Article

Ultralow contact resistance between semimetal and monolayer semiconductors

https://doi.org/10.1038/s41586-021-03472-9

Received: 1 July 2020

Accepted: 18 March 2021

Published online: 12 May 2021

Check for updates

Pin-Chun Shen[1,11]✉, Cong Su[2,3,4,5,11], Yuxuan Lin[1,6,11], Ang-Sheng Chou[7,8,11], Chao-Ching Cheng[7], Ji-Hoon Park[1], Ming-Hui Chiu[1,9], Ang-Yu Lu[1], Hao-Ling Tang[1,9], Mohammad Mahdi Tavakoli[1], Gregory Pitner[10], Xiang Ji[1], Zhengyang Cai[1], Nannan Mao[1], Jiangtao Wang[1], Vincent Tung[9], Ju Li[5], Jeffrey Bokor[4,6], Alex Zettl[2,3,4], Chih-I Wu[8], Tomás Palacios[1], Lain-Jong Li[7]✉ & Jing Kong[1]✉

Advanced beyond-silicon electronic technology requires both channel materials and also ultralow-resistance contacts to be discovered[1,2]. Atomically thin two-dimensional semiconductors have great potential for realizing high-performance electronic devices[1,3]. However, owing to metal-induced gap states (MIGS)[4–7], energy barriers at the metal–semiconductor interface—which fundamentally lead to high contact resistance and poor current-delivery capability—have constrained the improvement of two-dimensional semiconductor transistors so far[2,8,9]. Here we report ohmic contact between semimetallic bismuth and semiconducting monolayer transition metal dichalcogenides (TMDs) where the MIGS are sufficiently suppressed and degenerate states in the TMD are spontaneously formed in contact with bismuth. Through this approach, we achieve zero Schottky barrier height, a contact resistance of 123 ohm micrometres and an on-state current density of 1,135 microamps per micrometre on monolayer MoS_2; these two values are, to the best of our knowledge, the lowest and highest yet recorded, respectively. We also demonstrate that excellent ohmic contacts can be formed on various monolayer semiconductors, including MoS_2, WS_2 and WSe_2. Our reported contact resistances are a substantial improvement for two-dimensional semiconductors, and approach the quantum limit. This technology unveils the potential of high-performance monolayer transistors that are on par with state-of-the-art three-dimensional semiconductors, enabling further device downscaling and extending Moore's law.

The electrical contact resistance at a metal–semiconductor interface has been an increasingly critical, yet unsolved, issue for the semiconductor industry, hindering the ultimate scaling and the performance of electronic devices[9]. The main cause of this resistance is the energy barrier—the Schottky barrier—formed between the metal electrode and semiconductor[8], owing to (I) the energy difference between the metal work function and the semiconductor electron affinity, and (II) MIGS, resulting in Fermi-level pinning[4–7]. When a semiconductor is in close proximity to a metal surface, the extended wavefunction from the metal perturbs the environment of the semiconductor, leading to rehybridizations of the semiconductor's original wavefunctions. MIGS are a result of such perturbation, where new states in resonance with the metal states emerge in the bandgap (Fig. 1a), as compared to the original density of states (DOS) of the semiconductor (such as MoS_2) before contact (Fig. 1c). Resembling the contours of metal DOS after the band alignment, the density of the MIGS is contributed by the valence band (that is, donor-like, positively charged states) and the conduction band (that is, acceptor-like, negatively charged states)[10]. It has been theoretically and experimentally demonstrated that the Fermi level of the metal–semiconductor system is pinned at around the branching point of these two components (referred to as gap-state pinning; Fig. 1a), an energetically favourable state when free of residue charges[10]. If the Fermi level of the system lies inside the semiconductor bandgap, a Schottky barrier is unavoidable (Fig. 1d).

Two strategies have been developed to solve this issue: (I) reducing the Schottky barrier width by heavily doping the semiconductor so that the tunnelling current outweighs thermionic emission current at the Schottky barrier[11]; and (II) decoupling the metal–semiconductor interaction by introducing a thin dielectric, molecular layer, or van der Waals gap at the interface[6,12–14]. Whereas the first strategy is technologically challenging for two-dimensional (2D) materials, the second architecture is dominated by a non-negligible tunnelling barrier owing

[1]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. [2]Kavli Energy NanoSciences Institute, University of California, Berkeley, CA, USA. [3]Department of Physics, University of California, Berkeley, CA, USA. [4]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [5]Department of Nuclear Science and Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. [6]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA. [7]Corporate Research, Taiwan Semiconductor Manufacturing Company (TSMC), Hsinchu, Taiwan. [8]Graduate Institute of Photonics and Optoelectronics, National Taiwan University, Taipei, Taiwan. [9]Physical Science and Engineering Division, King Abdullah University of Science & Technology (KAUST), Thuwal, Saudi Arabia. [10]Corporate Research, Taiwan Semiconductor Manufacturing Company (TSMC), San Jose, CA, USA. [11]These authors contributed equally: Pin-Chun Shen, Cong Su, Yuxuan Lin, Ang-Sheng Chou. ✉e-mail: pcshenc@tsmc.com; lanceli1@hku.hk; jingkong@mit.edu

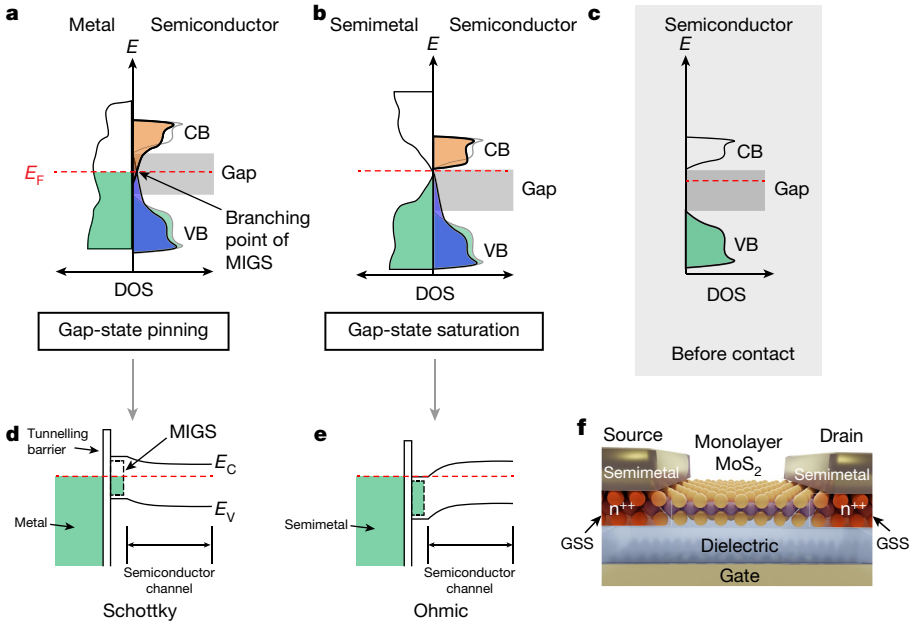


Fig. 1 | The concept of gap-state saturation at semimetal–semiconductor contact. a, The DOS of normal metal and semiconductor contact. The contributions of the conduction band (CB) and valence band (VB) to the metal-induced gap states (MIGS) are shaded as orange and blue areas, respectively. The Fermi level (E_F) is pinned at around the branching point of the MIGS, leading to gap-state pinning. Green area, electron-occupied states. **b**, The DOS of semimetal and semiconductor contact. Because the Fermi level of the semimetal aligns with the conduction band of the semiconductor, and the DOS at the Fermi level of the semimetal is near-zero, conduction-band-contributed MIGS are suppressed and the branching point is elevated into the conduction band. The MIGS, now mostly contributed by the valence band, are saturated, leading to gap-state saturation. **c**, The reference DOS of the semiconductor before contact. **d**, **e**, The band structure of metal–semiconductor contact (**d**), where a Schottky barrier is formed as a result of gap-state pinning, and semimetal and semiconductor contact (**e**), where ohmic contact is formed as a result of gap-state saturation. **f**, Schematic of a 2D FET with a monolayer semiconductor (MoS_2) channel and semimetal (Bi) contacts. The degenerate part of Bi-contacted MoS_2 due to gap-state saturation (GSS) is marked in orange colour.

to the increased metal–semiconductor distance[15,16]. Since these two strategies typically result in either high Schottky barriers (in the range of 100–400 meV), or interface tunnelling barriers with thickness greater than 1 nm in monolayer TMD transistors[6,13,14,17], the state-of-the-art contact resistance is around 1 kΩ µm, at least one order of magnitude larger than metal–Si contact[1].

Here we propose a strategy to reduce contact resistance by suppressing MIGS using semimetal–semiconductor contacts to avoid gap-state pinning (Fig. 1b). A semimetal has near-zero DOS at the Fermi level where few MIGS can be induced; therefore, if the Fermi level of a semimetal is close to the conduction band minimum of the semiconductor, the conduction-band-contributed MIGS are greatly reduced. As a result, the MIGS are greatly suppressed and are purely contributed by the valence band, thus can be filled up and saturated—this is referred to as gap-state saturation. In this way, the semiconductor in contact with semimetal will be in a degenerate state and free of a Schottky barrier at the interface (Fig. 1e).

To prove this hypothesis, we fabricate back-gated field-effect transistors (FETs) with various monolayer TMDs and form low-resistance contacts ('ohmic' contacts) by depositing semimetal bismuth (Bi) onto the contact windows of the 2D channel (Fig. 1f). Figure 2a compares typical transfer curves (drain-to-source current, I_{DS}, versus gate-to-source voltage, V_{GS}) of Bi-, Ni- and Ti-contacted transistors made of monolayer MoS_2 synthesized by metal–organic chemical vapour deposition (MOCVD). Among them, the Bi–MoS_2 FET clearly exhibits enhanced n-channel conduction with a high on/off current ratio of $>10^7$. Moreover, the Bi–MoS_2 interface shows linear output characteristics (I_{DS} versus drain-to-source voltage, V_{DS}) both at room temperature ($T = 295$ K) (Fig. 2b) and also at low temperatures (Fig. 2g and Extended Data Fig. 1a–c), indicating a good ohmic contact with a negligible Schottky barrier height. By contrast, the nonlinear output characteristics in both Ni- and Ti-contacted MoS_2 (Extended Data Fig. 1d, e, h) suggest the presence of barriers, which are similar to previous reports using either Schottky or interfacial layer contacts[13,14,17–21]. The contact resistance, R_C, for the Bi contact to monolayer MoS_2 is as low as 123 Ω µm at a carrier density (n_{2D}) of 1.5×10^{13} cm^{-2}, as shown in Fig. 2c (more data points are shown in Fig. 4h and Extended Data Fig. 2d).

It is observed that the electrical characteristics of Bi–MoS_2 FETs are dominated by the MoS_2 channel. The contact resistance ($2R_C$) in the Bi–MoS_2 FETs contributes to less than 5% of the total resistance (R_{TOT}) for a wide range of values of n_{2D} (Extended Data Fig. 2d). By lowering the temperature from room temperature to 77 K, the drain current density (I_{DS}) in two-terminal Bi–MoS_2 FETs increases owing to the enhanced electron mobility of MoS_2, whereas the I_{DS} in both Ni–MoS_2 and Ti–MoS_2 FETs is dramatically suppressed (Fig. 2d and Extended Data Fig. 1g), owing to the reduction of thermionic emission current at contact. The field-effect mobilities from two-terminal and four-terminal devices as a function of temperature (Fig. 2e) show identical characteristics, which suggests that the ultralow R_C makes two-terminal devices a simple yet powerful platform for characterizing intrinsic temperature-dependent transport properties of 2D semiconductors, whereas four-terminal devices were always required previously[14,20]. As a side note, the two-terminal field-effect electron mobility of the MOCVD-grown monolayer MoS_2 channel reaches 120 cm^2 V^{-1} s^{-1} at 77 K and 55 cm^2 V^{-1} s^{-1} at room temperature, outperforming ultrathin silicon (<2 nm) and germanium (<4 nm) on insulator devices[3].

We present three pieces of evidence of the absence of a Schottky barrier in Bi–MoS_2 FETs. First, R_C in the Bi–MoS_2 FETs is nearly independent of n_{2D} (determined by the overdrive voltage, $V_{GS} - V_T$, where V_T is the threshold voltage), suggesting a negligible, if not zero, energy barrier at the Bi–MoS_2 (Fig. 4h)[14,18–20,22]. Second, our temperature-dependent I_{DS} shows that the Bi–MoS_2 FETs operate in the barrier-free transport limit (see Methods).

