Structure of the moiré exciton captured by imaging its electron and hole

Title: Structure of the moiré exciton captured by imaging its electron and hole

Authors: Ouri Karni[1]†, Elyse Barré[2,3]†, Vivek Pareek[4]†, Johnathan D. Georgaras[5]†, Michael K. L. Man[4]†, Chakradhar Sahoo[4]†, David R. Bacon[4], Xing Zhu[4], Henrique B. Ribeiro[1,2], Aidan L. O'Beirne[6], Jenny Hu[1], Abdullah Al-Mahboob[4], Mohamed M. M. Abdelrasoul[4], Nicholas S. Chan[4], Arka Karmakar[4], Andrew J. Winchester[4], Bumho Kim[7], Kenji Watanabe[8], Takashi Taniguchi[9], Katayun Barmak[10], Julien Madéo[4], Felipe H. da Jornada[5], Tony F. Heinz[1,2], and Keshav M. Dani[4]*

Affiliations:

[1] Department of Applied Physics, Stanford University; Stanford, California, 94305, USA.

[2] SLAC National Accelerator Laboratory; Menlo Park, California, 94025, USA.

[3] Department of Electrical Engineering, Stanford University; Stanford, California, 94305, USA.

[4] Femtosecond Spectroscopy Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son; Okinawa, Japan 904-0495.

[5] Department of Material Science and Engineering, Stanford University; Stanford, California, 94305, USA.

[6] Department of Physics, Stanford University; Stanford, California, 94305, USA.

[7] Department of Mechanical Engineering, Columbia University; New York, New York 10027, USA.

[8] Research Center for Functional Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan.

[9] International Center for Materials Nanoarchitectonics, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan.

[10] Department of Applied Physics and Applied Mathematics, Columbia University; New York, New York 10027, USA.

† These authors contributed equally to this work.

*Corresponding author. Email: kmdani@oist.jp

Introductory paragraph:

Interlayer excitons (ILXs) – electron-hole pairs bound across two atomically thin layered semiconductors – have emerged as attractive platforms to study exciton condensation,[1–4] single-photon emission and other quantum-information applications[5–7]. Yet, despite extensive optical spectroscopic investigations[8–12], critical information about their size, valley configuration and the influence of the moiré potential remains unknown. Here, in a WSe_2/MoS_2 heterostructure, we captured images of the time- and momentum-resolved distribution of *both* the particles that bind to form the ILX: the electron and the hole. We thereby obtain a direct measurement of both the ILX diameter of ~5.2 nm, comparable to the moiré unit-cell length of 6.1 nm, and the localization of its center-of-mass (COM). Surprisingly, this large ILX is found pinned to a region of only 1.8 nm diameter within the moiré cell, smaller than the size of the exciton itself. This high degree of localization of the ILX is backed by Bethe-Salpeter equation calculations and demonstrates that the ILX can be localized within small moiré unit cells. Unlike large moiré

cells, these are uniform over large regions, allowing the formation of extended arrays of localized excitations for quantum technology.

Main Text:

Heterobilayers made by stacking different monolayers (1L) of transition-metal dichalcogenide (TMDC) semiconductors support ILX states, where electrons and holes residing in separate layers bind to one another to form a neutral excitation[8]. The utilization of these long-lived and tunable excitations for quantum applications or for studying new many-exciton phases and interactions requires critical knowledge about the real- and momentum-space structure of the ILX: the valley character of its constituent electron and hole; its size, namely the extent of the electron distribution around its hole counterpart; and the degree of the ILX confinement within the moiré potential. These attributes are key to determining the nature of the light-matter interactions with the ILX, such as their direct bandgap character and their polarization selection rules, as well as the ability of the ILX to form uniform periodic arrays of localized quantum emitters in the moiré pattern[5]. Additionally, the ILX size and localizability determine their density-dependent many-body physics – at low densities strong COM confinement inhibits the formation of condensates[13], while at higher densities, the size and localization determine Mott dissociation thresholds[2] or the formation of multi-ILX complexes[3,4].

So far, studies on moiré-localized ILXs have been focused on large moiré unit cells due to the prevailing assumption that moiré confinement requires a moiré-period much larger than the ILX itself[9,14]. Such systems were studied by optical spectroscopy that could only access a narrow sector of the momentum-space distribution of the ILX wavefunction[9,10,12,15]. While such optical

spectra were qualitatively consistent with a theoretical picture of ILXs spatially localized by the moiré-potential, these experiments did not yield the ILX size, and often relied on phenomenological moiré potentials that were not expected from first-principles calculations[9,16]. The valley configuration of the electron and hole of the ILX also remained contested[9,17]. Furthermore, such systems exhibit significant inhomogeneities in the moiré lattice periodicity due to strain fields[18], and hence cannot support the formation of extended arrays of localized ILX that are the starting point of recent quantum technology proposals. Revealing those hidden properties together with a direct measurement of the extent of the ILX COM localization requires knowing the full momentum-space distributions of *both* of its constituent particles. Recent time- and angle-resolved photoemission spectroscopy experiments on microscopic samples (TR-μ-ARPES) measured the size of the exciton in 1L WSe_2 via the momentum distribution of only the exciton-bound electron[19–21]. However, accessing the coordinates of the exciton-bound hole has remained beyond reach. Conceptually, one needs to first relate the hole coordinates in momentum-space to the photo-induced reduction in the 2D valence band (VB) electron density. Experimentally, measuring this small, *momentum-resolved*, reduction needs high signal to noise, and high-quality samples with narrow initial VB linewidths. The short excitonic lifetimes also add to these challenges, as substantial thermal broadening of the VB spectra occurs at early time-delays after photoexcitation, obscuring the hole footprint.

Here we show unprecedented images of *both* the ILX-bound electron and hole in momentum-space in a TMDC heterobilayer exhibiting a relatively small moiré pattern. From these distributions, we determine the ILX properties along both its defining coordinates: the relative electron-hole separation, and their COM. We extract the ILX size and its COM confinement, and

find, surprisingly, that even though the ILX diameter is comparable to the moiré-period, its COM coordinate is tightly confined by the moiré-potential. Further, we observe the direct bandgap nature of the ILX with the electron and hole residing in the K-valleys of the two layers, and the anomalous negative dispersion of the photoemitted electron, establishing the excitonic origin of the signal. Finally, the hole images, acquired independently of the photoemission matrix elements, offer a direct quantitative measurement of the ILX density and exhibit a broadening in momentum distribution when their density surpasses one ILX per moiré site, akin to excitons in quantum dots. In all, these findings promote the prospect of using small-period moiré patterns, which are homogeneous and more robust against strain than the large moiré lattices[18,22], to host arrays of quantum emitters.

Unperturbed heterobilayer band-structure

The studied heterobilayer consisted of 1L WSe_2 placed on top of 1L MoS_2 (Fig. 1d) with a relatively small moiré pattern period of 6.1 nm, based on the lattice mismatch and the crystallographic twist-angle between the layers[5]. Confirming the coupling between the WSe_2 and MoS_2 layers, the photoluminescence (PL) spectrum from the sample (Fig. 1b) exhibits an ILX peak at 1 eV[11], and its optical reflection spectrum exhibits the expected moiré pattern signatures: a splitting-up of the absorption feature associated with WSe_2 excitons into several resonances[16] (Extended Data Fig. 1).

We first characterized the unperturbed band-structure of the heterobilayer at 100 K using our TR-μ-ARPES capabilities[19,23], that are based on a photoelectron momentum microscope[24,25] (see Fig. 1a, Methods, and Extended Data Fig. 2 for apparatus details and distortion correction procedures). We observe a band-structure consistent with previous calculations (Fig. 2a)[11,26]. The

VB maxima (VBM) appear at the K valleys (defining the zero of our energy scale), with both the spin-orbit split bands of WSe_2 clearly visible and separated by 0.5 eV. At lower energies (-0.85 eV) a single broad band is seen (Extended Data Fig 3), corresponding to the two spin-orbit split VBs of MoS_2 that are unresolvable in our measurement due to inhomogeneous broadening. Figure 2b presents a view of the band-structure in 2D momentum space integrated over a small energy range (-0.2 to 0.2 eV) around the VBM. We clearly see the photoemission from the Γ, the K, and the K' valleys. The variations in photoemission intensity across the image originate from the differing photoemission matrix elements[27], which depend sensitively on the incidence angle and the polarization of the XUV pulse[28].

Signatures of the ILX in TR-ARPES

Having identified the signatures of the unperturbed heterobilayer, we turned to study the ILX and the momentum-resolved distributions of its electron and hole by photoexciting the sample resonantly at the A-exciton in WSe_2 with a 170 fs, 1.67 eV pump pulse. A rapid transfer of the electron from WSe_2 to MoS_2 results in the formation of the ILX (schematic in Fig. 1c). After a time-delay (Δt) of a few tens of picoseconds, we applied the XUV probe pulse to record the momentum- and energy-resolved photoemission spectra from the photoexcited heterobilayer after the ILX cooled and reached quasi-equilibrium.

The signatures of the ILX are immediately recognized (Fig. 2c, e) in the TR-μ-ARPES data: Photoemitted electrons appear in all the corners of the BZ 1eV above the VBM, in agreement with the PL associated with the ILX. It also appears at a distinctly lower energy than the free conduction band (CB) electrons that can be seen when exciting above the bandgap of MoS_2 and WSe_2 (Extended Data Fig. 5). Furthermore, the dispersion curve of the photoemitted signal at 1

eV clearly shows a negative curvature, similar to the negative dispersion of the K (and K') valley VB of WSe_2 (see Fig. 2f and Extended Data Fig. 4). The observation of this anomalous dispersion at long time delays is exclusively a hallmark of the excitonic origin of the ARPES response[19,29,30]. In contrast, similar signatures observed at short time delays can be associated with other phenomena such as Bloch-Floquet effects[31,32], or Shockley surface states[33]. Besides this signal, we observe a striking depletion of the ARPES signal around the WSe_2 VBM. This depletion, seen in all K and K' valleys (Fig. 2d), corresponds to the holes created in WSe_2 as part of the ILX (See Methods for supporting evidence). This momentum-resolved distribution of the hole within an exciton provides a new route to understand excitons, well beyond recent experiments that measure only the properties of the electron[19].

