \Gamma \) valley.

A negative applied field increases the conduction band offset, and it also decreases the type II bandgap. Also, in k-space, the type II bandgap transitions from indirect to direct. At a field of -0.2 V/ Å the direct, type II bandgap is 0.73 eV as shown in Figure S4a. The CBM of the trilayer remains at the K valley of the MoSe₂ composed of Mo \( d_{z^2} \) orbitals, while the VBM moves to the K valley of the WSe₂ composed of W \( d_{x^2-y^2} \) and \( d_{xy} \) orbitals. With respect to the 0-bias structure, the positions of the WSe₂ valence \( \Gamma \) and K valleys reverse such that the K valley is 10 meV higher than the \( \Gamma \) valley. The WSe₂ conduction band remains at K composed of \( d_{z^2} \) orbitals. The MoSe₂ also has a direct gap at the K valley, and the orbital compositions are the same as for WSe₂. The conduction band offset becomes equal to the bandgap at a field of approximately 0.3 V/ Å. Using the 12.94 Å thickness of the trilayer structure to convert the electric field into a voltage gives a voltage of 3.9 V. At this voltage, the electrons gain enough kinetic energy transitioning from the WSe₂ to the MoSe₂ to cause impact ionization. Furthermore, as the bias becomes more negative, the bandgap decreases, consistent with the picture described in the main text.
Supplementary Fig. 4 | Schematics of the MoSe₂/2L-WSe₂ electronic band structure. a, b, c, schematic illustration of the electronic band structure of the heterostructure under negative, zero, and positive electric field, respectively. Gray horizontal dashed lines show chemical potential of the individual layers, and the region in between the two vertical dashed lines is the junction. d. K-valley energies as a function of applied electric field. Positive electric field is in the direction of forward applied $V_{SD}$. Data colors match material assignments in a-c. Circles (inverted triangles) correspond to WSe₂ (MoSe₂) conduction band minimum. Squares (triangles) correspond to WSe₂ (MoSe₂) valence band maximum.
Under positive electric field, the movement of the bands becomes more complex. A positive field drives the type II bands into energetic alignment. The WSe$_2$ and MoSe$_2$ bands couple and shift with different valleys shifting by different amounts, and this alters the relative alignments of the different valleys. At a field of 0.2 V/Å, the CBM switches from the MoSe$_2$ layer to the WSe$_2$ layer as shown in Figure S4c. Furthermore, the CBM shifts in k-space from K to $\Delta$ (between $\Gamma$ and K) where it is composed of W $d_{x^2-y^2}$ and $d_{xy}$ orbitals. The VBM remains at $\Gamma$ of the WSe$_2$ composed of W $d_{z^2}$ orbitals. The band alignment switches from type II to type I with both the conduction and valence bands residing in the WSe$_2$ layer. The bandgap is 0.96 eV.

Figure S4d shows the movement of the conduction and valence K-valleys as a function of applied electric field. For negative electric fields, the plot shows the movement of the band edges, since the band edges reside at K at negative fields. At positive fields, the movement of the band edges is more complex due to the shifting of the different valley positions.


In this section, we obtain the temperature and $V_{SD}$ dependence of the interlayer current $I$ by a straightforward calculation based on the electron transmission formalism (Supriyo Datta 1995$^{23}$). Following Chul-Ho Lee, et al.$^{24}$, we model a 2D $n$-type, type II heterojunction in which most of the voltage drop in the device occurs across the vertical junction, leaving no appreciable potential barriers in the lateral transport direction within each semiconductor layer. In our model, we assume these contributions to the functional form of the $I-V_{SD}$ characteristics to be negligible, and later compare this assumption to experimental data. We then evaluate the current in forward and reverse bias by assessing the rates of electron transmission through the $n$-type heterojunction (Fig. S5), including the rate of interlayer $e$-$h$ pair multiplication. Our model assesses only carrier kinetic energy, and asks: at what kinetic energy does a carrier have sufficient excess energy to produce additional electron-hole pairs?