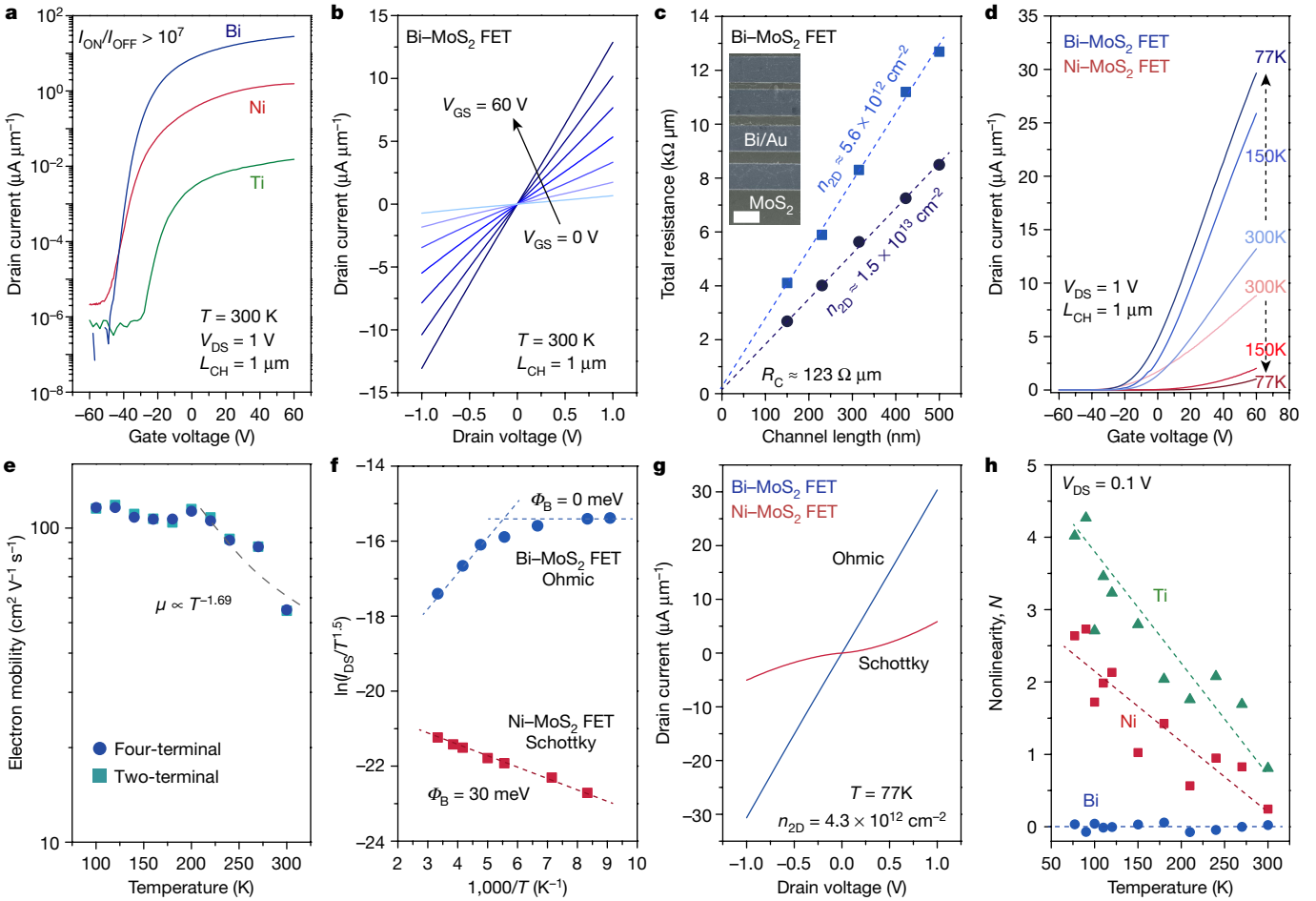


Fig. 2 | Comparison of ohmic and Schottky contacts in monolayer MoS_2 FETs. a, Comparison of room-temperature transfer characteristics (I_{DS}–V_{GS}) of typical monolayer MoS_2 FETs with Bi, Ni and Ti contacts on 300-nm-thick SiO_2 dielectrics. The Bi–MoS_2 FET presents an on/off current ratio (I_{ON}/I_{OFF}) of >10^7. **b**, Typical output characteristics (I_{DS}–V_{DS}) of Bi–MoS_2 FETs at room temperature. **c**, Contact resistance (R_C) extraction using the transfer-length method (TLM) for Bi–MoS_2 FETs on 100-nm-thick SiN_x dielectrics. Blue squares and black circles are total resistance versus channel length at carrier densities of 5.6×10^{12} and 1.5×10^{13} cm^{-2}, respectively. Inset, False colour SEM image of the TLM structure. Scale bar, 1 μm. **d**, Typical I_{DS}–V_{GS} of ohmic Bi–MoS_2 (blue) and Schottky Ni–MoS_2 (red) FETs on 300-nm-thick SiO_2 dielectrics at various temperatures. **e**, Two-terminal and four-terminal electron mobilities of the Bi–MoS_2 FET as a function of temperature. The consistency between these two methods suggests negligible R_C. The temperature dependence of the mobility $\mu \propto T^{\gamma}$ indicates the dominant optical phonon scattering[41]. **f**, Arrhenius plots of the ohmic Bi–MoS_2 (blue) and Schottky Ni–MoS_2 (red) FETs at a carrier density (n_{2D}) of 1.5×10^{12} cm^{-2} and V_{DS} of 1 V. A 30-meV and negligible barrier for Ni–MoS_2 and Bi–MoS_2 FETs, respectively, are extracted. **g**, Typical I_{DS}–V_{DS} of Bi–MoS_2 (blue) and Ni–MoS_2 (red) FETs with V_{GS} = 60 V and $n_{2D} \approx 4.3 \times 10^{12}$ cm^{-2} at 77 K. The full I_{DS}–V_{DS} characteristics can be found in Extended Data Fig. 1b, e. **h**, Nonlinearity ($N = (d^2I_{DS}/dV_{DS}^2)/2(dI_{DS}/dV_{DS})$) of the I_{DS}–V_{DS} for MoS_2 FETs with Bi, Ni and Ti contacts at various temperatures. Data are extracted at V_{DS} = 0.1 V at $n_{2D} \approx 4.3 \times 10^{12}$ cm^{-2}. N = 0 of the Bi–MoS_2 FET indicates its linear I_{DS}–V_{DS} characteristics, revealing ohmic contact at the Bi–MoS_2 junction.

For Schottky contacts, the Arrhenius plots (ln($I_{DS}/T^{1.5}$) versus 1,000/T; T, temperature) are linear with negative slopes (Fig. 2f and Extended Data Fig. 2a, b), from which the Schottky barrier heights (Φ_{SB}) at flatband are extracted to be 100 meV for Ni–MoS_2 and 150 meV for Ti–MoS_2 (Extended Data Fig. 1f, i). However, this analysis becomes invalid for Bi–MoS_2 FETs when the device is turned on (V_{GS} > −30 V). Instead, the saturation-like regime at lower temperatures (<200 K) suggests a zero contact barrier height for electron transport, and the positive slope in the range of 200–300 K can be attributed to the negative correlation between mobility and temperature. Finally, the I_{DS}–V_{DS} curves of Bi–MoS_2 FETs remain linear at low temperatures (Fig. 2g and Extended Data Fig. 1a–c). We define the nonlinearity, N, of the I_{DS}–V_{DS} relation as $N = (d^2I_{DS}/dV_{DS}^2)/2(dI_{DS}/dV_{DS})$. N = 0 corresponds to linear relation, where no barrier exists; larger N means increased nonlinearity, which is associated with a higher Schottky barrier. As shown in Fig. 2h, although the nonlinearities are close to zero for the Ni and Ti contacts at room temperature, they increase quickly as the temperature decreases; by contrast, the nonlinearity for the Bi contact remains zero for different temperatures.

To confirm the gap-state saturation of Bi–MoS_2, we carry out first-principles calculation based on the crystal structure identified from transmission electron microscopy (TEM), and the theoretical result is further substantiated by evidence from X-ray photoelectron spectroscopy (XPS) and Raman spectroscopy. The selected-area electron diffraction (SAED) measured on a freestanding MoS_2/Bi/Au stack (Fig. 3a) shows three sets of hexagonal patterns corresponding to the Bi, MoS_2 and Au crystals, respectively, indicating that the Bi (0001) plane is parallel to the plane of MoS_2 (Fig. 3b and Extended Data Fig. 3b–d). The diffraction pattern of MoS_2 proves its 2H semiconducting phase with no metallic phase transition[23]. Among the pseudocubic (metallic) and rhombohedral (semimetal) phases of Bi[24], the hexagonal SAED pattern has ruled out the possibility of the former. As a comparison, when Bi is directly deposited onto amorphous carbon, it becomes polycrystalline and partially oxidized into Bi_2O_3, evidenced by the powder-like diffraction rings with an extra set of Bi_2O_3 patterns located at 3.0 cm^{-1}, as shown in Extended Data Fig. 3f. From this, we conclude that no oxidation takes place at the Bi–MoS_2 interface.

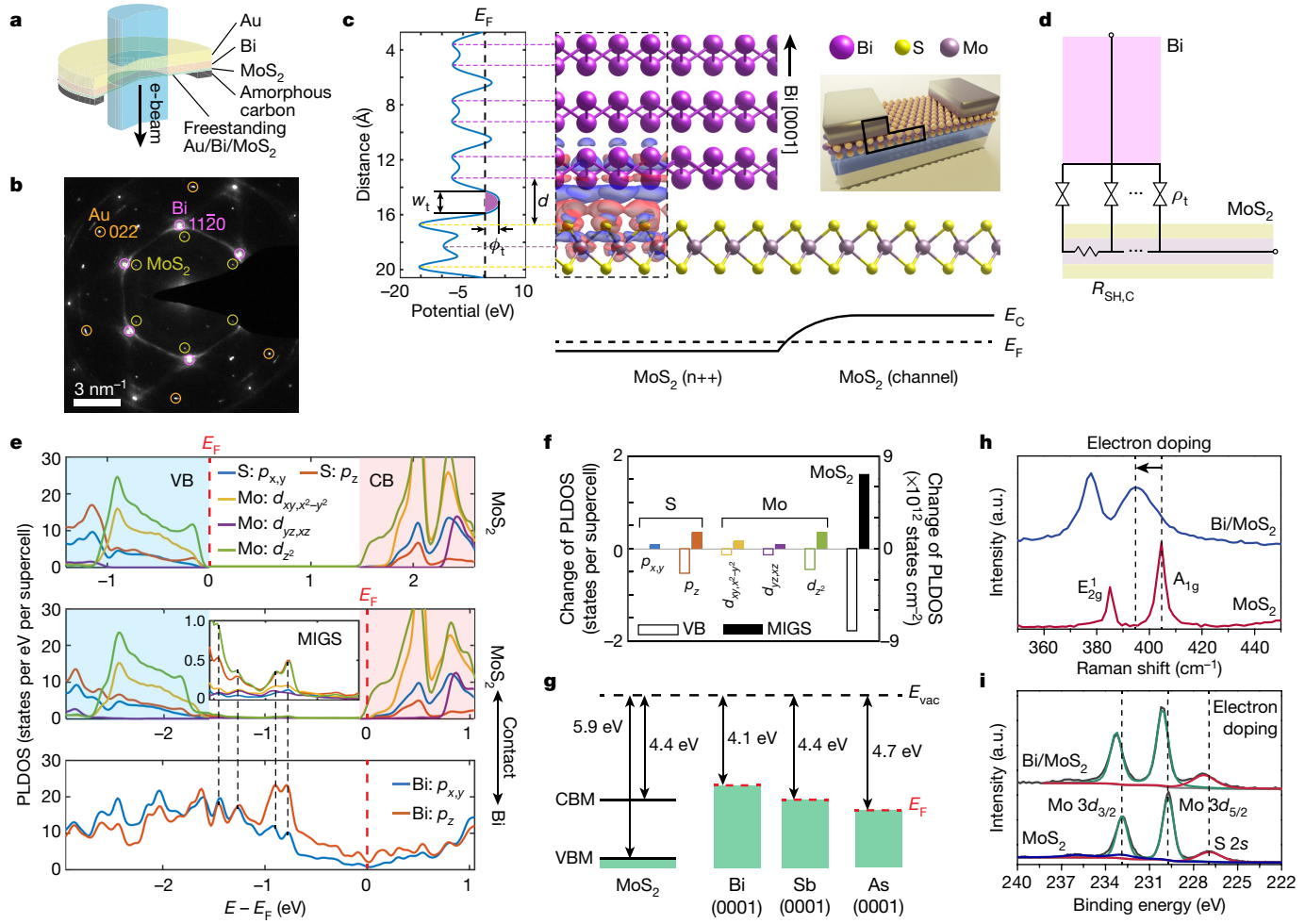


Fig. 3 | Crystal structure and mechanism of ohmic contact. a, Schematic of freestanding Au/Bi/MoS_2 on a meshed amorphous carbon (a-carbon) TEM grid for SAED. **b**, SAED patterns of MoS_2 (3.6 cm^{-1}), Bi (4.3 cm^{-1}) and Au (6.8 cm^{-1}) circled in yellow, pink and orange, respectively. **c**, The side view of Bi–MoS_2 (upper right, with the area marked in the inset three-dimensional render), and the corresponding electrostatic potential profile along the vertical direction (left). The electron tunnelling barrier is shaded in purple (width, w_t = 1.66 Å; height, Φ_t = 3.6 eV). The distance between Bi and S atomic layers is d = 3.4 Å. The differential charge density inside the region of dashed line, calculated by subtracting the pre-contact charge density from the post-contact charge density, is superposed in the atomic structure (red, positive; blue, negative). Bottom, the band profile of MoS_2. **d**, The equivalent circuit of the Bi–MoS_2 contact area with the space partitioned with respect to their atomic colours. ρ_t, tunnelling resistivity; $R_{SH,C}$, sheet resistances of MoS_2 in contact with Bi. **e**, PLDOS of MoS_2 before (upper panel) and after (middle panel) in contact with Bi (lower panel). The valence band is shaded in light blue and conduction band in light red. The Fermi level (E_F) is shifted from inside the gap (before Bi contact) to above the conduction band minimum (after Bi contact). Inset to middle, the magnified PLDOS in the bandgap, showing MIGS. **f**, The change of PLDOS of different orbitals in the valence band and MIGS areas marked in **e**. **g**, The band alignment of monolayer MoS_2 with examples of semimetals (Bi, Sb and As). **h**, **i**, The shift of Raman (**h**) and XPS (**i**) spectra comparing standalone MoS_2 and Bi-contact MoS_2. a.u., arbitrary units; CBM, conduction band minimum; VBM, valence band maximum.

We build a supercell (Fig. 3c) based on the above information and perform first-principles calculations using density functional theory (DFT). First, the resistance of the tunnelling barrier between Bi and MoS_2 is calculated to be low. The barrier has a width (w_t) of 1.66 Å and a height (Φ_t) of 3.6 eV marked in the electrostatic potential profile. This induces a tunnelling resistivity (ρ_t) of $1.81 \times 10^{-9}\ \Omega\ cm^2$, as shown in the equivalent circuit in Fig. 3d, which is very small compared with those made with thin dielectric gaps[6,14]. Second, the Fermi level of MoS_2 is shifted into the conduction band (Fig. 3e), yielding a degenerate MoS_2 in contact with Bi with $n_{2D} \approx 2 \times 10^{13}\ cm^{-2}$ and a small sheet resistance $R_{SH,C} \approx 15.6\ k\Omega$. The degenerate MoS_2 (I) reduces the overall resistance of the ρ_t–$R_{SH,C}$ network in Fig. 3d, and (II) more importantly, eliminates the Schottky barrier between Bi and MoS_2, resulting in a negligible Schottky barrier resistance. The overall contact resistance is calculated to be 130 Ω µm, in very good agreement with our measured values (see Methods for details).