To obtain the quantitative 2D momentum distributions of the holes, independent of the k-dependence of the photoemission matrix elements, we compared the TR-μ-ARPES spectra taken before and after the excitation (See Methods and Extended Data Fig. 6). The long lifetime of the ILX allows making this comparison at long time delays, by when the VB linewidths are comparable to the pre-excitation values, allowing the accurate extraction of the exciton-bound hole momentum-space distribution. At long delays, when the excitons are in quasi-equilibrium at the lattice temperature, the depleted signal at wavevector $\mathbf{k}$ in the VB provides the probability of finding a hole with momentum $-\hbar\mathbf{k}$ (See Supplementary Material for the derivation). We note that integrating this quantity over all k-space provides a direct measurement of the ILX density at any given delay (Fig. 3d) – an important quantity that can only be roughly estimated in time-resolved optical experiments. In addition to momentum-space images of the hole distribution, we also obtained the momentum-resolved images of the electron distribution directly from the

photoexcited TR-μ-ARPES spectrum[19]. We extracted quantitative parameters at different time delays by fitting these momentum-space electron and hole images to Gaussian distributions (Fig. 3a).

One of the extracted parameters is the Gaussian widths of these momentum-space distributions, which provide important information about the ILX. This can be understood by expressing the wavefunction of an idealized ILX, unconfined by any moiré potential, and with COM momentum $\hbar\mathbf{Q}$, as: $|X^{\mathbf{Q}}\rangle = \sum_{vc\mathbf{k}} A^{\mathbf{Q}}_{vc\mathbf{k}} c^{\dagger}_{c,\mathbf{k}+\alpha\mathbf{Q}} c_{v,\mathbf{k}-\beta\mathbf{Q}}|0\rangle$, where $\alpha = \frac{m_e}{m_e+m_h}$ and $\beta = \frac{m_h}{m_e+m_h}$. Here $A^{\mathbf{Q}}_{vc\mathbf{k}}$ is the envelope function of the unconfined ILX, $c^{\dagger}_{n,\mathbf{k}}$ is the creation operator of an electron in band *n* with wavevector **k**, and $|0\rangle$ is the un-photoexcited ground state of the system[29]. For $\mathbf{Q} = 0$, the momentum distributions of the electron and hole are identical and directly provide the size of the exciton, *i.e.*, the relative electron-hole coordinate in real-space (Fig. 3b)[19,29,30]. For finite **Q**, the electron and hole distributions are mainly offset in momentum by different amounts, $\alpha\mathbf{Q}$ and $\beta\mathbf{Q}$, respectively. Thus, a distribution of excitons with finite **Q** yields electrons and holes displaying distinct momentum distributions (Fig. 3c). In general, there are two processes that may lead to a distribution of excitons with finite **Q**: thermal effects and spatial confinement. For sufficiently low temperatures – smaller than the energy gap between different moiré-exciton bands, yet larger than the moiré-exciton bandwidth – thermal effects contribute a small and temperature-independent broadening that can be neglected (see Supplementary Material). This is indeed the case in our experiments, as the extents of the electron and hole distributions display no meaningful temperature dependence (see Extended Data Fig. 7). In contrast, the moiré potential localizes the COM coordinate of the ILX, resulting in a distribution of finite values of **Q**, which is encoded in the inequivalent electron and hole

momentum distribution. We note that the above discussion assumes low-enough exciton densities to avoid multi-exciton and state-filling effects, and sufficiently long time-delays to allow the excitons to cool to the lattice temperatures. In our measurements, these conditions are achieved at photoexcitation densities below 3×10^{12} cm^{-2} (Fig. 3e) and after a few tens of picoseconds (Fig. 3f). The density range corresponds to less than one ILX per moiré cell, in accord with the density threshold per moiré site indicated by previous PL reports of ILX localization[10]. The relaxed ILX then exhibits a hole distribution width of $\sigma_k^h = 0.042 \pm 0.002$ Å^{-1}, and an electron distribution width of $\sigma_k^e = 0.061 \pm 0.003$ Å^{-1}.

Structure of the ILX

To deduce the ILX size and spatial confinement from the measured momentum distributions of the electron and hole, we express the moiré-localized ILX $|X_m\rangle$ as a linear combination of unconfined ILXs $|X^{\mathbf{Q}}\rangle$ with different COM momenta $\hbar\,\mathbf{Q}$=$\hbar\,\mathbf{G}$, $|X_m\rangle = \sum_{\mathbf{G}} C(\mathbf{G})|X^{\mathbf{G}}\rangle$, where $\mathbf{G}$ is a reciprocal moiré lattice vector and $C(\mathbf{G})$ are expansion coefficients. The measured intensity distributions of the ILX-bound electron and hole, $I_e(\mathbf{k})$ and $I_h(\mathbf{k})$, are expressed in terms of the expansion coefficients $A^{\mathbf{Q}}_{vc\mathbf{k}}$ and $C(\mathbf{G})$, as $I_e(\mathbf{k}) = \sum_{\mathbf{G}}\left|C(\mathbf{G})A^{\mathbf{G}}_{\mathbf{k}-\alpha\mathbf{G}}\right|^2$ and $I_h(\mathbf{k}) = \sum_{\mathbf{G}}\left|C(\mathbf{G})A^{\mathbf{G}}_{\mathbf{k}+\beta\mathbf{G}}\right|^2$. For Wannier-like excitons, the simultaneous measurement of these distributions allowed us to extract both $A^{\mathbf{Q}=\mathbf{0}}_{vc\mathbf{k}}$ and $C(\mathbf{G})$. Their Fourier transforms provided the envelope functions of the relative electron-hole separation $\boldsymbol{r}$ as $\psi(\boldsymbol{r})$, and the COM coordinate $\mathbf{R}$ as $C(\mathbf{R})$, respectively. With this, one can approximate the envelope of the two body exciton wavefunction as a product of $\psi(\boldsymbol{r})$ and $C(\mathbf{R})$, with the real-space exciton radius and its COM localization being inversely proportional to the extent in momentum-space of $\left|A^{\mathbf{Q}=\mathbf{0}}_{vc\mathbf{k}}\right|^2$ and

$|C(\mathbf{G})|^2$, respectively (see Supplementary Material). Accordingly, we deduce an experimental ILX radius (*i.e.*, the root-mean-square (RMS) radius of its relative electron-hole separation distribution) of 2.6 ± 0.4 nm (5.2 ± 0.8 nm diameter), consistently larger than the 1.4 nm radius of the intra-layer exciton in WSe_2[19]. We also obtain an RMS localization of the COM distribution of 0.9 ± 0.1 nm, revealing an ILX whose COM coordinate is rather tightly pinned to the minimum of the moiré potential landscape. The higher temperature of our moiré confinement in MoS_2/Wse_2 heterobilayers, compared with previous results on $MoSe_2/Wse_2$ heterobilayers[10,34], is attributed to the former's band-structure, where the Q and K CB edges are spectrally further apart[11,17], leading to enhanced robustness of the moiré-localized ILX.

Our experimental results on the spatial size and localization of the ILX in a well-characterized sample motivate us to seek deeper physical insight into the nature of moiré-confined ILXs using state-of-the-art many-body perturbation theory calculations. We have accordingly carried out a set of first-principles and effective-Hamiltonian calculations. We first performed *ab initio* GW plus Bethe-Salpeter equation calculations[35,36] of the $\mathbf{Q} = 0$ ILX wavefunction distribution in an artificially commensurate MoS_2/Wse_2 heterobilayer to obtain $A_{vc\mathbf{k}}^{\mathbf{Q}=\mathbf{0}}$. This yielded an RMS radius of 2.4 nm (4.8 nm diameter) in real-space (See Fig. 4a). A further calculation employing an effective moiré ILX Hamiltonian[9,14] yielded a COM coordinate that is localized in real-space to within an RMS radius of 1.3 nm around the moiré potential minimum (Fig. 4b). Our calculations show that both $\left|A_{vc\mathbf{k}}^{\mathbf{Q}=\mathbf{0}}\right|^2$ and $|C(\mathbf{G})|^2$ are well-described by Gaussian distributions, and both the ILX size and its COM localization are in reasonable agreement with our experiment (see Methods for details). These conclusions remain unaltered when going beyond the effective moiré ILX Hamiltonian, as described in the Supplementary Material.

Discussion

Our simultaneous measurement of the electron and hole momentum distributions in a bound state provides previously inaccessible information about both the relative and COM coordinates that characterize the ILX two-body wavefunction. A measurement of the relative coordinate of the ILX provides, for the first time, its size – a critical parameter in determining the nature and thresholds of various many-exciton processes, such as the Mott transition[37] and exciton-exciton annihilation[38]. The COM coordinate of the exciton also plays an important role in its underlying physics. The observed localization, taken together with the onset of momentum-space broadening as the ILX density crosses one per moiré cell, resembles the behavior of excitons in impurity-based quantum dots[3,4]. Such observations support emerging schemes in quantum technology that utilize an array of quantum-dot-like states created by the periodic moiré potential[5,7]. Compared to larger moiré periods, our use of lattice-mismatched heterobilayers with a small moiré period possesses the crucial benefit of extended spatial homogeneity and robustness against strain fields[18]. Our observations also raise interesting questions about how the moiré period impacts exciton-exciton interactions and moiré-localization. Lastly, the COM distribution provides direct access to the temperature of an ensemble of excitons. Combined with the ability to read the exciton density directly from the data, our measurements provide an important new tool to study many-body excitonic states, their interactions with correlated electrons[39], and the corresponding phase diagrams.