Our phenomenological model was developed to qualitatively explain the temperature dependence of the $I-V_{SD}$ characteristics in the voltage regime just below the appearance of NDC. Like conventional heterojunction devices, the current-voltage characteristics are quite
complicated, exhibiting several regimes that are described by various microscopic processes. In conventional n-n$^+$ and p-n junctions for example, different regimes often exhibit different ideality factors $n$. The commonly used ideality factor is attributed to various mechanisms, but strictly speaking is a phenomenological constant included to express the proportionality between electric potential drop and applied voltage (see main text Reference 29). In comparison to our data (at the end of this section), we thus include a phenomenological parameter $\alpha$, which quantifies the ‘lever arm’ between electric potential energy and applied voltage $V_{SD}$.

The 2D heterojunction model requires that we first calculate the density of electronic states $g(E)$ in each of the TMD atomic layer semiconductors. Our starting model for individual TMD layers is the effective tight-binding two-valley Bloch Hamiltonian (Evan Sosenko, et al.$^{25}$, Di Xiao, et al.$^{26}$, Xiaodong Xu, et al.$^{27}$) which gives the energy spectrum (shown in Figure S5):

$$E_{\tau s}^n(k) = \frac{1}{2} \left( \lambda \tau s + n \sqrt{(2\alpha t k)^2 + (\Delta - \lambda \tau s)^2} \right). \quad (S1)$$

Here, $\Delta$ is the band gap energy, $2\lambda$ is the spin splitting in the valence band, the valley index $\tau=\pm 1$ corresponds to $\pm K$ points, and the spin index $s=\pm$ corresponds to the $z$-component of the spin. $n = \pm 1$ indexes the conduction and valence band, respectively.$^{25}$ We note that the energy spectrum is hyperbolic; $E(k)$ is approximately parabolic at low energies, but tends towards a constant linear slope at high electron energies.

The density of electronic states in the conduction band ($n = 1$) can be calculated directly from Equation (S1) and is given by:

$$g(E) = \frac{1}{\pi a^2 t^2} \left( E + E_{GAP} \right) = \frac{E_{GAP}}{\pi a^2 t^2} \left( 1 + E/E_{GAP} \right) \quad (S2)$$

Here, $E_{GAP}$ is the energy difference between the conduction and valence band in the $K$ and $K'$ valleys of a monolayer TMD, such as MoSe$_2$. From Equation (S2), the density of states approaches a constant value for electrons near the bottom of the bands, but depends linearly on the electron energy at high energies. Importantly, we note that this result is distinct from the density of states in conventional 2D electron gases, where $g(E)$ does not depend on energy due to parabolic band structure.
Supplementary Fig. 5 | Schematics of TMD band structure and interlayer electronic transport. 

a, energy band diagram of a 2D-TMD material with hyperbolic band structure, b, c, electronic band structure of the MoSe$_2$/2L-WSe$_2$ under forward and reverse bias respectively, the area inside the vertical dashed line illustrates the hetero-junction.
**Forward bias.** Figure S5 shows a schematic of the potential energy landscape for electrons in the heterojunction for \( V_{SD} > 0 \) V. For \( k_B T \gg \varepsilon_F \), we assume classical Boltzmann distributions to evaluate how many right-moving electrons (i.e., electrons that are transferred from MoSe₂ into WSe₂) transit the interface when a positive \( V_{SD} \) is applied. From the left side, we consider a charge carrier with energy that is just enough to get over the potential barrier at the interface \( \Delta E_c^0 \). From the diagram, this carrier has kinetic energy \( \varepsilon - \varepsilon_F = \Delta E_c^0 - eV_{SD} \) for \( V_{SD} > 0 \) V.

In the forward bias interlayer transport process, conduction occurs through electron states near the bottom of the conduction bands in both materials. In this case, the density of states in both layers is approximately independent of energy. The interlayer current results from counting only right-transiting electrons and is given by:

\[
I_+ = \int_{\varepsilon}^{\infty} e v \eta_0 g_W(E) g_{Mo}(E) f(E) dE
\]  

where \( v = v_F \) is the electron Fermi velocity, \( \eta_0 \) is the transmission coefficient, \( g_W(E) \) and \( g_{Mo}(E) \) are the density of states in WSe₂ and MoSe₂, respectively, and \( f(E) \) is the electron distribution in MoSe₂. Assuming \( eV_{SD}, \Delta E_c^0 \gg k_B T \), the interlayer current is then

\[
I_+ = \Gamma_+ (k_B T) \exp\left(-\frac{(\Delta E_c^0 - eV_{SD})}{k_B T}\right)
\]  

where

\[
\Gamma_+ = \frac{e v \eta_0}{(\pi \hbar^2)^2} \left(\Delta E_c^0 + \Delta E_h^0\right) \left(\Delta E_h^0 + \Delta E_v^0\right).
\]  