In contrast to other metals (such as Ni, Pt, Au, Ag, In)[25,26], the MIGS are greatly reduced in the case of Bi (Fig. 3e). The wavefunction of the p_z orbital of Bi is in resonance with the p_z and d_{z^2} orbitals of MoS_2 where the MIGS clearly follow the projected local density of states (PLDOS) of Bi (Fig. 3e, inset). However, unlike the others, the MIGS have been completely saturated with electrons, which leads to gap-state saturation. Several other observations include: (I) The charge transfer from Bi to MoS_2 is minimal based on the undistorted potential profile, and the induced electric dipole almost completely falls within the van der Waals gap (seen from the differential charge distribution of Fig. 3c). (II) The decrease of the valence band states has outpaced the increase of MIGS (Fig. 3f), pushing the Fermi level up into the conduction band. (III) Two factors are strongly correlated to the final contact type: (a) the nature of zero DOS at the Fermi level and saturated bonds on the surface of the semimetal, and (b) the work function of the semimetal compared with the electron affinity of MoS_2. The DFT results of the contact between V-group semimetals (Bi, Sb and As) and MoS_2 support this hypothesis: since the Fermi levels of Bi and Sb are equal or above the conduction band minimum of MoS_2 (Fig. 3g), they are predicted

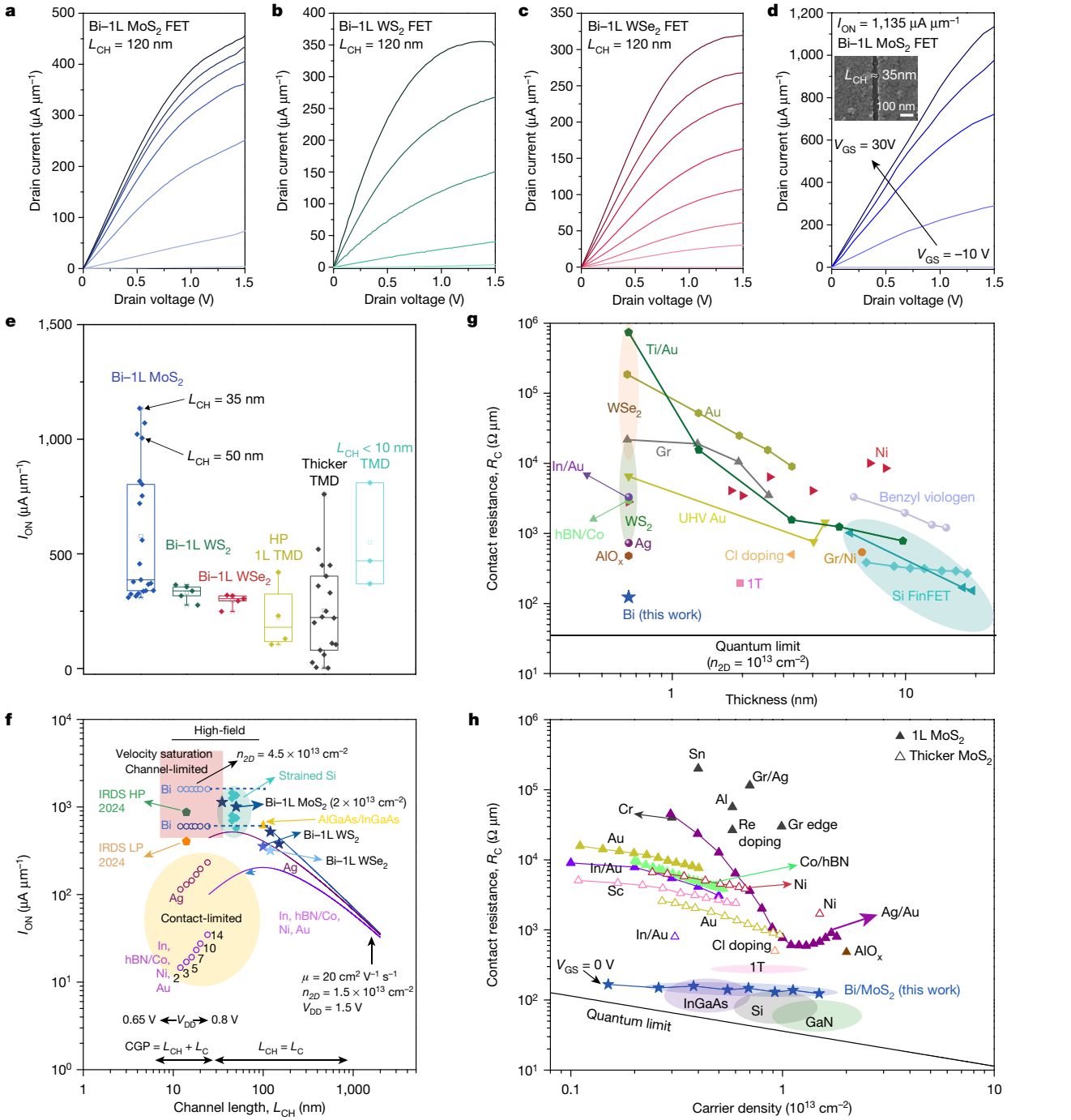


Fig. 4 | Benchmark of Bi-contacted 2D semiconductor technology. **a**–**c**, I_{DS}–V_{DS} curves of Bi–monolayer MoS_2 (MOCVD; **a**), Bi–monolayer WS_2 (exfoliated; **b**), and Bi–monolayer WSe_2 (exfoliated; **c**) FETs with L_{CH} of 120 nm. **d**, I_{DS}–V_{DS} curves of a 35-nm L_{CH} Bi–MoS_2 FET. V_{GS} changes from −10 V to 50 V for **a** and **b**, from −10 V to 60 V for **c**, and from −10 V to 30 V for **d**, all in steps of 10 V. Inset to **d**, SEM image of the 35-nm L_{CH} device. **e**, Statistics of I_{ON} for Bi-contacted monolayer TMD transistors, showing improved performance with respect to previous reports[18,22,30–34,42–45]. Data are extracted at the same V_{DS} of 1.5 V. **f**, Projected I_{ON} as a function of L_{CH} of monolayer TMD transistors with different contact technologies. The blue stars are the experimental data in this work. The solid lines are the projected I_{ON} for different metal contacts, including Bi (blue, $\rho_C \approx 8 \times 10^{-9}\,\Omega\,cm^2$), Ag (plum, $\rho_C \approx 3 \times 10^{-7}\,\Omega\,cm^2$), and In, hBN/Co, Ni or Au (purple, $\rho_C \approx 3 \times 10^{-6}\,\Omega\,cm^2$) at a fixed electron mobility μ of 20 $cm^2\,V^{-1}\,s^{-1}$ and n_{2D} of $1.5 \times 10^{13}\,cm^{-2}$. The projections at $L_{CH} < 26$ nm (14-nm technology nodes) takes into consideration both the L_C and the V_{DD} scalings. The relation between L_C and L_{CH} in this region is defined by the contacted gate pitch (CGP)[38]. The light and dark blue dashed lines are projections for Bi–MoS_2 transistors operating at velocity saturation, with n_{2D} of $4.5 \times 10^{13}\,cm^{-2}$ and $1.5 \times 10^{13}\,cm^{-2}$, respectively. The orange and green pentagons represent the lower bounds of the IRDS 2024 target for low-power (LP) and high-performance (HP) logic transistors[38], respectively. The turquoise diamonds and yellow triangle are I_{ON} for 90-nm technology node strained silicon[39,40], and 100-nm AlGaAs/InGaAs transistors[46]. **g**, Scaling of R_C with the thickness of MoS_2 and ultrathin Si fins or films. The ranges of R_C for monolayer WS_2 (shaded green) and WSe_2 (shaded orange) are also shown[2,13,14,18,22,30,32–34,42,45,47,48]. The black solid line indicates the quantum limit at $n_{2D} = 10^{13}\,cm^{-2}$. **h**, State-of-the-art contact technology for MoS_2 transistors plotted as a function of n_{2D}, showing the respective R_C of various semiconductor technologies (Si, III–Vs, and MoS_2)[1,13,14,17–22,33,34,47,49–51]. The black line represents the quantum limit of R_C, $\pi h/(4q^2k_F) \approx 0.036(n_{2D})^{-0.5}\,k\Omega\,\mu m$, which is determined by the quantum resistance ($h/2q^2 \approx 12.9\,k\Omega$, where h is Planck's constant and q is the elementary charge) and the number of conducting modes per channel width (k_F/π, where k_F is the Fermi wavevector) related to the 2D sheet carrier density (n_{2D}, in units of $10^{13}\,cm^{-2}$)[2]. 1L, monolayer.

to have negligible Schottky barrier contacts (Fig. 3e and Extended Data Fig. 4a); whereas the Fermi level of As–MoS_2 is still inside the bandgap.

Both Raman and XPS characterization results confirm that the monolayer MoS_2 under the Bi contacts turns to degenerate. First, the A_{1g} Raman vibration mode of Bi-contacted monolayer MoS_2 shows a ~10 cm^{-1} redshift, corresponding to an electron density of (2–5) × 10^{13} cm^{-2} and a Fermi level position at 40–100 meV above the conduction band minimum (Fig. 3h)[27,28]. In comparison, the A_{1g} peaks for the Ni–MoS_2 and Au–MoS_2 samples do not exhibit such shifts (Extended Data Fig. 5b). We note that all the three metal contacts result in similar shifts of the E_{2g} phonon mode, which is probably due to the strain induced at the MoS_2–metal interface[29]. Second, the blue shifts of both Mo 3*d* and S 2*p* peaks in the XPS spectra shown in Fig. 3i and Extended Data Fig. 5c reveal a 400-meV lifting of the Fermi level[30] when MoS_2 is in contact with Bi.

The proposed gap-state saturation mechanism at the Bi–MoS_2 interface applies to various 2D semiconductors and enables reliable and greatly improved transistor performance. We benchmark our transistor performance with the on-state current density (I_{ON}), a figure of merit in transistor scaling, and R_C, a key parameter limiting the scaling of I_{ON} and power supply voltage (V_{DD})[9]. Figure 4a–c shows typical output characteristics of monolayer MoS_2, WS_2, and WSe_2 transistors (channel length, L_{CH} = 120 nm) on 100-nm-thick SiN_x gate dielectrics. The linear relationships of I_{DS}–V_{DS} at low-field regime indicate their ohmic contacts, and this leads to high on/off current ratios (>10^7, Extended Data Figs. 6, 7e) and very high current-delivery capability (Fig. 4e) among monolayer TMDs at a drain voltage of 1.5 V[22,31–34]. The highest values of I_{ON} achieved in our study are 560 μA $μm^{-1}$ and 1,135 μA $μm^{-1}$ for a 120-nm L_{CH} and a 35-nm L_{CH} monolayer MoS_2 FETs, respectively (Extended Data Fig. 7d and Fig. 4d). Given that W-based TMDs have lower electron affinity than MoS_2, making it more difficult to make good n-type ohmic contacts to these materials[15], we perform first-principles calculation to confirm the formation of the degenerate state of WS_2 when contacted with Bi as shown in Extended Data Fig. 4c. In addition, it is observed that such barrier-free contacts can be formed on TMD crystals prepared by different methods including chemical vapour deposition (CVD), MOCVD, and mechanical exfoliation (Fig. 4a–c, Extended Data Figs. 7d–f, 8b–d), whereas a high-quality crystal is essential to avoid the undesired gap-state pinning (see DFT results in Extended Data Fig. 4b, experimental results in Extended Data Fig. 8 and additional discussion in Methods). Last but not least, the Bi contact is formed through a standard CMOS-compatible evaporation process, which promises good reliability and scalability. Figure 4e presents statistics of I_{ON} measured on more than 20 Bi–TMD transistors. Average I_{ON} values of 367, 331 and 300 μA $μm^{-1}$ at a V_{DS} of 1.5 V with narrow distributions are achieved for monolayer MoS_2, WS_2 and WSe_2 FETs.

Figure 4g, h summarizes the state-of-the-art contact methods for both monolayer and thicker MoS_2. The Bi contact to monolayer MoS_2 yields the lowest R_C. We expect an even lower R_C for Bi contacts with thicker MoS_2 than monolayer[2,33,353436,37], and the I_{ON} of our monolayer transistor has already exceeded the previous record for multilayer transistors (Fig. 4e). Furthermore, our measured R_C values are comparable to those reported in commonly used three-dimensional semiconductors, approaching the quantum limit[1,2].

Finally, we propose that the Bi–TMD FET technology could be readily scaled down to sub-10 nm technology nodes and potentially meet the requirement of the International Roadmap for Devices and Systems (IRDS) 2024 targets of logic transistors (pentagons in Fig. 4f)[38]. Our Bi–TMD transistors establish a new benchmark for monolayer TMD FETs (Fig. 4e), and deliver comparable performance to modern silicon transistors with similar dimensions (diamonds in Fig. 4f)[39,40]. With the consideration of critical scaling rules such as the contact length (L_C) scaling and supply voltage (V_{DD}) scaling, we conclude that our low-R_C contact is essential for maximized I_{ON} in aggressively down-scaled transistors. First, as the current transfer length of the contact (L_T) for Bi contact is as small as ~7 nm (Extended Data Fig. 9b), the L_C-scaling-associated R_C increase and I_{ON} suppression are mitigated (Fig. 4f). Second, the much lower R_C ensures that the drain voltage is dropped mostly across the channel, which greatly reduces the required minimum V_{DD} at each technology node for driving the MoS_2 channel to the velocity saturation regime, in order to reach its maximum current (Extended Data Fig. 9c, d). The current saturation for the I_{DS}–V_{DS} characteristics, as well as the constant spacing between the saturated I_{DS} for the same overdrive voltage interval (Fig. 4a–c, Extended Data Fig. 7f), suggests that the Bi–TMD FETs can already work in velocity saturation when L_{CH} = 120–150 nm and V_{DD} = 1.5 V. This promises a maximized I_{ON} for FETs with even shorter channels and lower V_{DD} (blue dashed lines and open circles in Fig. 4f). The saturation velocity v_{sat}, which determines the maximum I_{ON}, is extracted to be ~2.5 × 10^6 cm s^{-1} (see Methods). As a value of I_{ON} of 1,135 μA $μm^{-1}$ is achieved experimentally for a 35-nm-channel device, we envision that I_{ON} exceeding 1,800 μA $μm^{-1}$ in a 10-nm-channel monolayer TMD FET at a n_{2D} of 4.5 × 10^{13} cm^{-2} (Fig. 4f) could be realized in the foreseeable future, reaching the industrial goal of next-generation transistor technologies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03472-9.