References

1. Wang, Z. *et al.* Evidence of high-temperature exciton condensation in two-dimensional atomic double layers. *Nature* **574**, 76–80 (2019).

2. Wang, J. *et al.* Optical generation of high carrier densities in 2D semiconductor heterobilayers. *Sci. Adv.* **5**, eaax0145 (2019).
3. Li, W., Lu, X., Dubey, S., Devenica, L. & Srivastava, A. Dipolar interactions between localized interlayer excitons in van der Waals heterostructures. *Nat. Mater.* **19**, 624–629 (2020).
4. Kremser, M. *et al.* Discrete interactions between a few interlayer excitons trapped at a MoSe 2 –WSe 2 heterointerface. *Npj 2D Mater. Appl.* **4**, 1–6 (2020).
5. Yu, H., Liu, G.-B., Tang, J., Xu, X. & Yao, W. Moiré excitons: From programmable quantum emitter arrays to spin-orbit–coupled artificial lattices. *Sci. Adv.* **3**, e1701696 (2017).
6. Liu, X. & Hersam, M. C. 2D materials for quantum information science. *Nat. Rev. Mater.* **4**, 669–684 (2019).
7. Kennes, D. M. *et al.* Moiré heterostructures as a condensed-matter quantum simulator. *Nat. Phys.* **17**, 155–163 (2021).
8. Rivera, P. *et al.* Interlayer valley excitons in heterobilayers of transition metal dichalcogenides. *Nat. Nanotechnol.* **13**, 1004 (2018).
9. Tran, K. *et al.* Evidence for moiré excitons in van der Waals heterostructures. *Nature* **567**, 71 (2019).
10. Seyler, K. L. *et al.* Signatures of moiré-trapped valley excitons in MoSe 2 /WSe 2 heterobilayers. *Nature* **567**, 66 (2019).
11. Karni, O. *et al.* Infrared Interlayer Exciton Emission in MoS2/WSe2 Heterostructures. *Phys. Rev. Lett.* **123**, 247402 (2019).
12. Zhang, L. *et al.* Highly valley-polarized singlet and triplet interlayer excitons in van der Waals heterostructure. *Phys. Rev. B* **100**, 041402 (2019).

13. Fogler, M. M., Butov, L. V. & Novoselov, K. S. High-temperature superfluidity with indirect excitons in van der Waals heterostructures. *Nat. Commun.* **5**, 4555 (2014).
14. Wu, F., Lovorn, T. & MacDonald, A. H. Theory of optical absorption by interlayer excitons in transition metal dichalcogenide heterobilayers. *Phys. Rev. B* **97**, 035306 (2018).
15. Jin, C. *et al.* Identification of spin, valley and moiré quasi-angular momentum of interlayer excitons. *Nat. Phys.* **15**, 1140–1144 (2019).
16. Jin, C. *et al.* Observation of moiré excitons in WSe 2 /WS 2 heterostructure superlattices. *Nature* **567**, 76 (2019).
17. Gillen, R. & Maultzsch, J. Interlayer excitons in MoSe2/WSe2 heterostructures from first principles. *Phys. Rev. B* **97**, 165306 (2018).
18. Shabani, S. *et al.* Deep moiré potentials in twisted transition metal dichalcogenide bilayers. *Nat. Phys.* **17**, 720–725 (2021).
19. Man, M. K. L. *et al.* Experimental measurement of the intrinsic excitonic wave function. *Sci. Adv.* **7**, eabg0192 (2021).
20. Dong, S. *et al.* Direct measurement of key exciton properties: Energy, dynamics, and spatial distribution of the wave function. *Nat. Sci.* **1**, e10010 (2021).
21. Wallauer, R. *et al.* Momentum-Resolved Observation of Exciton Formation Dynamics in Monolayer WS2. *Nano Lett.* **21**, 5867–5873 (2021).
22. Li, H. *et al.* Imaging moiré flat bands in three-dimensional reconstructed WSe2/WS2 superlattices. *Nat. Mater.* **20**, 945–950 (2021).
23. Madéo, J. *et al.* Directly visualizing the momentum-forbidden dark excitons and their dynamics in atomically thin semiconductors. *Science* **370**, 1199–1204 (2020).

24. Schönhense, G., Medjanik, K. & Elmers, H.-J. Space-, time- and spin-resolved photoemission. *J. Electron Spectrosc. Relat. Phenom.* **200**, 94–118 (2015).
25. Medjanik, K. *et al.* Direct 3D mapping of the Fermi surface and Fermi velocity. *Nat. Mater.* **16**, 615–621 (2017).
26. Kunstmann, J. *et al.* Momentum-space indirect interlayer excitons in transition-metal dichalcogenide van der Waals heterostructures. *Nat. Phys.* **14**, 801 (2018).
27. Moser, S. An experimentalist's guide to the matrix element in angle resolved photoemission. *J. Electron Spectrosc. Relat. Phenom.* **214**, 29–52 (2017).
28. Beaulieu, S. *et al.* Revealing Hidden Orbital Pseudospin Texture with Time-Reversal Dichroism in Photoelectron Angular Distributions. *Phys. Rev. Lett.* **125**, 216404 (2020).
29. Rustagi, A. & Kemper, A. F. Photoemission signature of excitons. *Phys. Rev. B* **97**, 235310 (2018).
30. Ohnishi, H., Tomita, N. & Nasu, K. Direct determination of exciton wavefunction amplitudes by the momentum-resolved photo-electron emission experiment. *Int. J. Mod. Phys. B* **32**, 1850094 (2017).
31. Wang, Y. H., Steinberg, H., Jarillo-Herrero, P. & Gedik, N. Observation of Floquet-Bloch States on the Surface of a Topological Insulator. *Science* **342**, 453–457 (2013).
32. Mahmood, F. *et al.* Selective scattering between Floquet–Bloch and Volkov states in a topological insulator. *Nat. Phys.* **12**, 306–310 (2016).
33. Cui, X. *et al.* Transient excitons at metal surfaces. *Nat. Phys.* **10**, 505–509 (2014).
34. Liu, E. *et al.* Signatures of moiré trions in WSe2/MoSe2 heterobilayers. *Nature* **594**, 46–50 (2021).

35. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B* **34**, 5390–5413 (1986).
36. Rohlfing, M. & Louie, S. G. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B* **62**, 4927–4944 (2000).
37. Meckbach, L., Stroucken, T. & Koch, S. W. Giant excitation induced bandgap renormalization in TMDC monolayers. *Appl. Phys. Lett.* **112**, 061104 (2018).
38. Sun, D. *et al.* Observation of Rapid Exciton–Exciton Annihilation in Monolayer Molybdenum Disulfide. *Nano Lett.* **14**, 5625–5629 (2014).
39. Wu, F., Lovorn, T., Tutuc, E. & MacDonald, A. H. Hubbard Model Physics in Transition Metal Dichalcogenide Moir\'e Bands. *Phys. Rev. Lett.* **121**, 026402 (2018).

Main figure legends:

Fig. 1 - Sample structure and experimental configuration. a. Schematic representation of the TR-μ-ARPES setup and sample structure. The XUV probe pulse impinges on the sample at a delay Δt relative to the optical pump pulse. Photoemitted electrons from the excited sample are filtered through a micron-scale spatially selective aperture and resolved in energy and momentum by a 3D time-of-flight detector. We obtained a volumetric dataset of the band-structure of the entire Brillouin zone (BZ) with an energy range covering the relevant VB and the photoexcited electron features (Top). **b.** Low temperature (10 K) infrared PL of the 1L MoS_2 (yellow), WSe_2 (red), and the heterobilayer (blue). **c.** Schematic of the band-edge alignment of the layers and the electron-transfer process leading to formation of the ILX. **d**. The layer arrangement of the sample heterostructure. The heterobilayer is isolated from the conductive silicon substrate by a thin hexagonal boron nitride layer (see Methods).

Fig. 2 - Static and photoexcited TR-μ-ARPES measurements. a, c. Energy-momentum cuts along the K-Γ direction without (a) and 25 ps after (c) optical excitation. **b, d, e.** Momentum-space images around VBM (b), as well as 25 ps after photoexcitation around VBM (d) and around the ILX energy (e)**.** These ARPES intensity maps are formed by integrating over the energy ranges defined by the blue brackets in (a) and (c). Clear signatures of the ILX-bound electron and hole at 1 eV and at the VBM, respectively, can be seen after photoexcitation (c, d and e) in all K valleys. In (b), the dashed line marks the K_3-Γ-K_6 cut used in (a) and (c). In (c), the counts above 0.85 eV are enhanced by a factor of 50 for visibility. **f.** Emission from the ILX-bound electron exhibiting the anomalous negative dispersion [19] of the VB. This establishes the bound nature of the observed electrons and holes. Inset: Line-cuts of the dispersion of the ILX-bound electron (dashed) and VB (continuous) along k_x axis.