Here, we model the system to have constant transmission coefficient \( \eta_0 \) for carriers whose total kinetic energy is sufficient to transit the barrier between the TMD layers. According to Figure S5, the band gap energy in WSe₂ is approximately equal to \( \left(\Delta E_c^0 + \Delta E_h^0\right) \) and in MoSe₂ equal to \( \left(\Delta E_h^0 + \Delta E_v^0\right) \).

Equation (S4) indicates that in forward bias the normalized current \( I_+ / T \) should exhibit an exponential decrease with \( 1/T \), from which the activation energy \( \Delta E_c^0 \) can be extracted.
Equation (S4) can also be re-expressed in terms of a temperature-independent quantity that depends only on $\Delta E_c^0$ and $V_{SD}$:

$$(k_B T) \ln \left( \frac{I_+}{\Gamma} + k_B T \right) = -\Delta E_c^0 + eV_{SD}. \quad (S5)$$

The forward bias $I-V_{SD}$ characteristics, when rescaled as $T \ln(I_+/T)$ as in Equation (S5), should collapse to a single characteristic. The current $T \ln(I_+/T)$ vs. $V_{SD}$ exhibits temperature independent behavior with a linear increase as $V_{SD}$ increases.

The data for all temperatures (Fig. 3d) indeed exhibit this behavior, collapsing to a temperature-independent characteristic current $T \ln(I/T)$ . We infer the conduction band offset $\Delta E_c^0 = 0.2$ eV from an exponential fit of the forward bias current $I/T$ as a function of $1/T$ (Fig. 3d inset). The value obtained from transport measurements shows excellent agreement with band structure calculations $\Delta E_c^0$ (theory) = 0.20 eV (Supplementary Fig. S4). The energy $\Delta E_c^0$ is the maximum kinetic energy available to a hot electron at $V_{SD} = 0$ V (Fig. 1b).

**Reverse bias.** In reverse bias, the temperature and $V_{SD}$ dependence of the interlayer current serves as a signature of hot electron injection into MoSe$_2$, while the thermally activated behavior arises from the energetic threshold for $e-h$ pair multiplication. The electron-hole pair production process takes place in two steps: (1) a low-energy electron in WSe$_2$ is transferred into a high-energy state in MoSe$_2$. (2) The high-energy (hot) electron in MoSe$_2$ decays and produces an interlayer $e-h$ pair. As the reverse bias voltage is increased, electron-hole pairs generated by high-energy electrons are separated in the electric field and increase the device current. We first define the energy barrier for additional electron generation due to $e-h$ pair multiplication, and then evaluate the interlayer current $I$ that results from $e-h$ pair multiplication when $V_{SD} < 0$ V.

For interlayer transport from WSe$_2$ (right) to MoSe$_2$ (left), it is important to note that electrons need not overcome a potential energy barrier, but instead may gain the potential energy $\Delta E_c^0$. In conventional devices, this leads to reverse bias saturation current with weak dependence on $V_{SD}$ or temperature$^6$. However, interlayer $e-h$ pair production requires a hot electron with a minimum excess kinetic energy, and thus depends strongly on both $V_{SD}$ and
temperature. From the schematic in Figure S5, the minimum kinetic energy required for an electron to generate an interlayer e-h pair is given by:

\[ K_e^* = \Delta E_h^0 - \Delta E_c^0 + 2eV_{SD}, \]  
(S6)

where \( \Delta E_h^0 \) is the interlayer e-h pair excitation energy at \( V_{SD} = 0 \) V. We note that when \( V_{SD} < 0 \) V, the potential energy that can be gained by an electron is increased \( \Delta E_c = \Delta E_c^0 + e|V_{SD}| \), while the interlayer excitation energy is simultaneously decreased \( \Delta E_h = \Delta E_h^0 - e|V_{SD}| \), resulting in Equation (S6).