1. Chhowalla, M., Jena, D. & Zhang, H. Two-dimensional semiconductors for transistors. *Nat. Rev. Mater.* **1**, 16052 (2016).
2. Allain, A., Kang, J., Banerjee, K. & Kis, A. Electrical contacts to two-dimensional semiconductors. *Nat. Mater.* **14**, 1195–1205 (2015).
3. Akinwande, D. et al. Graphene and two-dimensional materials for silicon technology. *Nature* **573**, 507–518 (2019).
4. Louie, S. G. & Cohen, M. L. Electronic structure of a metal-semiconductor interface. *Phys. Rev. B* **13**, 2461–2469 (1976).
5. Nishimura, T., Kita, K. & Toriumi, A. Evidence for strong Fermi-level pinning due to metal-induced gap states at metal/germanium interface. *Appl. Phys. Lett.* **91**, 123123 (2007).
6. Kobayashi, M., Kinoshita, A., Saraswat, K., Wong, H.-S. P. & Nishi, Y. Fermi level depinning in metal/Ge Schottky junction for metal source/drain Ge metal-oxide-semiconductor field-effect-transistor application. *J. Appl. Phys.* **105**, 023702 (2009).
7. Sotthewes, K. et al. Universal Fermi-level pinning in transition-metal dichalcogenides. *J. Phys. Chem. C* **123**, 5411–5420 (2019).
8. Tung, R. T. The physics and chemistry of the Schottky barrier height. *Appl. Phys. Rev.* **1**, 011304 (2014).
9. Razavieh, A., Zeitzoff, P. & Nowak, E. J. Challenges and limitations of CMOS scaling for FinFET and beyond architectures. *IEEE Trans. NanoTechnol.* **18**, 999–1004 (2019).
10. Tersoff, J. Schottky barrier heights and the continuum of gap states. *Phys. Rev. Lett.* **52**, 465 (1984).
11. Sze, S. M. & Ng, K. K. *Physics of Semiconductor Devices* (Wiley, 2006).
12. Vilan, A., Shanzer, A. & Cahen, D. Molecular control over Au/GaAs diodes. *Nature* **404**, 166–168 (2000).
13. Wang, Y. et al. Van der Waals contacts between three-dimensional metals and two-dimensional semiconductors. *Nature* **568**, 70–74 (2019).
14. Cui, X. et al. Low-temperature ohmic contact to monolayer MoS_2 by van der Waals bonded Co/h-BN electrodes. *Nano Lett.* **17**, 4781–4786 (2017).
15. Liu, Y., Stradins, P. & Wei, S.-H. Van der Waals metal–semiconductor junction: weak Fermi level pinning enables effective tuning of Schottky barrier. *Sci. Adv.* **2**, e1600069 (2016).
16. Liu, Y. et al. Approaching the Schottky–Mott limit in van der Waals metal–semiconductor junctions. *Nature* **557**, 696–700 (2018).
17. Kim, C. et al. Fermi level pinning at electrical metal contacts of monolayer molybdenum dichalcogenides. *ACS Nano* **11**, 1588–1596 (2017).
18. English, C. D., Shine, G., Dorgan, V. E., Saraswat, K. C. & Pop, E. Improved contacts to MoS_2 transistors by ultra-high vacuum metal deposition. *Nano Lett.* **16**, 3824–3830 (2016).
19. Chee, S. S. et al. Lowering the Schottky barrier height by graphene/Ag electrodes for high-mobility MoS_2 field-effect transistors. *Adv. Mater.* **31**, 1804422 (2019).
20. Cao, Z., Lin, F., Gong, G., Chen, H. & Martin, J. Low Schottky barrier contacts to 2H-MoS_2 by Sn electrodes. *Appl. Phys. Lett.* **116**, 022101 (2020).
21. Das, S., Chen, H.-Y., Penumatcha, A. V. & Appenzeller, J. High performance multilayer MoS_2 transistors with scandium contacts. *Nano Lett.* **13**, 100–105 (2013).
22. Smithe, K. K., English, C. D., Suryavanshi, S. V. & Pop, E. Intrinsic electrical transport and performance projections of synthetic monolayer MoS_2 devices. *2D Mater.* **4**, 011009 (2016).
23. Qian, X., Liu, J., Fu, L. & Li, J. Quantum spin Hall effect in two-dimensional transition metal dichalcogenides. *Science* **346**, 1344–1347 (2014).
24. Nagao, T. et al. Nanofilm allotrope and phase transformation of ultrathin Bi film on Si(111)–7 × 7. *Phys. Rev. Lett.* **93**, 105501 (2004).

25. Zhong, H. et al. Interfacial properties of monolayer and bilayer MoS_2 contacts with metals: beyond the energy band calculations. *Sci. Rep.* **6**, 21786 (2016).
26. Kang, J., Liu, W., Sarkar, D., Jena, D. & Banerjee, K. Computational study of metal contacts to monolayer transition-metal dichalcogenide semiconductors. *Phys. Rev. X* **4**, 031005 (2014).
27. Chakraborty, B. et al. Symmetry-dependent phonon renormalization in monolayer MoS_2 transistor. *Phys. Rev. B* **85**, 161403 (2012).
28. Michail, A., Delikoukos, N., Parthenios, J., Galiotis, C. & Papagelis, K. Optical detection of strain and doping inhomogeneities in single layer MoS_2. *Appl. Phys. Lett.* **108**, 173102 (2016).
29. Moe, Y. A., Sun, Y., Ye, H., Liu, K. & Wang, R. Probing evolution of local strain at MoS_2–metal boundaries by surface-enhanced Raman scattering. *ACS Appl. Mater. Interfaces* **10**, 40246–40254 (2018).
30. Yang, L. et al. Chloride molecular doping technique on 2D materials: WS_2 and MoS_2. *Nano Lett.* **14**, 6275–6280 (2014).
31. Liu, W. et al. Role of metal contacts in designing high-performance monolayer n-type WSe_2 field effect transistors. *Nano Lett.* **13**, 1983–1990 (2013).
32. Yeh, C.-H., Cao, W., Pal, A., Parto, K. & Banerjee, K. in 2019 *IEEE International Electron Devices Meeting (IEDM)* 23.24.21–23.24.24 (IEEE, 2019); https://ieeexplore.ieee.org/abstract/document/8993600.
33. English, C. D., Smithe, K. K., Xu, R. L. & Pop, E. in *2016 IEEE International Electron Devices Meeting (IEDM)* 5.6.1–5.6.4 (IEEE, 2016); https://ieeexplore.ieee.org/abstract/document/7838355.
34. McClellan, C. J., Yalon, E., Smithe, K. K., Suryavanshi, S. V. & Pop, E. High current density in monolayer MoS_2 doped by AlO_x. *ACS Nano* **15**, 1587–1596 (2021).
35. Kwon, J. et al. Thickness-dependent Schottky barrier height of MoS_2 field-effect transistors. *Nanoscale* **9**, 6151–6157 (2017).
36. Li, S.-L. et al. Thickness scaling effect on interfacial barrier and electrical contact to two-dimensional MoS_2 layers. *ACS Nano* **8**, 12836–12842 (2014).
37. Wang, Q., Shao, Y., Gong, P. & Shi, X. Metal–2D multilayered semiconductor junctions: layer-number-dependent Fermi-level pinning. *J. Mater. Chem. C* **8**, 3113–3119 (2020).
38. Badaroglu, M. et al. More Moore. In *International Roadmap for Devices and Systems 2017* https://irds.ieee.org/images/files/pdf/2017/2017IRDS_MM.pdf (IEEE, 2017).
39. Ghani, T. et al. A 90-nm high volume manufacturing logic technology featuring novel 45-nm gate length strained silicon CMOS transistors. In *IEEE International Electron Devices Meeting 2003* 11.16.11–11.16.13 (IEEE, 2003); https://ieeexplore.ieee.org/abstract/document/1269442.
40. Thompson, S. et al. A 90-nm logic technology featuring 50-nm strained silicon channel transistors, 7 layers of Cu interconnects, low-*k* ILD, and 1 μm^2 SRAM cell. In *International Electron Devices Meeting 2002* 61–64 (IEEE, 2002); https://ieeexplore.ieee.org/abstract/document/1175779.
41. Kim, S. et al. High-mobility and low-power thin-film transistors based on multilayer MoS_2 crystals. *Nat. Commun.* **3**, 1011 (2012).
42. Liu, Y. et al. Pushing the performance limit of sub-100-nm molybdenum disulfide transistors. *Nano Lett.* **16**, 6337–6342 (2016).
43. Nourbakhsh, A. et al. MoS_2 field-effect transistor with sub-10-nm channel length. *Nano Lett.* **16**, 7798–7806 (2016).
44. Yang, L., Lee, R., Rao, S. P., Tsai, W. & Ye, P. in *2015 73rd Annual Device Research Conference (DRC)* 237–238 (IEEE, 2015).
45. Jung, Y. et al. Transferred via contacts as a platform for ideal two-dimensional transistors. *Nature Electron.* **2**, 187–194 (2019).
46. Nguyen, L. D., Tasker, P. J., Radulescu, D. C. & Eastman, L. F. Characterization of ultra-high-speed pseudomorphic AlGaAs/InGaAs (on GaAs) MODFETs. *IEEE Trans. Electron Dev.* **36**, 2243–2248 (1989).
47. Smithe, K. K., Suryavanshi, S. V., Muñoz Rojo, M., Tedjarati, A. D. & Pop, E. Low variability in synthetic monolayer MoS_2 devices. *ACS Nano* **11**, 8456–8463 (2017).
48. Yue, D., Kim, C., Lee, K. Y. & Yoo, W. J. Ohmic contact in 2D semiconductors via the formation of a benzyl viologen interlayer. *Adv. Funct. Mater.* **29**, 1807338 (2019).
49. Guimarães, M. H. et al. Atomically thin ohmic edge contacts between two-dimensional materials. *ACS Nano* **10**, 6392–6399 (2016).
50. Smets, Q. et al. Ultra-scaled MOCVD MoS_2 MOSFETs with 42 nm contact pitch and 250 μA/μm drain current. In *2019 IEEE International Electron Devices Meeting (IEDM)* 23.2.21–23.2.24 (IEEE, 2019).
51. Gao, J. et al. Transition-metal substitution doping in synthetic atomically thin semiconductors. *Adv. Mater.* **28**, 9735–9743 (2016).

Methods

MOCVD of monolayer MoS_2

Monolayer MoS_2 films are grown using low-pressure metal–organic chemical vapour deposition (MOCVD). Molybdenum hexacarbonyl ($Mo(CO)_6$, 98%, Sigma Aldrich) and diethyl sulfide ($C_4H_{10}S$, 98%, Sigma Aldrich) are selected as the precursors of molybdenum (Mo) and sulfur (S), respectively. With argon (Ar) as the carrier gas, the precursors are supplied in the vapour form into the chamber using a homemade bubbler system. The monolayer MoS_2 films are deposited on 300-nm-thick SiO_2/Si wafers at 320 °C for 15 h with flow rates of 100 standard cubic centimetres per minute (sccm) for Ar, 0.6 sccm for $Mo(CO)_6$, and 2.0 sccm for $C_4H_{10}S$. The typical Raman spectrum for the as-grown MOCVD monolayer MoS_2 is shown in Extended Data Fig. 5a.

CVD of monolayer MoS_2

Perylene-3,4,9,10-tetracarboxylic potassium salt molecules are used as the seeding promoter and are coated onto two clean SiO_2/Si pieces, serving as the seed reservoirs to provide the seeding molecules during the MoS_2 growth. The target substrate of a 300-nm-thick SiO_2/Si wafer is suspended between those two seed reservoirs. All of these three substrates are faced down and placed on a crucible containing molybdenum oxide (MoO_3, 99.98%) powder precursor. This MoO_3 precursor is put in the middle of a quartz tube reaction chamber and another sulfur powder (99.98%) precursor is placed upstream in the quartz tube. Before heating, the CVD system is purged using 1,000 sccm of Ar (99.999% purity) for 5 min. Next, the flow rate of Ar is switched to 20 sccm as the carrier gas for the MoS_2 growth, and the temperature of the reaction chamber is increased to 625 °C at a rate of 30 °C min^{-1}. The monolayer MoS_2 is synthesized at 625 °C for 3 min under atmospheric pressure.

CVD of monolayer WS_2

A simple method for deposition of monolayer WS_2 crystals at atmospheric pressure is used. Tungsten trioxide (WO_3) powder is sprayed onto a piece of SiO_2/Si wafer, acting as a WO_3 reservoir during the deposition. The SiO_2/Si target substrate is positioned face-up and downstream, 1 cm away from the WO_3 reservoir. A crucible containing sulfur powder is placed upstream. Prior to the growth, the reaction chamber is purged using 1,000 sccm of Ar for 5 min. Then the furnace temperature is ramped to 800 °C at a rate of 39 °C min^{-1} and the deposition of monolayer WS_2 crystals is implemented at 800 °C for 5 min with 50 sccm of Ar carrier gas.

CVD of monolayer WSe_2

An $(NH_4)_2WO_4$ aqueous solution (2 mg ml^{-1}) is spin-coated (2,500 rpm for 1 min) onto a SiO_2/Si substrate, and then placed in the centre of the furnace. Se powder (30 mg, 99.5%, Sigma Aldrich) is loaded upstream about 17 cm away from the centre. Before the growth, the tube is flushed with 300 sccm of Ar for 10 min to eliminate residual oxygen and moisture. During the growth, the temperature is increased to 900 °C at a rate of 50 °C min^{-1} and the growth lasts for 10 min with 30 sccm of Ar and 10 sccm of H_2 flow as the carrier gases. After the growth, the furnace is rapidly cooled to room temperature.

Mechanically exfoliation of monolayer WS_2 and WSe_2

Monolayer WS_2 and WSe_2 flakes are mechanically exfoliated onto the 100-nm-thick SiN_x dielectrics by the standard scotch tape technique. Prior to device fabrication, the exfoliated TMD flakes are immersed in acetone for 3 h to remove the tape residues. Raman spectroscopy are performed on the selected TMD flakes to confirm their monolayer characteristics for further device fabrication as shown in Extended Data Fig. 5a.

Transfer of monolayer TMDs on dielectric/Si substrates

The monolayer TMD crystals grown by MOCVD or CVD are transferred onto the dielectric/p^{++}-Si substrates for device fabrication using a wet transfer process. First, poly(methyl methacrylate) (PMMA) is spin-coated onto the monolayer TMD samples. Then, the PMMA/TMD stacks are released from the SiO_2/Si growth substrate by etching in a concentrated potassium hydroxide (KOH) aqueous solution at 90 °C. The freestanding PMMA/TMD stacks are picked up, rinsed with deionized water three times for 2 h, and then attached onto the target substrates. To dry the samples and enhance the adhesion, the PMMA/TMD stacks are baked on the hotplate at 70 °C for 20 min and 130 °C for another 20 min. Finally, the sample is immersed in cold acetone for at least 6 h to remove the PMMA.