Fig 3 – Time- and momentum-resolved distributions of the ILX electrons and holes. a. Distributions around the K valley (K_6 from Fig. 2). Point clouds and heat maps represent the data. Teal surfaces are Gaussian fits. The coordinate (k_x,k_y)=(0,0) is set to the VB maximum. Delay values are indicated above each panel. The normalized photoexcited electron distribution (top row) exhibits a much larger width than the hole distribution (bottom row). **b.** Schematic of the momentum distributions (yellow) of the electron and hole constituting an unconfined exciton with zero COM momentum, showing identical extents. The black curves indicate the CB and VB with their different curvatures. **c.** Similar schematic for a confined exciton, composed of multiple non-zero Q states (blue, yellow, and red). The electron and hole each carry only a fraction of Q that is proportional to their respective effective mass, resulting in unequal momentum distributions as shown. **d.** Hole density, indicative of ILX density, as function of delay. The exponential fit yields a lifetime of 115 ± 34 ps, much longer than that of intralayer excitons. **e.**

The quasi-equilibrium photoexcited electron and hole distribution widths at different densities, indicating the onset of a broadening at $\sim 3 \times 10^{12}$ cm^{-2}, corresponding to more than one ILX per moiré lattice. **f.** Width of the momentum distribution of the electron and hole versus time-delay showing the two distinct values for quasi-equilibrium at long-time delays. The Gaussian fits were performed to obtain estimation errors of 3% in the population, and 5% in the widths, as depicted by the error bars.

Fig. 4 - ILX wavefunctions in real-space. a. Modulus squared of the ILX wavefunction $|\psi(\mathbf{r}_e, \mathbf{r}_h)|^2$ (red) as a function of the electron coordinate $\mathbf{r}_e$, for the hole coordinate $\mathbf{r}_h$ fixed as indicated in black. The moiré confinement was not included in this plot. **b.** The calculated real-space envelope function squared for the COM coordinate of the ILX, $|C(\mathbf{R})|^2$, on top of the calculated spatial landscape of the moiré potential V(**R**).

Methods

Sample fabrication

1L MoS_2 (2D Semiconductors) and 1L WSe_2 (flux-grown, Columbia University, NY, USA) flakes were exfoliated onto transparent polydimethylsiloxane (PDMS) stamps and identified by their optical contrast and their photoluminescence (PL). To isolate the sample and the ILX from the conducting bare silicon substrate (with a 1-2 nm thick native oxide), while still preventing the sample from charging during the photoemission measurement, a flat and thin (18 nm, measured by AFM)) hexagonal boron nitride flake (hBN-National Institute for Material Science, Japan) was directly exfoliated on the substrate prior to stacking the heterobilayer[19,23] (see Fig. 1d). The heterostructure (clear area of 25x8 μm^2) was assembled by dry-transfer of the monolayer flakes onto the hBN, and then vacuum-annealed for 2 hr at 150° C. The crystallographic alignment of

1L WSe_2 and 1L MoS_2 was established judging by the sharp edges of the flakes. After fabrication, a polarization-resolved second-harmonic generation measurement determined the twist angle to be 2.2±0.8°. The lattice mismatch between the layers is 3.8 %.

Optical characterization of the samples

The PL spectra of the monolayers and their heterostructure were measured in a home-built confocal microscope setup. The sample was held at a controlled temperature in a cryostat (Montana Instruments). Laser light at 532 nm wavelength was used for excitation and was focused to a diffraction-limited spot (~1 μm) on the sample by a microscope objective (Olympus LCPLN50xIR). The infrared PL was collected by the same objective; it was detected and analyzed by a Jobin Yvon iHR550 spectrometer equipped with a Spectrum One InGaAs array. For reflection-contrast measurements (shown in Extended Data Fig. 1) white light from a stabilized tungsten-halogen light source (Thorlabs SLS201L) was delivered to the sample, and its reflection collected, through the same optical setup. To analyze the reflection spectrum in the visible range a Synapse CCD camera was employed as a detector.

Time- and angle-resolved photoemission spectroscopy

Time- and angle-resolved photoemission spectroscopy is performed by coupling ultrafast pulses with an angle-resolved, energy-resolved electron microscope (Metis 1000, SPECS GmbH) [19,23]. To monitor carrier dynamics, a pump-probe scheme has been developed with a tunable optical pump pulse that photo-excites the sample and an XUV probe pulse that photo-emits the electrons. Both pulses originate in a Yb-doped fiber laser producing 250 fs, 50 μJ pulses at a

wavelength of 1030 nm and a repetition rate of 2 MHz. The optical pump pulse, at 1.67 eV (740 nm), is produced by a non-colinear optical parametric amplifier (NOPA) pumped by part of the fundamental laser power. The high-energy XUV photons are produced by high-harmonic generation (HHG), achieved by frequency doubling the remaining of the fundamental laser light, and then focusing it into Kr gas. One of the high harmonics, at an energy of 21.67 eV is selected using a combination of Al and Sn filters. Both the pump and the probe are focused onto the region of interest on the sample. The pump energy was tuned in resonance with one of the absorption peaks of WSe_2 in the heterobilayer at the relevant experimental temperatures ranging between 80 K to 100 K. The probe energy is sufficient to photo-emit electrons with momenta covering the entire BZ of the sample, which emerge with different energies and at different angles. The photoelectrons are guided by the electrostatic lenses of the electron microscope and are detected by a time-of-flight detector, which measures the energy of the emitted electrons. Finally, the electrons are imaged either in real space (spatial- and energy-resolved) or in *k*-space (momentum- and energy-resolved) for different pump-probe time delays. The obtained data sets are corrected for the distortions in the apparatus using a two-stage procedure detailed in the following Methods section. More details of the TR-μ-ARPES instrumentation can be found in Ref. [19,23]

Distortion correction procedures in the momentum microscope

In the momentum microscope, the mapping of momentum-space images is done in two key elements. The first element is the cathode objective lens. As shown in Extended Data Fig. 2a at the front-end of the microscope, the objective lens, floating at a high potential, is used to collect photoemitted electrons from the sample and then to focus electrons emitted at different angles

onto the back focal plane, forming the angle resolved momentum space image. The second element is the rest of the imaging optics which serve to project a magnified momentum space image onto the time-of-flight detector. Perfect imaging optics should project an exact copy of the momentum space image with fixed magnification factor and without any distortion. However, in the actual system, aberrations in the lenses are unavoidable, which can induce radial distortion such as barrel and pincushion distortion to the projected images. In addition, imperfection in machining, mechanical misalignment, and external effects such as stray magnetic field can also induce translation and shearing distortion, scaling variation on different image directions, and other higher order nonlinear distortions. Below we introduce the two-part image processing routines that we use to remove such distortions from the acquired images.

1. Removing the distortions caused by the imaging optics in the momentum microscope. Visualizing these distortions is easily done by inserting a grid at the back focal plane of the objective lens. In Extended Data Fig. 2b we show an image of the grid overlaid on some band structure. The grid has a square pattern with known periodicity. We use it to calibrate the magnification of the imaging optics and remove distortion in the images. We generally make sure the alignment of the lenses in the microscope is at optimal setting, meaning the electrons are going through the optical axis of the lenses, ensuring minimal distortion due to misalignments. Observing the yellow square box in the image, we see that the distortion of the grid is quite small. Tiny amount of pincushion distortion is visible. The grids are also slightly elongated in the horizontal direction. To correct these image distortions, we apply a polynomial distortion model, commonly known as the Brown-Conrady model[40]. We utilize the regular pattern of the grid to obtain the correction coefficient of the model. After that we

apply geometric transformation to remap the distorted image into a rectified, undistorted output image (Extended Data Fig. 2c).

In general, the action of the lenses, and hence the distortion field, changes as a function of electron kinetic energy. The lens action could change very strongly when particular lens elements, such as the high pass energy filter, are active in the microscope. Extended Data Fig. 2d shows a cut of the 3D data cube, plotting the momentum k_x versus electron kinetic energy E_{kin}, in this image we can see strong variation of magnification in the momentum space image with respect to changing E_{kin}. To correct this, we deduce the distortion correction coefficient at different kinetic energies using the same algorithm as above and apply an energy dependent distortion correction on the 3D data cube. This procedure produces a distortion free image as shown in Extended Data Fig. 2e. This grid distortion correction procedure removes the intrinsic distortion originating from the lenses of the momentum microscope. It is sample independent and does not require any standard sample with known band structure to work properly. The procedure is required only when the microscope alignment is changed. Since the alignment grid can provide a lot of reference points for the distortion field, it also facilitates correction of high order distortion if necessary.

2. Removing distortions from the objective lens and sample alignment: In general, grid distortion correction on its own is sufficient to remove most of the image distortion we get in the momentum microscope. However, this procedure cannot correct distortions before the back focal plane of the objective lens, including distortion from the objective lens, sample tilt and inhomogeneous field from the sample. To account for these distortions, we use high symmetry features in the band structure of the measured sample and apply a method similar

to what has been discussed previously reported[41]. The positions of valence band (VB) maxima at high symmetry valleys, such as the K and Γ valleys in the measured data, provide seven reference points. Their positions are indicated by the maxima of the band dispersions to avoid the misleading influences of the nonuniform photoemission matrix element. The symmetry of the crystal tells us exactly where the valleys should be. With seven reference points, we can detect distortions such as shearing, or direction scaling, and eliminate them by geometric transformation. We also take care of any energy-dependent distortion by mapping out the distortion field at different electron kinetic energy. This is done by varying the sample bias voltage U_{bias}. The energy of the photoemitted electrons is giving by $E_{kin} = h\nu - E_b - W + U_{bias}$, where $h\nu$ is the photon energy, E_b is binding energy, and W is work function of the measured material. When we adjust U_{bias}, we effectively shift the kinetic energy of the photoemitted electrons entering the microscope. This lets us shift the kinetic energy of the VB edges. In our measurement, we are interested in an energy range from 3 eV above to -3 eV below the top of the VB edges. This range includes the exciton-bound electrons and part of the VBs. We shift the spectrum up and down by 1eV steps and take a minimum of seven reference points in energy. At each U_{bias}, we take a snapshot of the 3D ARPES image and determine the position of the high symmetry points. Based on the seven energy steps, we can interpolate the positions of the valleys and the distortion coefficients, and thus correct for any image distortion at any intermediate energies.