In reverse bias, the energy difference between the conduction bands of the two materials is large. Due to the electric field at the interface, low-energy electrons in WSe\(_2\) - with approximately constant density of states - are transferred into hot electrons in MoSe\(_2\) at high kinetic energy with density of states proportional to \( E \) (Eq. (S2)). The current due to interlayer e-h pair multiplication \( I_M \) is then

\[ I_M \approx \int_{-\infty}^{\infty} ev\gamma_0 \left( \frac{E_{GAP}}{\pi a^2 t^2} \right) \left( \frac{E}{\pi a^2 t^2} \right) \exp\left( -\frac{(E-E_F)}{k_B T} \right) dE; \]  
(S7a)

\[ = \frac{ev\gamma_0 \left( \Delta E_c^0 + \Delta E_h^0 \right)}{\left( \pi a^2 t^2 \right)^2} \int_{-\infty}^{\infty} E \exp\left( -\frac{(E-E_F)}{k_B T} \right) dE, \]  
(S7b)

where we assume that the product of the transmission coefficient and impact excitation rate \( \gamma_0 \) is approximately constant above the kinetic energy threshold. Computing the integral over the entire energy range and noting \( k_B T >> \epsilon_F \), gives

\[ I_M = \Gamma_M \left( k_B T \right)^2 \left( \exp\left( -K_e^*/k_B T \right) - 1 \right); \]  
(S8a)

\[ \Gamma_M = \frac{ev\gamma_0 \left( \Delta E_c^0 + \Delta E_h^0 \right)}{\left( \pi a^2 t^2 \right)^2}. \]  
(S8b)

In contrast to the forward bias current, there is an additional power of \( k_B T \) in the reverse bias current, which arises from the density of available electron states in MoSe\(_2\).

Equation (S8) indicates that in the \( V_{SD} \) and temperature regime in which e-h pair multiplication is the dominant source of current, the reverse bias current can be re-expressed in
terms of a temperature-independent quantity that depends on $\Delta E^0_h$. The reverse bias $I-V_{SD}$ characteristics, when rescaled as $T \ln(I_M/T^2)$ as in Equation (S8), should collapse to a single line. In this case, the voltage dependence of $T \ln(I_M/T^2)$ exhibits universal behavior, with a linear increase with $V_{SD}$.

While Equation (S8) gives the temperature scaling of the current due to $e\cdot h$ pair multiplication, an additional straightforward analysis is required to obtain $\Delta E^0_h$. First, note that Equation (S8) contains a $V_{SD}$-independent term and a $V_{SD}$-dependent term. In order to isolate the activation energy $\Delta E^0_h$, we simply take the derivative $dI_M/dV_{SD}$ of Equation (S8), giving

$$\frac{|dI_M/dV_{SD}|}{(2k_BT)} = \exp(-K_e^*/k_BT).$$  

(S9)

Equation (S9) indicates that the quantity $(dI_M/dV_{SD})/T$ vs. $1/T$ should exhibit an exponential decrease, from which the activation energy $K_e^*$ can be extracted.

**Comparison of the model to experimental Data.** Our simple model agrees well with the temperature dependence observed in Figure 3, but may not capture all details of the interlayer transport processes at all biases. Importantly, we note that the model agrees well only within the voltage regime leading up to the onset of NDC ($-3 \text{ V} < V_{SD} < -0.5 \text{ V}$). Within this voltage regime, $\alpha$ remains constant and can be extracted from the temperature-dependent data such as that shown in Figure 3c. The obtained value for the device of Figure 2, $\alpha = 0.2$, is typical for TMD heterostructure devices studied here, and likely arises from a combination of contact resistance, resistance drop across the individual materials (see main text Reference 22), as well as intralayer recombination processes. Figure S6 shows $I-V_{SD}$ characteristics over the range of voltages that can be well described by our model. The rescaled current $T\ln(I/T^2)$ changes approximately linearly with $V_{SD}$ over the range $-3 \text{ V} < V_{SD} < -0.5 \text{ V}$ and for increasing values of $V_G$ near the onset threshold. $T\ln(I/T^2)$ also increase linearly with $V_G$ over the range of voltages $-37 \text{ V} < V_G < -31 \text{ V}$, which is precisely the onset regime. The linear dependence on $V_{SD}$, and the corresponding constant value of $\alpha$ over this voltage regime strongly suggests that the potential drop across the interface scales proportionally to the applied voltage in a similar way over the entire voltage range near the onset to NDC.