Device fabrication and characterization

Monolayer TMD crystals are confirmed by Raman and photoluminescence characterization and then selected for transistor fabrication. Electron-beam (e-beam) lithography is used to define the channel and the source/drain contacts with PMMA e-beam resists (MicroChem). Metallization is implemented by e-beam evaporation of 20-nm bismuth with a well controlled deposition rate of 0.5 Å s^{-1}, followed by an Au capping layer (10–100 nm at 2 Å s^{-1}) at ~10^{-6} torr. Lift-off process is carried out in hot acetone. No annealing or chemical doping treatment is performed on the devices. The channel widths for the devices in this study are in the range of 2 to 10 μm. All electrical characterization is conducted in a vacuum environment (10^{-5}–10^{-6} torr) in a Lakeshore probe station using a semiconductor parameter analyser (Keysight B1500A). The electrical resistivity of the evaporated bismuth film on monolayer MoS_2, and on SiN_x are measured to be 9.0×10^{-6} Ω m and 9.5×10^{-6} Ω m, respectively. The gate dielectrics for the monolayer TMD transistors studied in this work include 300-nm-thick SiO_2 (NOVA Electronic Materials) and 100-nm-thick SiN_x (MTI Corporation). To estimate the sheet carrier density and carrier mobility of the devices accurately, the capacitances for the dielectrics are measured at 1 MHz with a power device analyser (Keysight B1505A) on separate metal–insulator–metal capacitors at room temperature.

Sample preparation for Raman and XPS characterizations

First, a 20-nm bismuth thin film is deposited on a continuous monolayer MoS_2 film grown on a SiO_2/Si wafer followed by an Au capping layer using e-beam evaporation. Next, the heterostructure of Au–Bi–MoS_2 can be peeled off by a thermal tape, as the MoS_2 adheres more strongly to the thermal tape than to the silica substrate. Finally, the sample is inverted to expose the continuous MoS_2 film on top of the bismuth film for characterization. In this way, a pristine Bi–MoS_2 interface can be studied without oxidation of bismuth. This method allows us to directly carry out the Raman and XPS characterizations on the Bi–MoS_2 interface.

Raman spectroscopy

Raman spectroscopy of monolayer MoS_2 flakes is carried out on a confocal Raman system (HR800, Horiba Scientific) with a laser wavelength of 523 nm at a laser power of 2.5 mW and accumulation time of 0.5 s. The emitted Stokes Raman signal is collected by a 0.9 numerical aperture of 100× objective (Carl Zeiss Microscopy) with a 1,800 lines per mm grating for the measurements. The spectrum is calibrated by the silicon characteristic peak at 520.6 cm^{-1} from an undoped silicon wafer.

XPS analysis

The XPS measurement is carried out by using a PHI Versaprobe II XPS instrument with monochromated Al Kα source (1,486.6 eV) and a spot size of 200 μm. A 50-W gun power and 15-kV operation voltage are used during spectrum acquisition. During the measurement, samples are flooded with electron and Ar ion guns to compensate the surface charging. All the XPS spectra presented in this work are calibrated by the C1 s peak at 284.8 eV. The XPS spectra are analysed and fitted by Gaussian/Lorentzian mix function.

TEM analysis

The TEM is performed using an FEI Tecni (G2 Spirit TWIN) under 120 kV. The SAED images are taken with a 1-μm selected-area aperture. The Bi/MoS_2 freestanding sample is prepared by direct e-beam evaporation of Bi and Au on top of freestanding monolayer MoS_2 sample on a Protochips C-Flat TEM grid (2/4) as illustrated in Extended Data Fig. 3a, e. The streaks connecting the diffraction pattern of Bi in SAED originate from the diffusive scattering of grain boundaries of Bi.

Extraction of Schottky barriers

A 2D Schottky FET can be regarded as two Schottky diodes connected back-to-back. Most of the applied drain-to-source voltage (V_{DS}) drops at the reverse-biased contact. Therefore, for an n-channel FET the transistor behaviour is dominated by the source side. The drain current density (I_{DS}, in units of μA μm^{-1}) thermally injected from the metal contact into the 2D channel through a reverse-biased Schottky barrier can be expressed as:

$$I_{DS} = A^*_{2D} T^{1.5} \exp\left(-\frac{\Phi_B}{k_B T}\right)\left[1 - \exp\left(\frac{-V_{DS}}{k_B T}\right)\right], \quad (1)$$

where $A^*_{2D} = q(8\pi k_B^3 m^*)^{0.5}/h^2$ is the Richardson constant for a 2D system (m^*, electron effective mass), T is the temperature, k_B is Boltzmann's constant, q is the elementary charge, and Φ_B is the effective contact barrier height at a given gate–source voltage (V_{GS}). If $V_{DS} \gg k_B T$, equation (1) can be simplified to

$$I_{DS} \approx A^*_{2D} T^{1.5} \exp\left(-\frac{\Phi_B}{k_B T}\right). \quad (2)$$

In this way, the effective energy barrier at a given V_{GS} can be extracted by finding the slope in the Arrhenius plots, as shown in Extended Data Fig. 2a, using the following equation:

$$\ln(I_{DS}/T^{1.5}) = \frac{-\Phi_B}{k_B T} + c, \quad (3)$$

where c is a constant. Φ_{SB} is then extracted at the flatband condition $(V_{GS} = V_{FB})^2$, as shown in Extended Data Fig. 1c, f, i.

However, the Arrhenius plots of Bi-contacted MoS_2 FETs display an opposite trend to the thermionic emission model (equation (3)). As shown in Extended Data Fig. 2b, the Arrhenius plots of the Bi–MoS_2 transistors can be divided into two regimes: (i) a positive slope regime where the mobility increases with a decreasing temperature (150–300 K) and (ii) a saturation-like regime where the mobility reaches constant at even lower temperatures (77–150 K). For an ideal transistor with a zero contact barrier, the thermionic emission model gives rise to a zero slope in the Arrhenius plot when ignoring the contribution of the channel resistance. However, in our devices the drain current is dominated by the channel resistance, so the temperature dependence of the MoS_2 mobility needs to be considered, which contributes to the increase in drain current density (I_{DS}). Therefore, the positive slopes within the range of 150–300 K originate from the enhancement of MoS_2 mobility owing to the reduced phonon scattering. Once the mobility gradually reaches a constant, owing to scattering of long-range Coulomb impurities or short-range atomic defects in the range of 77–150 K, the slopes of the Arrhenius plots tend to saturate, which implies the contact barrier-free nature of the Bi–MoS_2 FETs.

Extraction of contact resistance through transfer-length method (TLM)

In a two-terminal device, the major resistance components originate from the contact resistance (R_C) and the channel resistance (R_{CH}). As a result, the total device resistance (R_{TOT}, in units of kΩ μm) normalized by channel width (W) can be expressed as $R_{TOT} = 2R_C + R_{CH} = 2R_C + R_{SH}L_{CH}$, where R_{SH} is the sheet resistance of the semiconductor channel (in units of kΩ per square, kΩ □$^{-1}$) and L_{CH} is the channel length. The total device resistance varies linearly with the L_{CH} if R_C (in units of kΩ μm) and R_{SH} are spatially homogeneous in the device. Therefore, by measuring the total resistances of the devices with various L_{CH}, the R_{TOT} can be plotted as a function of L_{CH}. The residual resistance at $L_{CH} = 0$ corresponds to the total contact resistance ($2R_C$) of the device.

Accordingly, we extract the R_C of the Bi–MoS_2 transistors for a given carrier density (n_{2D}) by plotting R_{TOT} versus L_{CH} as shown in Fig. 2c and Extended Data Fig. 2c, d. The vertical intercept at $L_{CH} = 0$ of a linear fit yields the $2R_C$ for the two-terminal Bi–MoS_2 devices. Also, the R_{SH} of the MoS_2 channel for a certain n_{2D} can be calculated from the slope of the linear fit. The effective mobility is then calculated by $\mu = 1/(qn_{2D}R_{SH})$. The n_{2D} induced by electrostatic gating is estimated by assuming a simple linear charge dependence on the gate voltage overdrive $n_{2D} = C_{ox}(V_{GS} - V_T)/q$, where C_{ox} is the capacitance per unit area of the gate dielectric and is experimentally characterized by standard capacitance–voltage (C–V) measurement ($C_{ox} \approx 6 \times 10^{-8}$ F cm^{-2} for the 100-nm-thick SiN_x dielectrics used in this study), and V_T is the threshold voltage linearly extrapolated from the transfer characteristic curve of the device. As shown in Fig. 2c and Extended Data Fig. 2d, the good linear fits to the plot of R_{TOT} versus L_{CH} indicate uniform channel materials and electrical contacts.

The accuracy of the R_C extraction can be improved by: (I) a more efficient gate with higher gate capacitance (100-nm SiN_x instead of 300-nm SiO_2), so that the carrier density—and thus the sheet resistance (slopes of Fig. 2c and Extended Data Fig. 2d)—can be substantially reduced; (II) shorter channel lengths so that the data points are closer to the y-axis intersection ($2R_C$); and (III) samples with minimal variation in terms of V_T and μ. With the consideration of these factors we estimated the mean and the fitting uncertainty of the R_C value of our best Bi–MoS_2 device to be 123 ± 63 Ω μm (mean ± 1σ).

Extraction of field-effect mobility from two-terminal and four-terminal devices

From the transfer curves (I_{DS}–V_{GS}) of a two-terminal MoS_2 device, the field-effect mobility is extracted using the expression $\mu = 1/(qn_{2D}R_{SH}) \approx (dI_{DS}/dV_{GS}) \times [L_{CH}/(WC_{ox}V_{DS})]$. Note that this field-effect mobility typically represents the lower limit because of contact resistance in the devices.

In a four-terminal configuration, a bias current (I_{DS}) is applied between the two outer electrodes (D and S), while the voltages on the two inner electrodes are measured (V_1 and V_2). Ideally, the current path in the material should not be affected by the inner sense electrodes, which allows for an accurate assessment of the channel resistivity and thus the intrinsic mobility. We calculate the channel resistance as $R_{SH} = (V_2 - V_1)/I_{DS} \times (W/L_{21})$, where L_{21} is the length between the inner voltage contacts. The four-terminal field-effect mobility can be calculated by $\mu = 1/(qn_{2D}R_{SH}) = (dI_{DS}/dV_{GS}) \times [L_{21}/(WC_{ox}V_{21})]$, where V_{21} is the voltage difference between the two inner electrodes.

First-principles calculations

We build a model with three layers of Bi with its (0001) surface in contact with monolayer MoS_2. A vacuum slab of 10 Å is used for separating the cells to mimic the 2D system. The layer thickness of Bi is proved to be adequate given that very little electronic structure is changed at the second Bi layer (counting from Bi–MoS_2 interface), and almost no electronic structure is changed on the third layer demonstrated in the differential charge plot in Fig. 3c. The supercell of MoS_2 is a 5 × 5 replicate of its unit cell, and the supercell of Bi is a 3 × 3 replicate of its redefined unit cell. The lattice mismatch between these two supercells is only 0.1%, making the structure stable during the ionic relaxation steps in the first-principles calculations.

The first-principles calculations for geometric optimization and the electronic properties of crystal structure are carried out using DFT and projector augmented wave (PAW) method implemented in the

Vienna Ab-initio Simulation Package (VASP)[52]. The semilocal generalized gradient approximation in the form of Perdew–Burke–Ernzerhof (PBE), and the PAW pseudopotentials are adopted. During the ionic optimization steps, the Hellman–Feynman forces of single atoms are optimized to be less than 0.02 eV Å, where the energy cutoff is set to be 400 eV and only gamma point is used for *k*-space sampling. In the DOS calculation, *k* points for the supercell are chosen to be $9 \times 9 \times 1$ and an energy cutoff is set to be 400 eV, and the single electronic step is converged to 1×10^{-5} eV. Dipole correction along the *z* axis is implemented in all the DFT calculations, which is especially important for electrostatic potential calculations. The orbital PLDOS is a projection of total DOS into each orbital and further into the sphere of atoms natively defined in VASP code. Spin–orbit coupling is not included in generating Fig. 3 but has been tested in a smaller Bi–MoS_2 system where the Fermi level is proved to be still located at the conduction band minimum. The one-dimensional electrostatic charge potential along the *z* axis is calculated from Poisson's equation where the charge density per unit length is acquired by integrating the charge along the *xy* plane from the DFT calculation. According to Bader charge analysis, there are 0.88 electrons per supercell transferred from Bi to MoS_2, which is equivalent to an electron doping of 4×10^{11} cm^{-2}.

Analysis of contact resistance with transmission line model

The contact resistance (R_C) measured from the TLM originates from two components: (I) transport through the metal–semiconductor energy barrier, and (II) lateral access resistance under the contact due to the sheet resistance (R_{SH}) of the channel material. We employ the transmission line model for R_C, expressed as[11,33]:

$$R_C = \sqrt{\rho_C R_{SH}} \coth\left(\frac{L_C}{L_T}\right) \approx \sqrt{\rho_C R_{SH}}, \text{ if } L_C \gg L_T, \quad (4)$$

where R_C is normalized by W in units of Ω μm, ρ_C is the specific contact resistivity accounting for the vertical interlayer transport under the contact, L_T is the current transfer length, and L_C is the physical contact length (1 μm in this study). R_{SH} here is extracted from the slopes of the TLM plots.

We estimate that $\rho_C \approx 8.8 \times 10^{-9}$ Ω cm^2 at $n_{2D} = 1.5 \times 10^{13}$ cm^{-2} for the Bi–MoS_2 FETs on 100-nm-thick SiN_x at room temperature. Based on the definition of the current transfer length L_T

$$L_T = \sqrt{\frac{\rho_C}{R_{SH}}}, \quad (5)$$

we evaluate the best L_T to be around 7 nm at room temperature, much smaller than the L_C used in our devices, justifying the use of the approximation in equation (4). These results suggest that the dimension of the contacts for the Bi–MoS_2 FETs can be reduced to ~7 nm without performance degradation resulting from the current-crowding effect.

Extended Data Fig. 9a plots the fractions of the total contact resistance ($2R_C$) and the intrinsic channel resistance (R_{CH}) with respect to the total device resistance (R_{TOT}) as a function of L_{CH} using $R_C = 123$ Ω μm and mobility ~20 cm^2 V^{-1} s^{-1} as extracted from our Bi–MoS_2 FETs on 100-nm SiN_x (Fig. 2c). Extended Data Fig. 9b shows $2R_C$ versus L_{CH} for different contact technologies. It can be seen that R_C does not dominate the performance of Bi–MoS_2 FETs until the L_{CH} reaches ~7 nm, providing a substantial improvement in the scaling limit of TMD transistors.