Procedure to extract band energy dispersions

In order to extract band dispersions either of the VB or of the exciton-bound electron feature the following procedure was followed:

1. For each point in [k_x, k_y], the energy spectrum of the measurement (known in other reports as Energy Distribution Curve, or EDC) over the relevant energy range was selected, integrating over a small environment of 3x3 k-pixels (in the case of VB) or 5x5 k-pixels (in the case of the ILX-bound electron band), to improve the signal-to-noise ratio, yet without changing our momentum-resolution significantly.

2. The spectra were fitted with Gaussian curves. In the VB, we assigned three Gaussians to the WSe_2 two spin-split VB and to MoS_2 VBs. For the ILX-bound electron band we used a single Gaussian. The Gaussians peak energies and their variation across the k-plane represent the extracted dispersions. Extended Data Fig. 4 shows the result for the ILX-bound electron band. Three spectra are presented, alongside their Gaussian fits, and the resulting peak energy curve as function of the k-vector, around the center of the exciton distribution. Alternatively, instead of this peak energy, one can also use the spectral center-of-mass of the EDC in the relevant energy range (Fig. 2f), yielding near identical results.

<u>Supporting evidence for the ILX origin of the photoexcited electron and hole</u>

The main evidence that tie the photoexcited electron and hole to the ILX are the match of their energy difference to the ILX PL peak, the negative curvature of the electron band dispersion (Extended Data Fig. 4), and the long lifetime. We further support this assignment as follows:

1. Following an above band gap excitation, at 2.5 eV, which excites both constituent layers of our heterobilayer, we initially observe a broad cluster of photoexcited electrons at energies above 1.3 eV (See Extended Data Fig. 5a and b). With clear signs of normal, CB-like,

dispersion, these are identified as unbound electrons in the CB. In contrast, after 50 ps, the excited electrons look differently (Extended Data Fig. 5c and d): they are concentrated at the K point, at 1 eV energy, and show an oppositely curved dispersion (Extended Data Fig. 4). These differences demonstrate clearly that the observed electrons are now bound to holes as excitons.

2. At the same delay, a careful examination of the VBs signals in comparison to their equilibrium spectrum reveals the holes at the top, WSe_2, VB, whereas the VBs associated with MoS_2 show no reduction in counts (Extended Data Fig. 5d). Thus, it is clear that the exciton-bound electrons 1 eV above the VB maximum (VBM), associated with MoS_2 (they are too low in energy to be located in WSe_2, whose intralayer excitonic features sit above 1.5 eV, including defect states[42]), must be bound to the WSe_2 holes. This fulfils the definition of this state as the interlayer exciton.

3. From our Gaussian fit procedure, the centers of the electron and hole distributions are found with a relative shift in momentum (See Extended Data Fig. 5e and f). It is consistent with the ILX assignment of these distributions, since the ILX-bound electrons and holes are expected to be offset in momentum depending on the lattice and orientation mismatch between the layers and the corresponding band extrema. We note that this shift does not affect our conclusions about the ILX spatial attributes as these depend only on the measured momentum width of the electron and hole distributions.

<u>Procedure to deduce the hole momentum distribution from the experimental data</u>

549 This procedure is a series of steps:

550 1. Separate the ARPES signal from the highest VB from that of the other VBs:

551 The photo-excited holes appear only at the highest-lying VB. To obtain a clear picture of
552 their distribution in momentum space, it is necessary to compare the readings of that band
553 with and without photoexcitation, independently of the other VBs. To that end, the highest
554 band energy and counts for each (k_x, k_y) in a region of interest (0.2 $Å^{-1}$ x 0.2 $Å^{-1}$ environment
555 around the K_6 point) are extracted by fitting the local photoemission spectrum with Gaussian
556 peaks (shown in Extended Data Fig. 6a and b for the unexcited case, and Extended Data Fig.
557 6c and d for the excited case). The highest energy peak is assigned to the upper VB, and its
558 count rate and energy are used thereafter.

559 2. Correcting the momentum offset between measurements:

560 Using the dispersion relation as determined above, we align the position of the band edge in
561 both the excited and un-excited data sets. We do so by fitting the dispersion with a two-
562 dimensional paraboloid around the center of the window of interest.

563 3. Computing the hole distribution:

564 Sharing the same coordinate grid in k-space, we can correct for variations in the experimental
565 conditions between measurements by normalizing the data by the count rate at an area that is
566 unaffected by the photoexcitation (namely, outside the hole feature). This environment is
567 marked by red squares in Extended Data Fig. 6e and f, showing the counts of the top VB in
568 our region of interest, for the un-excited and excited measurements, respectively. Subtracting

the ratio of normalized excited and un-excited ARPES signals from one yields the hole occupancy distribution:

$$f[k_x, k_y] = 1 - \frac{C_{un-pump}[k_{x0}, k_{y0}]}{C_{pump}[k_{x0}, k_{y0}]} \frac{C_{pump}[k_x, k_y]}{C_{un-pump}[k_x, k_y]}$$

Such a distribution is demonstrated in Extended Data Fig. 6g, as well as in the distributions presented (for different delay points) in Fig. 3a of the main text. As both pumped and unpumped images have similar photoemission matrix element profiles across the region of interest, their ratio removes the dependence of the resulting hole distribution on the photoemission matrix element.

Calculating ILX density from the data

The actual hole (and therefore, ILX) density is given by multiplying their occupancy distribution, $f[k_x, k_y]$, with the density of hole states in momentum and summing over the entire distribution:

$$n = \frac{2}{L^2} \iint f[k_x, k_y] \frac{dk_x dk_y}{\frac{4\pi^2}{L^2}}$$

$dk_{x/y}$ is the k coordinate grid spacing in the $k_{x/y}$ directions. In our data $dk_x = dk_y = 0.0066$ Å^{-1}. The prefactor of 2 stands for the two valleys that share the BZ.

Ab initio calculations

We performed our mean-field density-functional theory (DFT) calculations with a Perdew-Burke-Ernzerhof exchange-correlation functional using the Quantum Espresso code [43]. The calculations were done on a unit cell in the plane-wave basis using the Optimized Norm-Conserving Vanderbilt pseudopotentials from the SG-15 dataset [44]. We first considered a unit cell for an artificially periodic MoS_2/WSe_2 heterobilayer, for which we take the lattice constant

as the average of the individual MoS_2 and WSe_2 experimental lattice constants, 3.22 Å. The distance between repeated unit cells in the out-of-plane direction was taken as 25 Å and the Coulomb interaction was truncated [45] to effectively remove interaction between the repeated simulation cells along the out-of-plane direction.

We relaxed the atomic positions for the MoS_2/WSe_2 heterostructure in the R_h^X stacking [5] including van der Waals interactions within the frameworks of Dione *et al*. [46] and Cooper [47], and obtained a separation of 6.2 Å between the center of MoS_2 and WSe_2 layers. The dielectric matrices used to solve the Bethe-Salpeter equation (BSE) to obtain excitonic states were evaluated up to a cutoff of 6 Ry on a 90x90x1 Monkhorst-Pack *k*-point grid and computed with the BerkeleyGW software package [48].

To solve for the BSE, we first evaluated the electron-hole interaction kernel on a coarse 90x90x1 *k*-grid with a fully relativistic framework, which was then interpolated onto a patch of a 600x600x1 *k*-point grid around the K-valley. We diagonalized the BSE Hamiltonian with the BerkeleyGW package keeping four CBs and four VBs. Similarly to previous calculations[49], the sample patch only includes the 9593 *k*-points within 0.2 $Å^{-1}$ of the K point of the BZ and properly includes all the qualitative features of the desired ILXs.

Model-Hamiltonian calculations of the exciton confinement

In the absence of any moiré effects, the excitons that we computed from first principles in the previous section can be labeled with respect to a principal band and wavevector (**Q**) quantum

numbers. To a good approximation, because of vanishingly small exchange terms for ILX, these excitons can be described with an effective-mass approximation. The energy of the lowest moiré-unconfined ILX is given by $E_{\mathbf{Q}} = E_0 + \hbar^2 Q^2/(2M)$, where E_0 is the energy of the lowest ILX and M is the total exciton mass. However, the presence of a moiré pattern with a periodicity larger than the lattice constants of the constituent monolayers allows for multiple excitons with distinct wavevectors (and, to lesser extent, band indices) to mix. This mixing can be rationalized in terms of an effective moiré potential, $\Delta(\mathbf{R})$, that describes the local variation of the bandgap across the moiré unit cell stemming from a different atomic registry at each point **R**.

We follow refs. [14] and [9] and describe the effective Hamiltonian for an ILX with wavevector **q** inside the moiré BZ that experiences a moiré potential as

$$H_{\mathbf{q}} = E_0 + \frac{\hbar^2 |\mathbf{q} - \kappa|^2}{2M} + \Delta(\mathbf{R}),$$

where the momentum offset of κ (κ is the K point of the hexagonal moiré BZ) is due to the finite twist angle between the two individual TMDC layers [50].

The moiré potential energy is a smooth periodic function with extrema at specific high-symmetry points of the moiré lattice and can be approximated in the lowest-order harmonic expansion [9,14] as

$$\Delta(\mathbf{R}) \approx \sum_{j=1}^{6} V_j \exp\left(i\mathbf{G}_j \cdot \mathbf{R}\right),$$

where $\mathbf{G}_j$ are the first-shell reciprocal lattice vectors of the moiré BZ [9,14]. We first fit the potential parameters V_j by performing DFT calculations on a series of high-symmetry stacking configurations and obtain a total potential width (maximum-to-minimum) $\Delta(\mathbf{R})$ of ~100 meV,

similarly to previous calculations[5,51], and rescale it to 125 meV to account for the underestimate of DFT-derived potentials[51]. The potential parameters are $\left|V_j\right|_{j=1..6} = 12.5$ meV.