Of importance, the voltage drop along the lateral direction of the interface may depend
Figure S6

Supplementary Fig. 6 | Interlayer current-voltage characteristics in the transport regime near the onset to NDC. a, Reverse bias $I-V_{SD}$ characteristics, rescaled as described in Section S4. b, $I-V_G$ characteristics with the same rescaling shown at various values of $V_{SD}$. 
on the device current, which in turn influences the voltage drop across the junction. While this is not explicitly treated in our model, the details of this effect are captured qualitatively in the phenomenological factor $\alpha$. While a detailed treatment would require microscopic understanding of the interlayer transfer rates and the exact origins of $\alpha$, we exclude this subtle effect since $\alpha$ remains constant over the voltage regime of interest, thus allowing us to capture the basic temperature dependent behavior. At very low bias ($V_{SD} < 0.5$ V), the regime in which the current does not match the behavior in our model, $\alpha$ is difficult to extract since the current does not increase exponentially. As suggested in Chul-Ho Lee, et al.\textsuperscript{24}, this likely results from the lateral voltage drop or the non-exponential behavior due to recombination effects.

At very high voltage (above NDC), we observe an additional feature of carrier multiplication: an enhancement of the device noise due to impact excitation. At $T = 360$ K, the efficiency of impact excitation increases, leading to significant noise at high voltage. While this behavior may result from random current fluctuations due to carrier multiplication, it is not well described by the hot-carrier enhanced impact excitation model \textit{below} the threshold. Future work will focus closely on exploring the impact excitation noise factor to explore noise in the carrier multiplication process.

S5. Photoresponse of the Atomic Layer Hetero-Junction

For the assessment of the multiplication factor, we were very careful to photoexcite only within the WSe$_2$ absorption edge while avoiding optical absorption in the (overlapping) low energy tail of the MoSe$_2$ absorption peak. Absorption in MoSe$_2$ may result in complicated interlayer transfer processes and additional state filling that may skew the determination of the multiplication factor. From Supplementary Figures 2 and 3, we show that the heterostructure exhibits somewhat weak, but easily measured photoluminescence and photocurrent response at photon energy 1.38 eV (corresponding to $\lambda = 900$ nm). While this energy is slightly below, but still within the absorption peak of the $e_W$ transition, we chose this wavelength to avoid as much as possible any absorption in the low energy tail of the MoSe$_2$ absorption edge, which begins to become non-negligible as the photon energy increases above 1.4 eV (Figure S3).
Figure S7 shows photocurrent $I-V_{SD}$ characteristics taken at increasing photon energies and $T = 300\text{K}$ (in the regime of inefficient impact excitation) with the laser power fixed at $P = 17 \ \mu\text{W}$. While the overall photoresponse at 1.38 eV is comparatively weak, the photocurrent results primarily from absorption in WSe$_2$ rather than combined absorption within the absorption bands of WSe$_2$ and MoSe$_2$. As the absorption of MoSe$_2$ becomes important at higher energies, the overall photoresponse increases proportionally. At higher photon energies, a careful assessment of the multiplication factor is less dependable, as it must include the additional effects of current flow and interlayer transfer caused by absorption in both layers. Future work will examine the detailed photon energy dependence of impact excitation process by probing the intricate dependence of photocurrent on photon energy, voltage, and temperature.
Supplementary Fig. 7 | Wavelength dependence of the interlayer photocurrent-voltage characteristics at $T = 300$ K. Photocurrent (absolute value) vs. $V_{SD}$ measured as a function of decreasing wavelength (increasing photon energy) at fixed power $= 17$ $µ$W. Photon energies are $E_{ph} = 1.38$ eV, 1.45 eV, 1.57 eV, 1.62 eV, and 1.74 eV from bottom to top. For comparison, the black line trace is the same data shown in main text Figure 5a ($T = 300$ K, $E_{photon} = 1.38$ eV).
Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.
Supporting references and notes