Equivalent circuit model for the contact resistance of Bi–MoS_2 FET

Based on the first-principles calculation and the experimental results, we built an equivalent circuit for the Bi–MoS_2 contact region (Fig. 3d). The contact resistance is composed of a network of the van der Waals gap tunnelling resistors (with tunnelling-specific resistivity, ρ_t) and the Bi-contacted MoS_2 resistors (with sheet resistance $R_{SH,C}$). Because the barrier height at the Bi–MoS_2 interface is negligible, we do not consider the contact resistance contribution due to the thermionic emission at the Schottky barrier. The tunnelling current density (J_t) through the van der Waals barrier can be obtained through Simmon's model[53,54],

$$J_t = \frac{q}{4\pi^2 \hbar w_t^2}\left\{\left(\Phi_t - \frac{qV}{2}\right)\exp\left[-2\frac{(2m_e)^{1/2}}{\hbar}\alpha w_t\left(\Phi_t - \frac{qV}{2}\right)^{1/2}\right] - \left(\Phi_t + \frac{qV}{2}\right)\exp\left[-2\frac{(2m_e)^{1/2}}{\hbar}\alpha w_t\left(\Phi_t + \frac{qV}{2}\right)^{1/2}\right]\right\}, \quad (6)$$

where w_t is the tunnelling gap width, Φ_t is the tunnelling barrier height, α is an empirical factor that is associated with the shape of the barrier ($\alpha = 1$ for an ideal square barrier), V is the bias voltage, q is the electron charge, $\hbar$ is the reduced Planck's constant, and m_e is the free-electron mass. At low bias ($qV \ll \Phi_t$), the tunnelling-specific resistivity is given by

$$\rho_t = \left(\frac{dJ_t}{dV}\right)^{-1} \approx \frac{4\pi^2 \hbar w_t^2}{q^2} \frac{\exp\left(2\frac{(2m_e)^{1/2}}{\hbar}\alpha w_t \Phi_t^{1/2}\right)}{\frac{(2m_e)^{1/2}}{\hbar}\alpha w_t \Phi_t^{1/2} - 1}, \quad (7)$$

According to our DFT calculation results, $w_t \approx 1.66$ Å and $\Phi_t \approx 3.6$ eV. We therefore estimate that $\rho_t \approx 1.81 \times 10^{-9}$ Ω cm^2 for our Bi–MoS_2 contact.

In addition, $R_{SH,C}$ is estimated to be 15.6 kΩ assuming that the 2D carrier density of Bi-contacted MoS_2 is around 2×10^{13} cm^{-2}. Then the contact resistance is $R_C = (\rho_t R_{SH,C})^{1/2} + R_{SB0}$, where $R_{SB0} = k_B T/(qA^*_{2D} T^{3/2})$ is the residual contact resistance (according to the thermionic model, equation (1)) for a zero Schottky barrier height. The contact resistance for Bi–MoS_2 is calculated to be 0.13 kΩ μm, in very good agreement with our measured values.

Velocity saturation, critical channel length and scaling rules in Bi–TMD transistors

In the linear regime, the transistor on-state current density (I_{ON}) is approximately determined by the total device resistance ($R_{TOT} = 2R_C + R_{CH}$), and I_{ON} increases linearly with the applied V_{DS} at a given R_{CH} which can be modulated by the gate voltage. By taking the contact resistance into account, the effective drain-to-source voltage (V'_{DS}) dropped across the TMD channel is $V'_{DS} = V_{DS} - 2R_C I_{ON}$. As the V_{DS} increases and/or the channel length reduces, at a certain point where the lateral field becomes greater than the critical field strength (F_C) in the TMD material, the conducting electron in the TMD channel is accelerated to its saturation velocity (v_{sat}) and the I_{ON} saturates to a maximum of $I_{ON} = n_{2D} q v_{sat}$. In this velocity saturation regime, the I_{ON} in the transistor thus scales only linearly with n_{2D} induced into the TMD channel through electrostatic gating. Accordingly, for I_{ON} of 450 μA μm^{-1} shown in Fig. 4a, the v_{sat} of monolayer MoS_2 is extracted to be ~2.5×10^6 cm s^{-1} at an n_{2D} of ~10^{13} cm^{-2}.

At a fixed V_{DS} and gate bias, there is a transition from the linear regime to the velocity saturation regime when the L_{CH} of the transistor reduces to the critical channel length, L_{cr}, and the TMD channel reaches its critical field strength

$$F_C = \frac{V'_{DS}}{L_{cr}} = \frac{V_{DS} - 2R_C n_{2D} q v_{sat}}{L_{cr}}. \quad (8)$$

Taking MoS_2 as an example, F_C is 1.15×10^5 V cm^{-1} (ref. [55]). Therefore, for a Bi–MoS_2 transistor with $R_C \approx 123$ Ω μm biased at a V_{DS} of 1.5 V with an n_{2D} of 1.5×10^{13} cm^{-2}, the L_{cr} is ~117 nm, which agrees with our experimental results. On the other hand, with the same device dimension and bias conditions, a higher R_C would drive the TMD channel to operate below F_C, or in linear regimes, with smaller I_{DS} than the saturation current; therefore, a larger V_{DS} is required to reach the maximum I_{ON} (saturation current)

at a similar device dimension, limiting the allowed minimum power supply voltage.

For future transistor scaling, a low supply voltage (V_{DD}) is required. We further consider the minimum V_{DS} to bias the Bi–MoS_2 transistor in its velocity saturation regime while satisfying the IRDS target I_{ON} for high-performance and low-power logic transistors. Owing to the short current transfer length of the Bi contact ($L_T \approx 7$ nm), the R_C would remain low without major current crowding even with a reduced contacted gate pitch (CGP = $L_{CH} + L_C$) used in the future technology node (that is, $L_T < L_C$)[38]. Thus, the minimum V_{DS} for Bi–TMD transistors to meet the target I_{ON} in different future technology nodes is shown in Extended Data Fig. 9c, d:

$$V_{DS,\ min} = F_C L_{CH} + 2R_C I_{ON}. \quad (9)$$

The projection, as shown in Fig. 4e, suggests that the Bi contact to TMDs can potentially meet the IRDS requirements for future energy-efficient electronics. An n_{2D} of 4.5×10^{13} cm^{-2} is assumed based on the use of a 3-nm ultrathin gate dielectric with a dielectric constant (κ) of 15 and an overdrive of 1.6 V. This condition corresponds to a vertical electric field of 5.3 MV cm^{-1} in the dielectric, which lies below its breakdown strength (F_{BD}) of 6.2 MV cm^{-1} ($F_{BD} = 35\kappa^{-0.64}$)[56].

Bi contacts for other monolayer TMDs and effects of TMD quality

We would like to point out that a high sample quality is a prerequisite for the proposed gap-state saturation mechanism. Both our DFT calculation (Extended Data Fig. 4b) and experiment observations (Extended Data Fig. 8 and Extended Data Table 1) indicate that a high density of structural defects such as chalcogen vacancies tend to obstruct the formation of ohmic contact while the gap-state pinning mechanism becomes dominant at the contact interface.

To study the effects of material quality, monolayer TMD crystals with different sample conditions are contacted with Bi electrodes and their electrical characteristics are measured. Since CVD normally exhibits a high variation in local concentrations of precursors along the growth substrate, CVD-grown TMD crystals typically show a larger variation in the sample quality. Extended Data Fig. 8a shows the output characteristics of a device based on a CVD-grown monolayer MoS_2 crystal possessing poor sample quality (that is, non-clean surface with curved edges). The lower I_{DS} and the nonlinear I_{DS}–V_{DS} curves resulting from this device suggest the presence of a contact barrier and imply that gap-state pinning takes over the band-alignment mechanism at the MoS_2 surface, which is further confirmed by DFT calculation (Extended Data Fig. 4b). In our experiments, we note that MOCVD-grown MoS_2 crystals exhibit a higher homogeneity and reproducibility and low variation in sample quality, probably owing to the well controlled flow rate of the precursors during the deposition (Extended Data Fig. 8b).

For the case of monolayer WSe_2, similar variation in the contact performance is also observed in CVD WSe_2 devices. The aged sample with a defective surface (that is, holes and cracks) even turns to p-type conduction with a low I_{DS}, which manifests the strong gap-state pinning at the metal–TMD interface where the Fermi level is aligned closer to the valence band of WSe_2, as shown in Extended Data Fig. 8e, f. This gap-state pinning effect is mitigated when a WSe_2 crystal with reasonable quality is used and the device behaviour changes to ambipolar, indicating that the Fermi level is pinned upward at a position closer to the conduction band minimum (Extended Data Fig. 8g, h). As Bi electrodes are in contact with a high-quality, freshly exfoliated WSe_2, the ohmic characteristics become prevalent at the Bi–WSe_2 interface, giving rise to a good ohmic contact and the considerably enhanced n-type conduction (Extended Data Fig. 8i, j). Note that good ohmic contacts can be also formed on CVD monolayer TMDs when a high-quality sample is obtained (Extended Data Fig. 8c, d).

Data availability

All data needed to evaluate the conclusions herein are present in the Article.

52. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).
53. Simmons, J. G. Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film. *J. Appl. Phys.* **34**, 1793–1803 (1963).
54. Simmons, J. G. Electric tunnel effect between dissimilar electrodes separated by a thin insulating film. *J. Appl. Phys.* **34**, 2581–2590 (1963).
55. Fiori, G., Szafranek, B. N., Iannaccone, G. & Neumaier, D. Velocity saturation in few-layer MoS2 transistor. *Appl. Phys. Lett.* **103**, 233509 (2013).
56. Kim, J. J. et al. Intrinsic time zero dielectric breakdown characteristics of HfAlO alloys. *IEEE Trans. Electron Dev.* **60**, 3683–3689 (2013).

Acknowledgements P.-C.S., J.B. and J.K. acknowledge financial support from the Center for Energy Efficient Electronics Science (NSF award no. 0939514), which provided funding for development of high-performance monolayer TMD transistors. P.-C.S., Y.L., J.-H.P., A.-Y.L., T.P. and J.K. acknowledge the US Army Research Office through the Institute for Soldier Nanotechnologies at MIT, under cooperative agreement no. W911NF-18-2-0048. C.S., J.-H.P., J.W. and J.K. acknowledge the support from the US Army Research Office (ARO) under grant no. W911NF-18-1-0431. C.S. is currently supported by the Kavli Energy NanoScience Institute/Heising–Simons Fellowship, Berkeley, California, USA. C.S. and A.Z. are supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, of the US Department of Energy under contract no. DE-AC02-05-CH11231, within the sp2-Bonded Materials Program (KC2207), which provided for TEM characterizations. C.S. and A.Z. are further supported by the National Science Foundation under grant no. DMR-1807233, which provided funding for development of TEM image-processing methods. J.L. acknowledges support by the Office of Naval Research MURI through grant no. N00014-17-1-2661. Y.L. and J.B. acknowledge support by the Office of Naval Research MURI programme N00014-16-1-2921, and the NSF award RAISE TAQS under grant no. DMR-1839098. Y.L. and N.M. acknowledge support by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under award DE-SC0020042. A.-S.C. and C.-I.W. acknowledge support from the Ministry of Science and Technology of Taiwan (MOST 108-2622-8-002-016). J.W. and J.K. acknowledge support from the joint development project (JDP) from TSMC. V.T. and M.-H.C. are indebted to the support from the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under award no: OSR-2018-CARF/CCF-3079. H.-L.T. acknowledges partial support from the Ministry of Science and Technology of Taiwan (MOST-108-2917-I-564-036). We acknowledge H.-S.P. Wong for guidance. We thank Y. Guo, E. Shi and H. Wang for technical assistance with materials characterizations, and Y.-T. Shao for discussions on SAED pattern simulation.

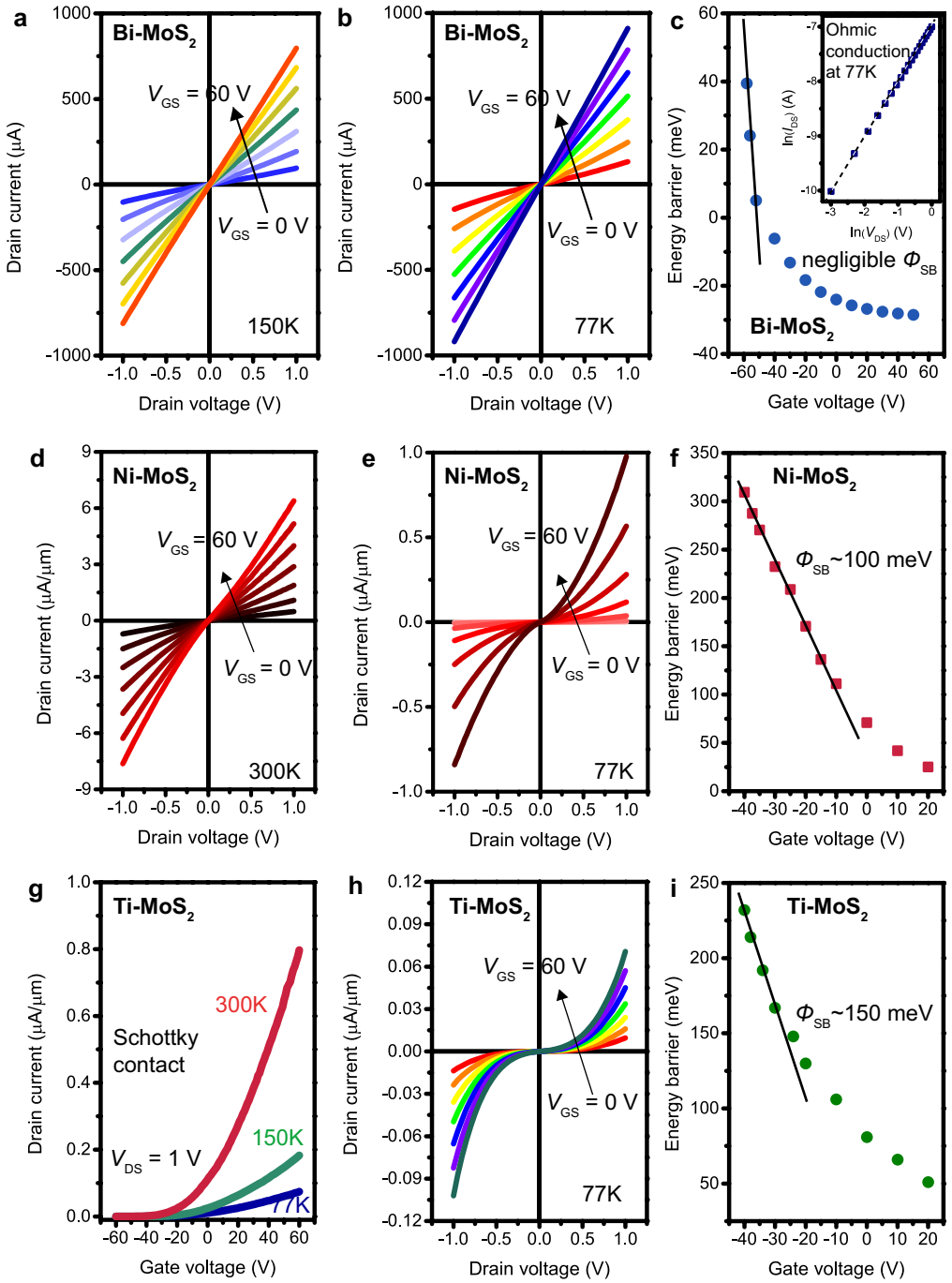
Author contributions J.K. and L.-J.L. supervised the project. P.-C.S. and J.K. proposed the project. P.-C.S., C.S., Y.L. and J.K. designed the experiments. P.-C.S. carried out the device fabrication. P.-C.S., H.-L.T. and Y.L. performed the electrical characterization supervised by T.P. C.S. carried out the TEM measurements and analysis and first-principles calculations supervised by A.Z. and J.L. P.-C.S., Y.L. and C.S. conducted the device modelling and data analysis. A.-S.C., C.-C.C. and G.P. carried out additional fabrication and characterization of the short-channel devices supervised by L.-J.L. The work of A.-S.C. is also co-supervised by C.-I.W. Y.L. and J.W. performed the SEM measurements. J.-H.P., P.-C.S., Z.C. and N.M. contributed to the growth, exfoliation and transfer of materials supervised by J.K. M.-H.C., A.-Y.L., M.M.T., and P.-C.S. carried out the materials characterizations. P.-C.S., C.S., Y.L. and J.K. wrote the manuscript. All authors discussed the results and provided constructive comments on the manuscript.