The model Hamiltonian expressed as such is solved in a plane-wave basis, where we retain moiré reciprocal-lattice vectors $\mathbf{G}$ with maximum magnitude of $|\mathbf{G}| = 10\mathbf{G}_1$. For each moiré-confined ILX with wavevector $\mathbf{q}$, we obtain the ILX envelope function $C_{\mathbf{q}}(\mathbf{G})$, as defined in the main text, by diagonalizing $H_{\mathbf{q}}$. The envelope function squared of the lowest-energy eigenstate can be approximated as an effective 2D Gaussian distribution whose width $\sigma_{\mathrm{COM}}^{r}$ can be deduced from its momentum distribution – more specifically, from the variance $\langle|\mathbf{G}|^2\rangle_{\mathbf{q}=\kappa} \equiv \sum_{\mathbf{G}} |\mathbf{G}|^2 \left|C_{\mathbf{q}=\kappa}(\mathbf{G})\right|^2 = 2\sigma_{\mathrm{COM}}^2$. We find numerically $\sigma_{\mathrm{COM}} \approx 0.053$ Å^{-1}. At 0 K, this width is directly related to the reciprocal of the COM localization radius (see Supplementary Material).

To account for thermal effects, we must consider the momentum distribution of $\left|C_{\mathbf{q}}(\mathbf{G})\right|^2$ for a range of values of $\mathbf{q}$ in the moiré BZ, the variance of which is related to an effective, thermally averaged width $\sigma_{\mathrm{COM},T}$. Because we are in a regime where $W < k_b T < \Delta E$, where W is the bandwidth of the lowest-energy moiré exciton (~ 2 meV, from our calculations), T is the temperature, and ΔE is the energy difference between the first and second moiré-exciton bands (~ 30 meV, from our calculations), we may assume that the lowest exciton band across the moiré BZ is uniformly occupied while higher bands are empty. From our numerical calculations, we find that finite-temperature effects increase σ_{COM} to a value $\sigma_{\mathrm{COM},T}$ that is less than 10% larger. Furthermore, for sufficiently large σ_{COM}, one can approximate the thermal effects as

652 $\sigma^2_{\mathrm{COM},T} \approx \sigma^2_{\mathrm{COM}} + \sigma^2_T$, where $\sigma_T = \sqrt{7/6}\,|\mathbf{G}_1|/4 \approx 0.032$ Å^{-1} is an effective thermal broadening
653 contribution to the width of the COM distribution in reciprocal space and which only depends on
654 the geometry of the moiré BZ.
655
656
657
658
659 **Additional Methods references:**
660

661 40. Brown, D. Decentering distortion of lenses.
662 https://www.semanticscholar.org/paper/Decentering-distortion-of-lenses-
663 Brown/2ef001c656378a1c5cf80488b35684742220d3f9 (1966).
664 41. Xian, R. P., Rettig, L. & Ernstorfer, R. Symmetry-guided nonrigid registration: The case for
665 distortion correction in multidimensional photoemission spectroscopy. *Ultramicroscopy* **202**,
666 133–139 (2019).
667 42. Ye, Z. *et al.* Efficient generation of neutral and charged biexcitons in encapsulated WSe2
668 monolayers. *Nat. Commun.* **9**, 3718 (2018).
669 43. Giannozzi, P. *et al.* QUANTUM ESPRESSO: a modular and open-source software project
670 for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
671 44. Schlipf, M. & Gygi, F. Optimization algorithm for the generation of ONCV
672 pseudopotentials. *Comput. Phys. Commun.* **196**, 36–44 (2015).
673 45. Ismail-Beigi, S. Truncation of periodic image interactions for confined systems. *Phys. Rev. B*
674 **73**, 233103 (2006).
675 46. Dion, M., Rydberg, H., Schröder, E., Langreth, D. C. & Lundqvist, B. I. Van der Waals
676 Density Functional for General Geometries. *Phys. Rev. Lett.* **92**, 246401 (2004).

47. Cooper, V. R. Van der Waals density functional: An appropriate exchange functional. *Phys. Rev. B* **81**, 161104 (2010).
48. Deslippe, J. *et al.* BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* **183**, 1269–1289 (2012).
49. da Jornada, F. H., Qiu, D. Y. & Louie, S. G. Nonuniform sampling schemes of the Brillouin zone for many-electron perturbation-theory calculations in reduced dimensionality. *Phys. Rev. B* **95**, 035109 (2017).
50. Alexeev, E. M. *et al.* Resonantly hybridized excitons in moiré superlattices in van der Waals heterostructures. *Nature* **567**, 81 (2019).
51. Zhang, C. *et al.* Interlayer couplings, Moiré patterns, and 2D electronic superlattices in MoS2/WSe2 hetero-bilayers. *Sci. Adv.* **3**, e1601459 (2017).

Acknowledgments:

We thank the OIST engineering support section, and Yuki Yamauchi from OIST Facilities Operations and Use section for their support. This work was supported in part by funding from the Femtosecond Spectroscopy Unit, Kick-start fund KICKS and OIST Innovative Technology - Proof of Concept Program at the Okinawa Institute of Science and Technology, Graduate University. Sample fabrication made use of the facilities in the Stanford Nano Shared Facilities (SNSF), supported by the National Science Foundation under award ECCS-1542152. The computational work was supported by the Center for Computational Study of Excited State Phenomena in Energy Materials, which is funded by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, under

Contract No. DE-AC02-05CH11231, as part of the Computational Materials Sciences Program. We acknowledge the use of computational resources at the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US DOE under the above contract.

Funding:

AMOS program, Chemical Sciences, Geosciences, and Biosciences Division, Basic Energy Sciences, U.S. Department of Energy (TFH, OK, HBR)

Gordon and Betty Moore Foundation's EPiQS Initiative through grant number GBMF9462 (JH, ALO)

Koret Foundation (OK)

Natural Science and Engineering Research Council (NSERC) of Canada, fellowship PGSD3-502559-2017 (EB)

Femtosecond Spectroscopy Unit – Okinawa Institute of Science and Technology Graduate University (VP, MKLM, CS, DRB, XZ, AAM, MMMA, NSC, AK, AJW, JM, KMD)

Okinawa Institute of Science and Technology Graduate University Innovative Technology Research – Proof of Concept Program (KMD)

JSPS KAKENHI Grant No. JP17K04995 (MM)

Kick-start fund KICKS – Okinawa Institute of Science and Technology Graduate University (KMD NSC)

JSPS KAKENHI Grant No. 21H01020 (JM)

FAPESP for post-doctoral fellowships, grant number 2018/04926-9 and 2017/20100-00 (HBR).

National Science Foundation Materials Research Science and Engineering Center DMR-1420634 and DMR- 2011738 (KB, BK)

Elemental Strategy Initiative, conducted by the MEXT, Japan, Grant No. JPMXP0112101001 (KW, TT)

JSPS KAKENHI Grant No. 19H05790 and JP20H00354 (KW, TT)

Department of Energy Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, under Contract No. DE-AC02-05CH11231 (FHJ, JDG)

Author contributions:

OK, EB and KMD conceived the project. KW and TT, BK and KB supplied raw materials for sample fabrication. EB and OK, supported by ALO, HBR, and JH, fabricated the samples. VP, MKLM, and CS collected the data with assistance from DRB, XZ, AAM, MMMA, AK, AJW and JM. VP, MKLM, CS, NSC and XZ performed preliminary analysis. EB and OK analyzed the data. FHJ and JDG performed theoretical calculations. JDG was supervised by FHJ. ALO, EB, HBR, JH, and OK were supervised by TFH. KMD supervised the project. All authors contributed to discussions and manuscript preparations.

Competing interests:

J.M., M.K.L.M., and K.M.D. are inventors on a patent application related to this work filed by the Okinawa Institute of Science and Technology School Corporation (US 2020/0333559 A1 published on October 22, 2020). The authors declare no other competing interests.

Corresponding Author:

Keshav M. Dani: kmdani@oist.jp

Data and materials availability:

All data are available in the main text or the supplementary materials.

Extended Data figure legends:

Extended Data Fig. 1. Reflection contrast spectra of the sample. Top: reflection of the 1L WSe_2 region, showing its lowest excitonic resonance at 1.7 eV, marked with a grey dashed line. **Middle:** the same, for 1L MoS_2, with its lowest resonance at 1.95 eV. **Bottom:** the same, for the heterobilayer region. It shows multiple resonances, marked I, II, and III instead of a single line in the vicinity of the WSe_2 resonance. This signature has been categorized as evidence for the emergence of the moiré pattern[16]. All measurements were done at 80 K.

Extended Data Fig. 2. Distortion correction procedures in the momentum-microscope a. Schematic of the imaging optics in the momentum-microscope. To correct for the distortions in the imaging optics, a physical grid was inserted at the back focal plane of the objective lens. **b.** A cut in the k_x-k_y plane in the presence of the grid as imaged by the detector before correction.

Overlaid on top is a yellow square box representing a perfect square. It is visible that the pattern of the grid is not perfectly square. A weak pincushion distortion is also visible. **c.** The same, after correction, showing the grid lines conforming to the yellow square **d.** A 2D projection of the grid at different electron kinetic energies E_{kin}, before corrections. The variation of the distortion and magnification with different E_{kin} is obvious. **e.** The same, after energy dependent grid distortion corrections.

Extended data Fig. 3. Static ARPES data with VB assignment near the K-points. A re-scaled version of Fig. 2a of the main text, highlighting the MoS_2 VBs around the K-valleys.

Extended data Fig 4. Extracting the exciton-bound electron dispersion. a. k_x ARPES cut of the photoexcited electron signal around the K point. The dashed yellow line is the VB dispersion. **b.** Selected spectra (EDC) along the dashed orange, green, and red lines marked in (a), together with the Gaussian fits, demonstrating the negative dispersion of the signal.

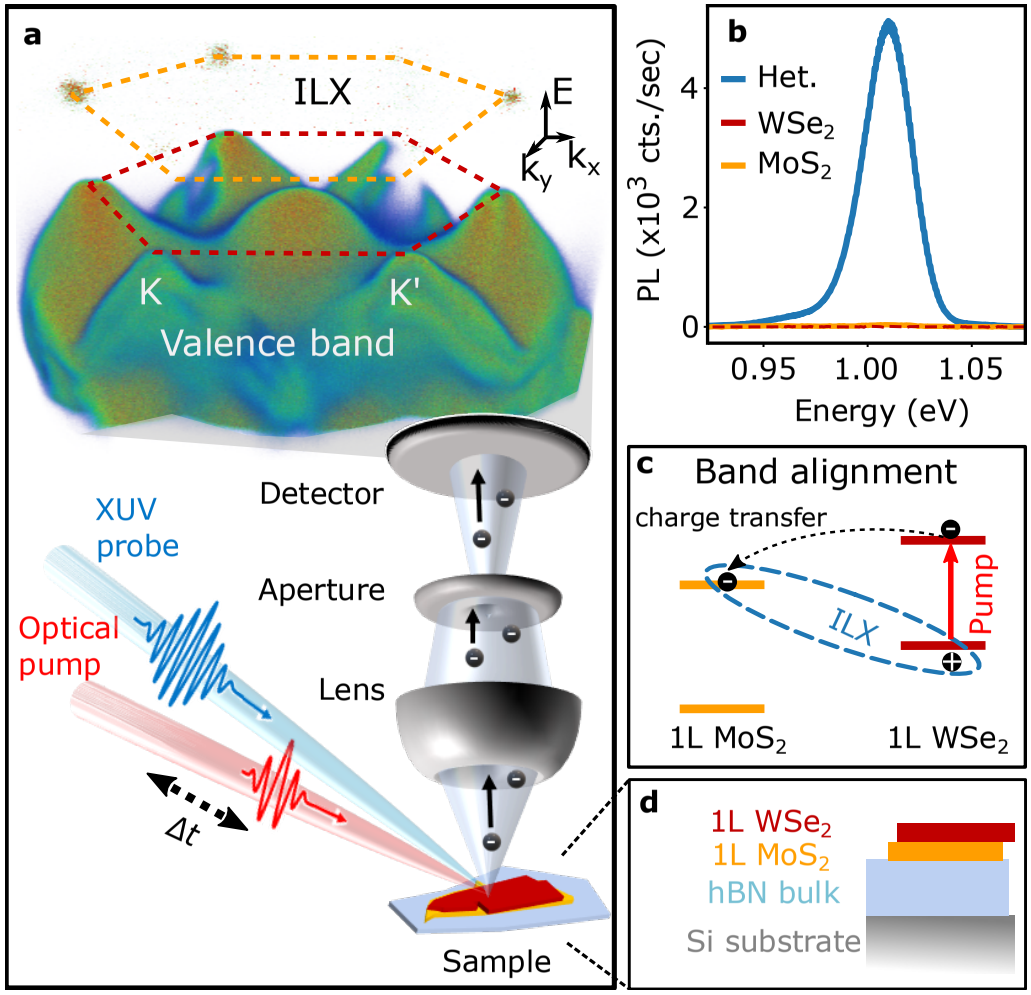
Extended Data Fig. 5. TR-μ-ARPES data with above band-gap photoexcitation. a. A momentum-slice along the Γ-K axis of the BZ at t=0 ps. The photoexcited electrons are scattered over a wide momentum and the shaded energy range above the dashed orange line, with a clear CB dispersion around the K point (yellow dashed indicator). **b.** Normalized EDC at the K point. The green plot is the equilibrium data. The black plot refers to the data at t=0 ps. **c.** The same as (a), at t=50 ps. The photoexcited electrons are concentrated at 1 eV energy (magenta dashed line) with an anomalous dispersion curvature. **d.** The same as (b), for t=50 ps. The red plot refers to the data at 50 ps delay. This highlights the spectral differences between the unbound (at t=0 ps) and exciton-bound (at later times) electrons showing up at different energies. The dashed magenta lines at energies below 0 eV indicate the band-edge energies for each VB. In comparison to the equilibrium VB EDC, at t=50 ps a reduction in counts is clearly registered for

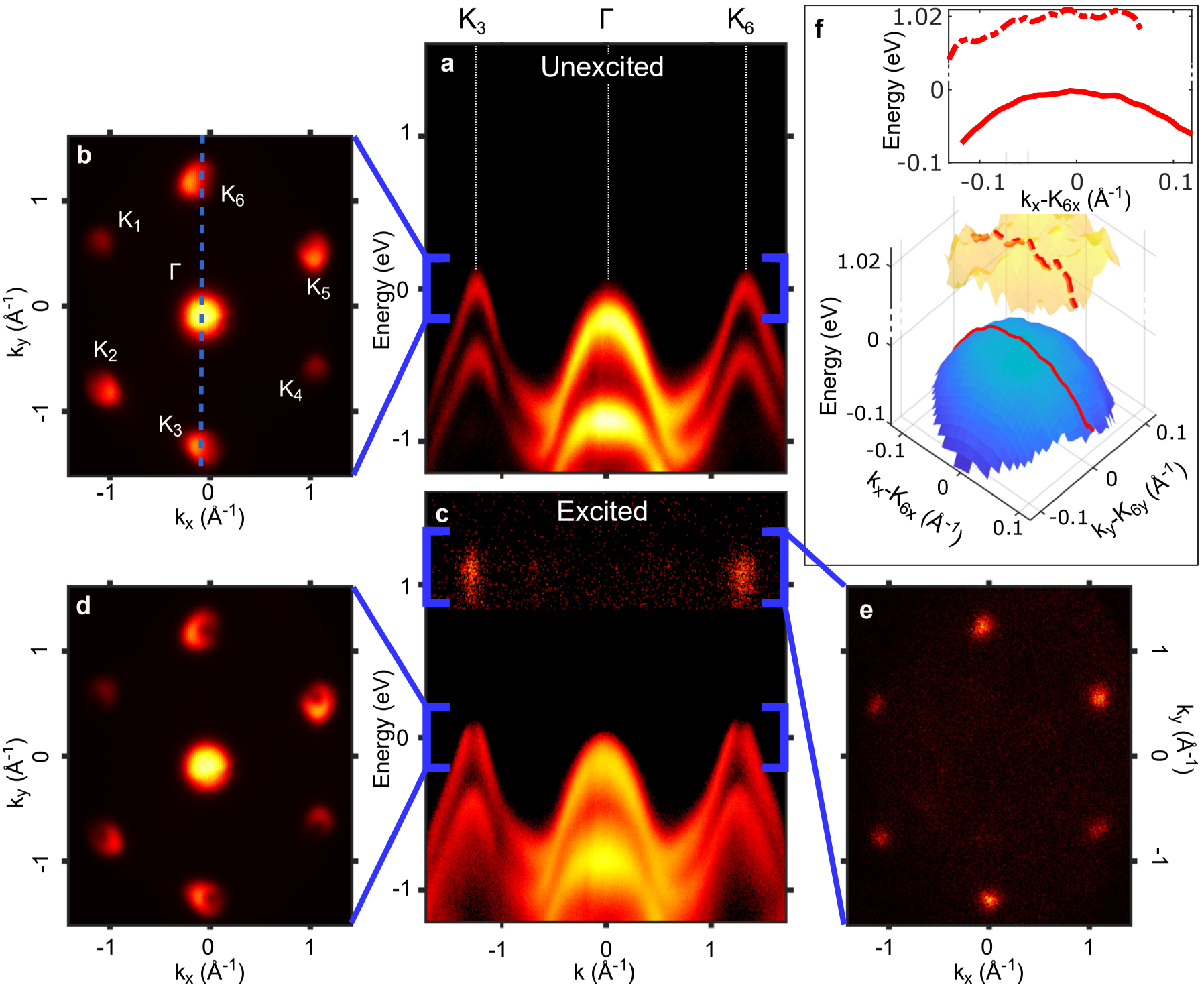
VB 1, associated with WSe_2, whereas none is registered for the other VBs, deducing that no holes are accumulating in MoS_2. **e.** Average momentum deviation from the VBM, $(k_x,k_y)=(0,0)$ $Å^{-1}$, of the ILX-bound hole distributions, determined at various time delays using the Gaussian fit. The distribution are clustered on average around $(k_x,k_y)=(0.002,0.0088) \pm (0.0056,0.0042)$ $Å^{-1}$, effectively (within a single pixel error) at the VBM. **f**. The same for the ILX-bound electrons. Their distributions are broadly clustered around $(k_x,k_y)=(-0.0139,0.0381) \pm (0.0019,0.0035)$ $Å^{-1}$, deviating from the hole momentum. This is attributed to the expected momentum-mismatch between the ILX-bound electrons and holes of a moiré exciton.