Competing interests P.-C.S. and J.K. are co-inventors on a patent application (provisional filling number US 63/024,141) related to the research presented in this paper.

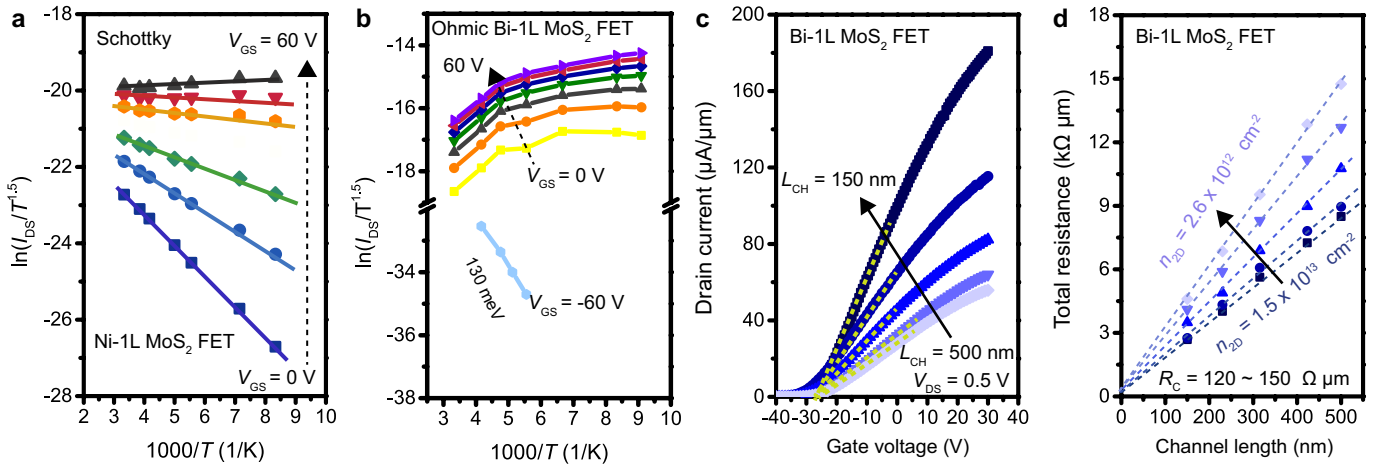
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03472-9.

Correspondence and requests for materials should be addressed to P.-C.S., L.-J.L. or J.K.

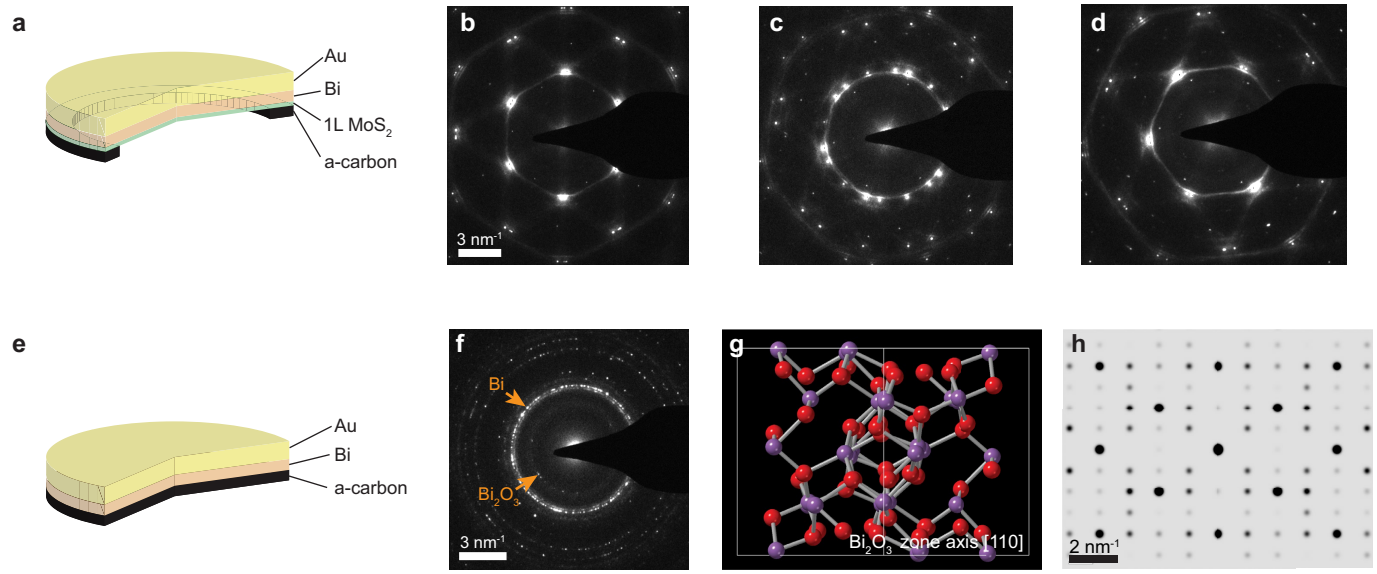
Peer review information *Nature* thanks Hyeon-Jin Shin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



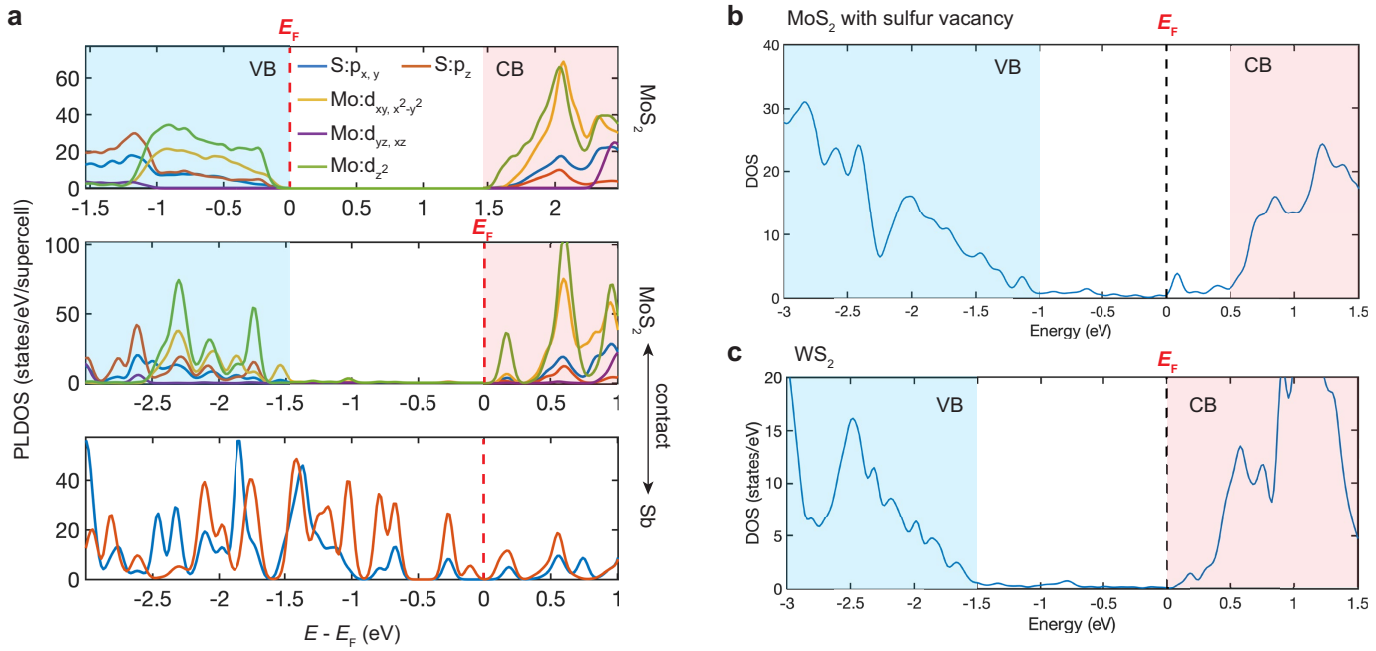
Extended Data Fig. 1 | Temperature-dependent electrical characteristics. **a**, **b**, Typical I_{DS}–V_{DS} curves at 150 K (**a**) and 77 K (**b**) for the Bi–MoS_2 FET. The device exhibits linear output characteristics at all the temperatures measured. **c**, Schottky barrier height (Φ_{SB}) extraction for the Bi–MoS2 FET, showing a negligible contact barrier. Inset, logarithmic plot of the I_{DS}–V_{DS} curve at 77 K and $n_{2D} \approx 4 \times 10^{12}\,cm^{-2}$, demonstrating ohmic contact in the Bi–MoS_2 FETs. **d**, **e**, Typical I_{DS}–V_{DS} curves at room temperature (**d**) and 77 K (**e**) for the Ni–MoS_2 FET. The nonlinear output characteristics at low temperatures suggest the existence of a Schottky barrier at the Ni–MoS_2 junction. **f**, Schottky barrier (Φ_{SB}) extracted by equation (3) as a function of the gate voltage for the Ni–MoS_2 FET. Φ_{SB} is around 100 meV at the flatband voltage (the elbow of the curve)[2]. **g**, Typical I_{DS}–V_{GS} curves of the Ti–MoS_2 FET. **h**, Typical I_{DS}–V_{DS} curves at 77 K for the Ti–MoS_2 FET. Similar to the Ni–MoS_2 device, the Ti–MoS_2 FET exhibits both drain–current suppression and obviously nonlinear output characteristics at low temperatures, owing to the presence of a Schottky barrier at the Ti–MoS_2 interface. **i**, Extracted Φ_{SB} for the Ti–MoS_2 FET as a function of the gate voltage, which is around 150 meV at the flatband voltage.



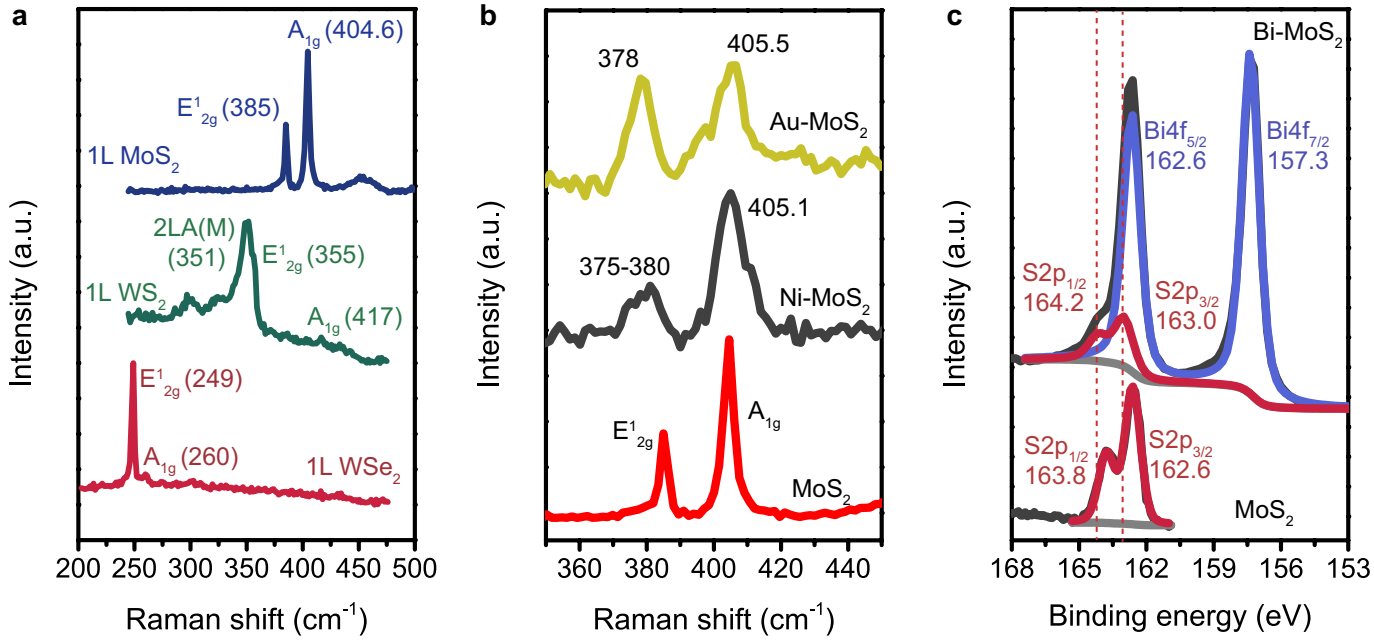
Extended Data Fig. 2 | Arrhenius plots and extraction of contact resistance. **a**, **b**, Arrhenius plots of Ni-contacted (**a**) and Bi-contacted (**b**) monolayer (1L) MoS_2 FETs. The two transistors yield opposite slopes derived from equation (3), reflecting different metal–semiconductor junction configurations. The good agreement between the data extracted from the Ni–MoS_2 FET and the thermionic emission model suggests that there is thermally activated electronic transport at an energy barrier, that is, a Schottky barrier at the Ni–MoS_2 interface. By contrast, the deviation from the thermionic emission model and nearly saturated slopes at low temperatures observed in the Bi–MoS_2 FET indicate the disappearance of an energy barrier for electron injection. The light blue curve represents the off state of the Bi–MoS_2 FET biased at a negative gate voltage of −60 V. The device at this condition shows a negative slope in the Arrhenius plot and the effective barrier height is extracted to be ~130 meV. This barrier originates from the energy difference between the Fermi level of the degenerate MoS_2 underneath Bi and the conduction band minimum of the depleted MoS_2 channel. **c**, Transfer characteristics, I_{DS}–V_{GS}, of Bi-contacted monolayer MoS_2 FETs on 100-nm-thick SiN_x with various channel lengths (L_{CH}) at a V_{DS} of 0.5 V for the TLM study. **d**, Plots of total device resistance R_{TOT} (normalized by width) versus L_{CH} for the Bi–MoS_2 FETs at various carrier densities, from which the total contact resistance ($2R_C$) can be extracted from the y-axis intercepts. Symbols are experimental data and lines are linear fits in **a** and **d**.