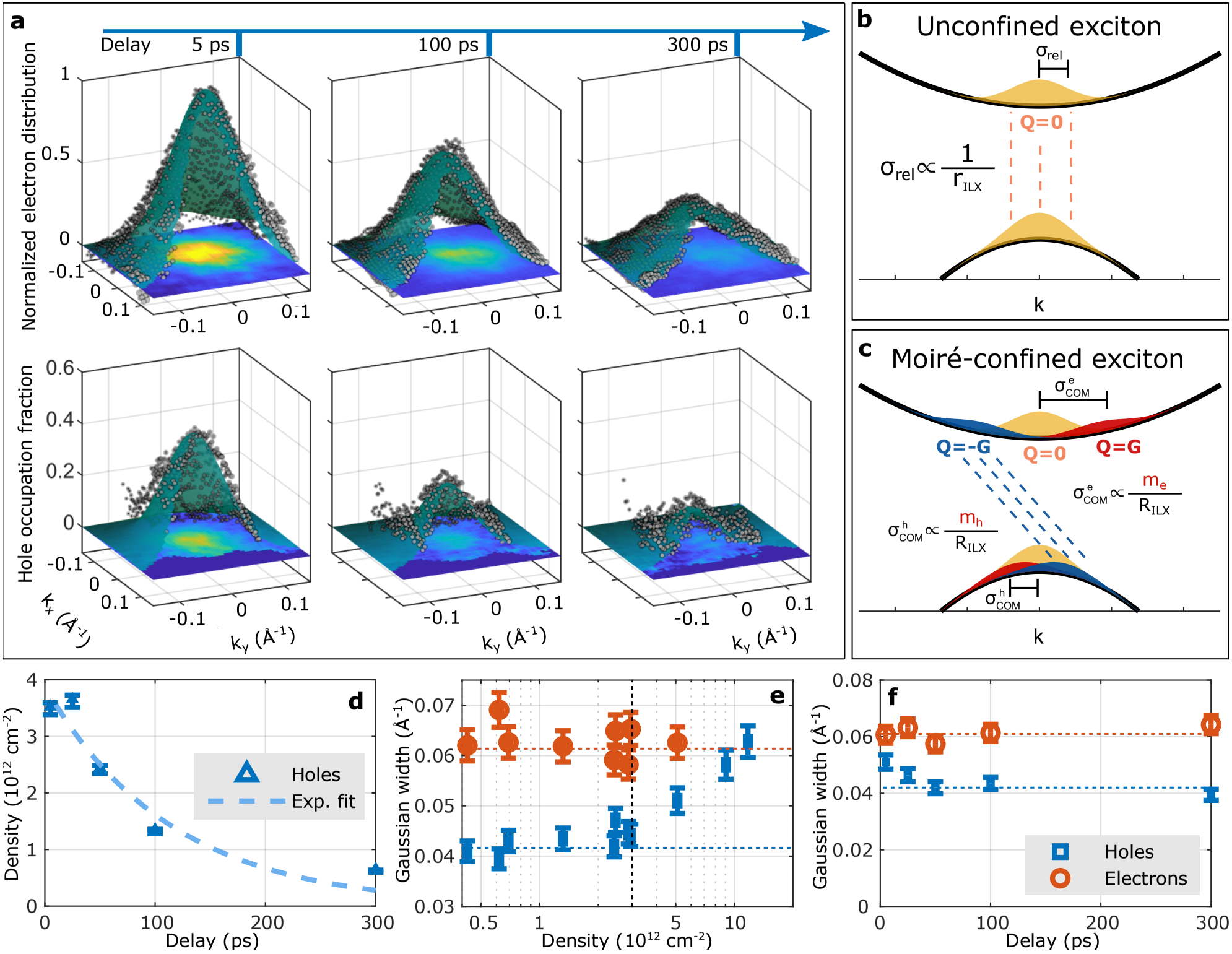
Extended Data Fig. 6. Extracting the photoexcited hole distribution. a. APRES energy-momentum cut along K-Γ direction around a specific K point for unexcited conditions. **b.** Fitting the VB with three gaussians in energy near the center of the plot in (a). **c.** The same as (a), after the photoexcitation. **d.** The same as (b), after the photoexcitation. **e.** Heatmap of the photoemitted counts associated with the top VB of the unperturbed sample. **f.** The same, after the excitation. Red squares mark the region used to normalize the counts from each measurement. **g.** The hole occupation distribution map resulting from the comparison between panels (e) and (f).

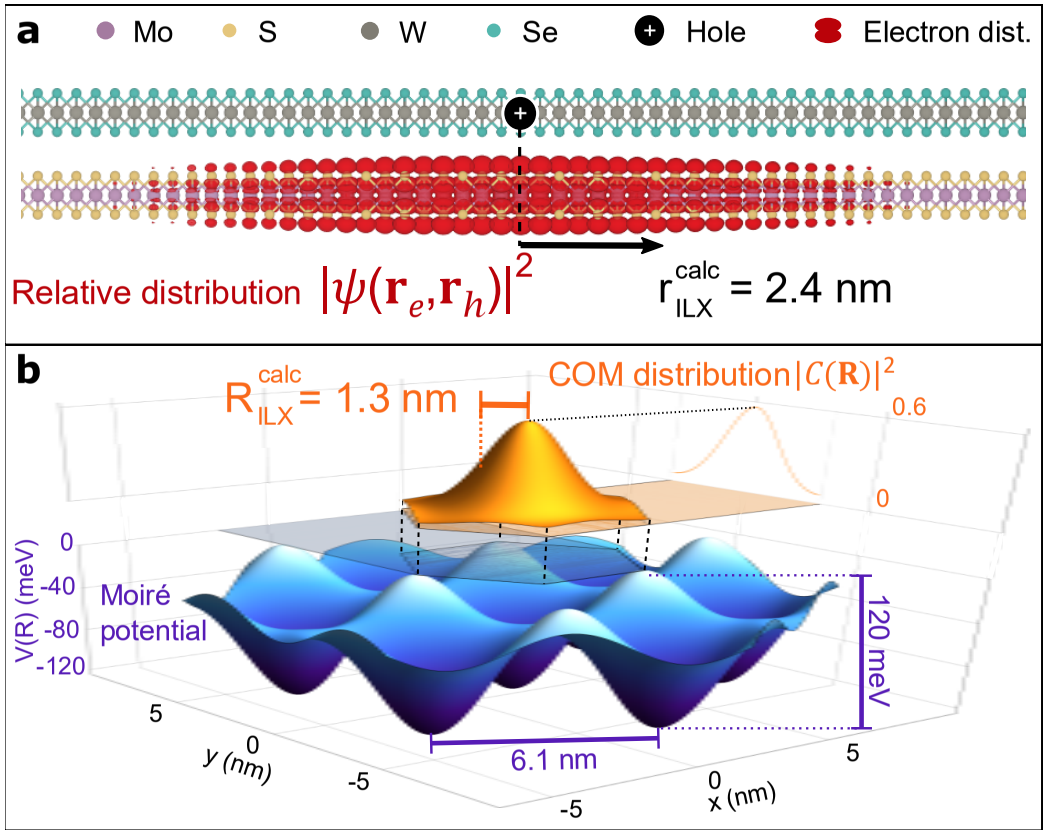
Extended Data Fig. 7. Temperature dependence of ILX density and momentum-distribution widths at 50 ps delay. a. ILX density, acquired for two excitation powers, at 100 K and 300 K. The reduction of ILX density with temperature is in line with the expected shortening of their lifetime in elevated temperatures. **b.** Fitted Gaussian widths of the ILX-bound hole and electron distributions at 100 K and 300 K. The width of the ILX-bound hole distribution does hardly change, while the electron distribution displays some broadening within our experimental accuracy. This is consistent with the analysis in the Supplementary Material that predicts little temperature dependence for the widths of $I^e(\mathbf{k})$ and $I^h(\mathbf{k})$. Both excitation powers show very

similar distribution widths, ruling out low signal-to-noise issues or many-body effects. The later is also consistent with the ILX densities being lower than their broadening onset described in Fig. 3e.


a
ILX
E
k_y
k_x
K
K'
Valence band
Detector
XUV
probe
Aperture
Optical
pump
Lens
Δt
Sample
b
Het.
WSe_2
MoS_2
PL (x10^3 cts./sec)
4
2
0
0.95
1.00
1.05
Energy (eV)
c
Band alignment
charge transfer
ILX
Pump
1L MoS_2
1L WSe_2
d
1L WSe_2
1L MoS_2
hBN bulk
Si substrate


a
Unexcited
K3
Γ
K6
Energy (eV)
b
K1
K2
K3
K4
K5
K6
Γ
ky (Å-1)
kx (Å-1)
c
Excited
k (Å-1)
d
e
f
Energy (eV)
kx-K6x (Å-1)
ky-K6y (Å-1)


a
Delay
5 ps
100 ps
300 ps
Normalized electron distribution
Hole occupation fraction
k_x (Å^{-1})
k_y (Å^{-1})
b
Unconfined exciton
σ_{rel}
Q=0
$\sigma_{rel} \propto \frac{1}{r_{ILX}}$
k
c
Moiré-confined exciton
σ^e_{COM}
Q=-G
Q=0
Q=G
$\sigma^e_{COM} \propto \frac{m_e}{R_{ILX}}$
$\sigma^h_{COM} \propto \frac{m_h}{R_{ILX}}$
σ^h_{COM}
k
d
Density (10^{12} cm^{-2})
Delay (ps)
Holes
Exp. fit
e
Gaussian width (Å^{-1})
Density (10^{12} cm^{-2})
f
Gaussian width (Å^{-1})
Delay (ps)
Holes
Electrons


a
Mo
S
W
Se
Hole
Electron dist.
Relative distribution $|\psi(\mathbf{r}_e,\mathbf{r}_h)|^2$
r_{ILX}^{calc} = 2.4 nm
b
R_{ILX}^{calc} = 1.3 nm
COM distribution $|C(\mathbf{R})|^2$
0.6
0
V(R) (meV)
0
-40
-80
-120
Moiré potential
120 meV
6.1 nm
5
0
-5
y (nm)
5
0
-5
x (nm)