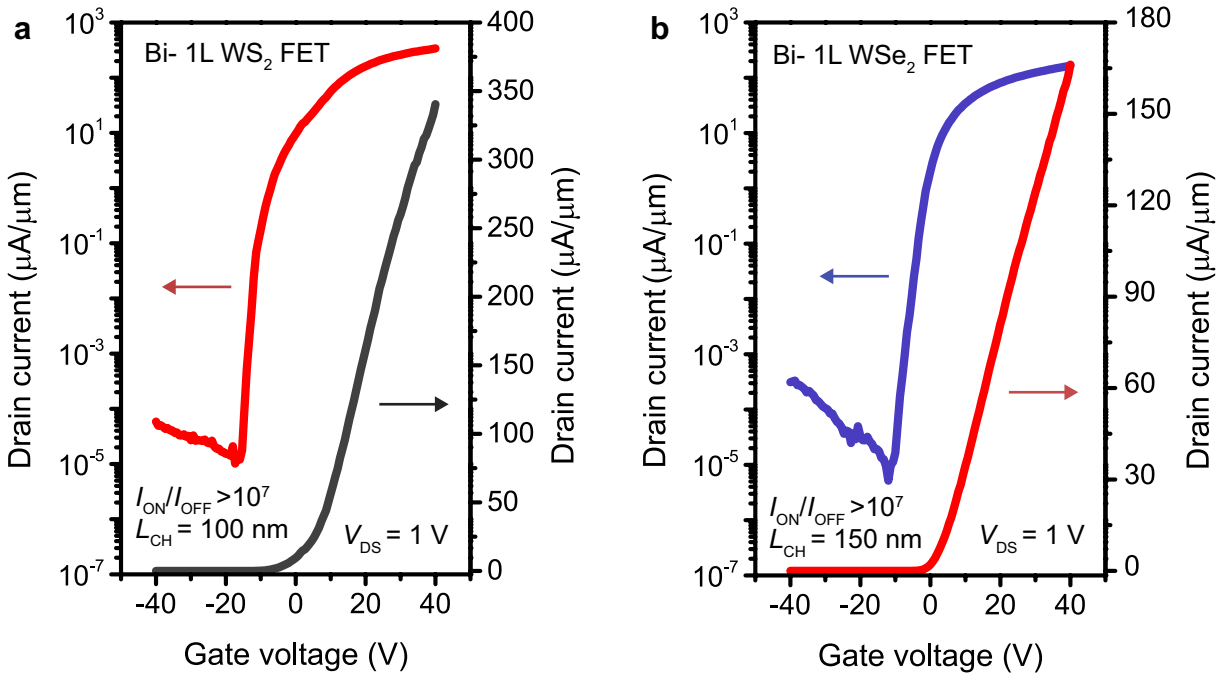
Extended Data Fig. 3 | SAED patterns of the freestanding Au/Bi/monolayer–MoS_2 and Au/Bi/amorphous carbon. a, **e**, Schematics of the Au/Bi layer deposited directly on the monolayer (1L) MoS_2 (**a**) and amorphous carbon (a-carbon; **e**) in the TEM grid. **b**–**d**, SAED patterns of Au/Bi/1L–MoS_2 at three different locations. The [0001] zone axis of Bi is always observed in parallel to the electron beam throughout the whole sample. The diffraction spots of MoS_2 at 3.6 nm^{-1} can be clearly identified. The in-plane rotations of MoS_2 with respect to the Bi (0001) plane are 30° (**b**), 4° (to the nearest Bi diffraction spots; **c**), and 8° (**d**). For most of the areas, Bi demonstrates homogeneous orientation, as shown in **b** and **d**, but polycrystalline areas can also be found, as shown in **c**. The selected-area aperture is 1 µm. **f**–**h**, The diffraction ring located at 3.0 nm^{-1} is identified to be from Bi_2O_3 polycrystal, as confirmed from the atomic structure of Bi_2O_3 viewing at zone axes [110] (**g**), and its simulated diffraction pattern (**h**), demonstrating the diffraction pattern at 3.0 nm^{-1}.



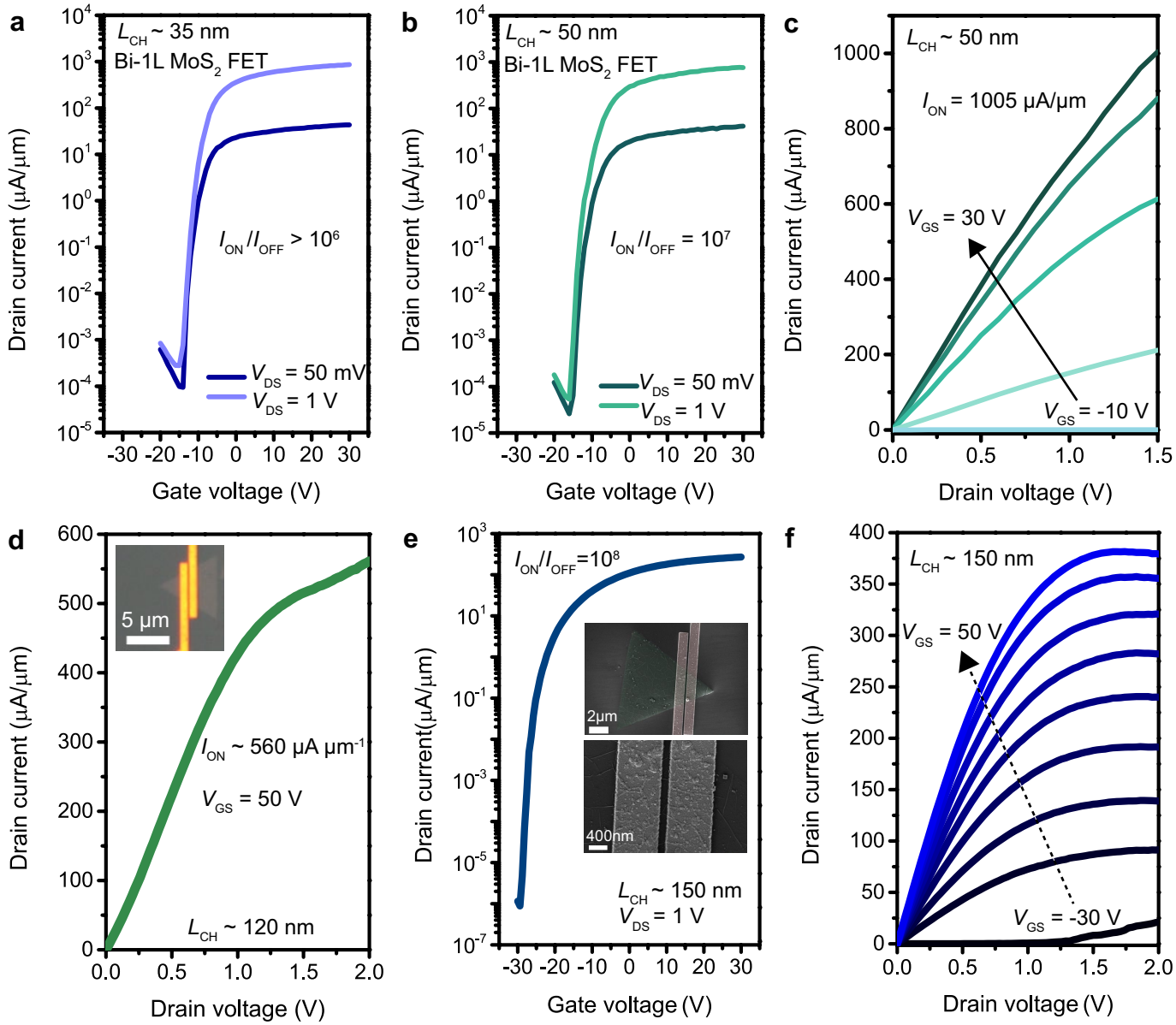
Extended Data Fig. 4 | DFT results for Sb–MoS_2, Bi–MoS_2 with sulfur vacancy and Bi–WS_2. a, PLDOS of MoS_2 before (upper) and after (lower) contact with Sb. The valence band (VB) is shaded in light blue and conduction band (CB) in light red. The Fermi level (E_F) is shifted from the valence band maximum inside the gap (before Bi contact) into the conduction band (after Bi contact). **b**, **c**, LDOS of MoS_2 with a sulfur vacancy (**b**) and WS_2 (**c**) when in contact with Bi. The Fermi level is pinned at the sulfur vacancy defect state inside the bandgap. This implies that a high-quality TMD crystal with a low defect density is critical to form ohmic contact to Bi. The result of LDOS of WS_2 in contact with Bi, predicting that ohmic contact can also be formed at the Bi–WS_2 interface owing to gap-state saturation.



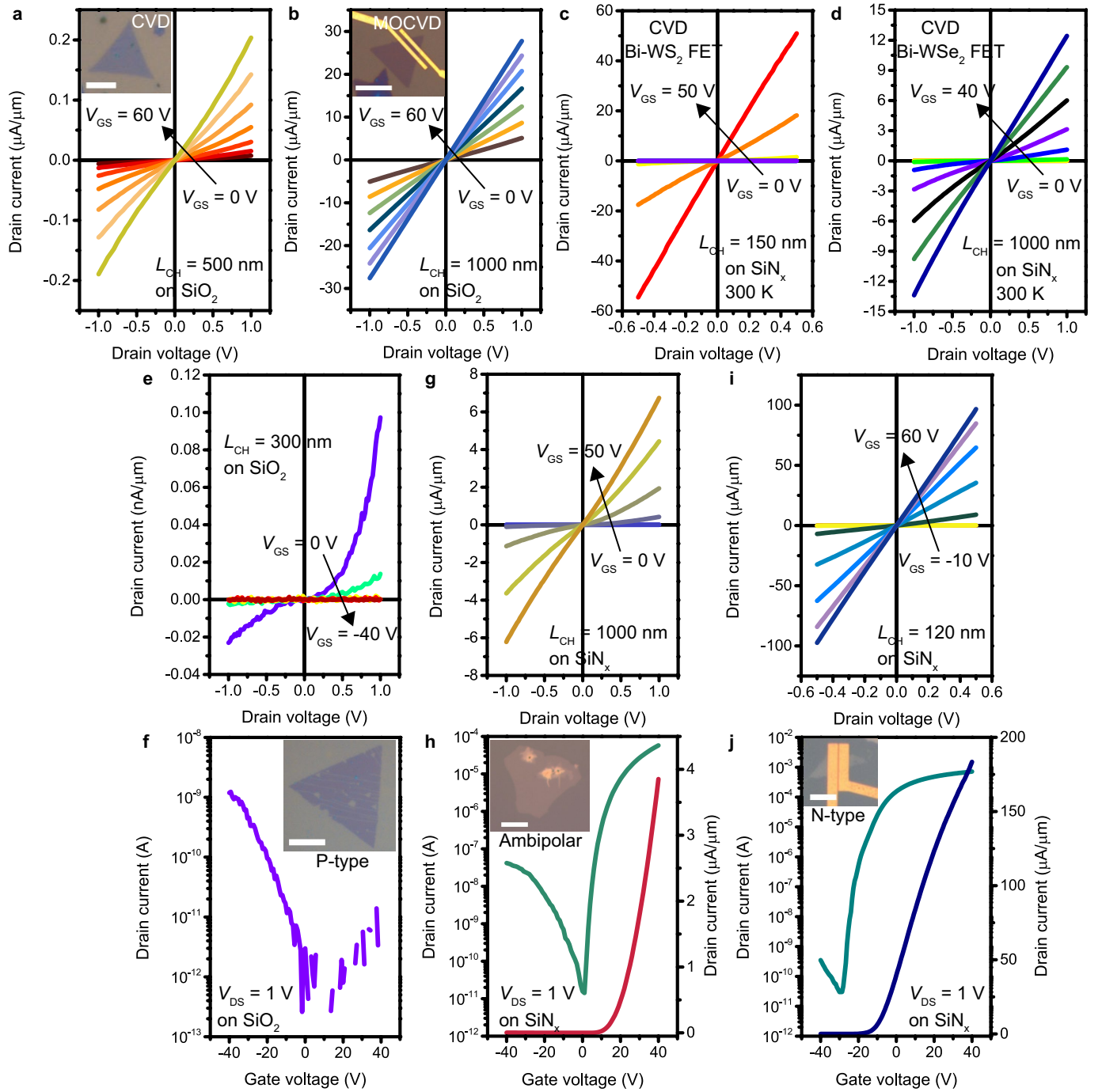
Extended Data Fig. 5 | Characterization of transition metal dichalcogenide monolayers. a, Raman characterization of MOCVD-grown monolayer MoS_2 (blue) and mechanically exfoliated WS_2 (green) and WSe_2 (red) monolayers for device fabrication. **b**, Raman characterization of Ni–MoS_2 and Au–MoS_2 interfaces. Samples are prepared using the mechanically tape-assisted exfoliation. No substantial shifts in A_{1g} are observed for Ni and Au contacts. The shift in E^1_{2g} is prevalently observed in the metal–MoS_2 system, probably originating from the strain induced at the metal–MoS_2 boundary. **c**, Deconvolution of the XPS spectra of S 2*p* and Bi 4*f* for pristine monolayer MoS_2 and Bi-contacted MoS_2. The blueshifted core-level binding energies for the Bi-contacted MoS_2 indicate the upward shift of its Fermi level induced by the Bi contact, which is in good agreement with the DFT calculation and the Raman spectroscopy analysis. Moreover, the absence of characteristic peaks for Bi_2O_3 suggest that the Bi contact is free of oxidation when in contact with MoS_2, which is consistent with the TEM results (Fig. 3b and Extended Data Fig. 3).



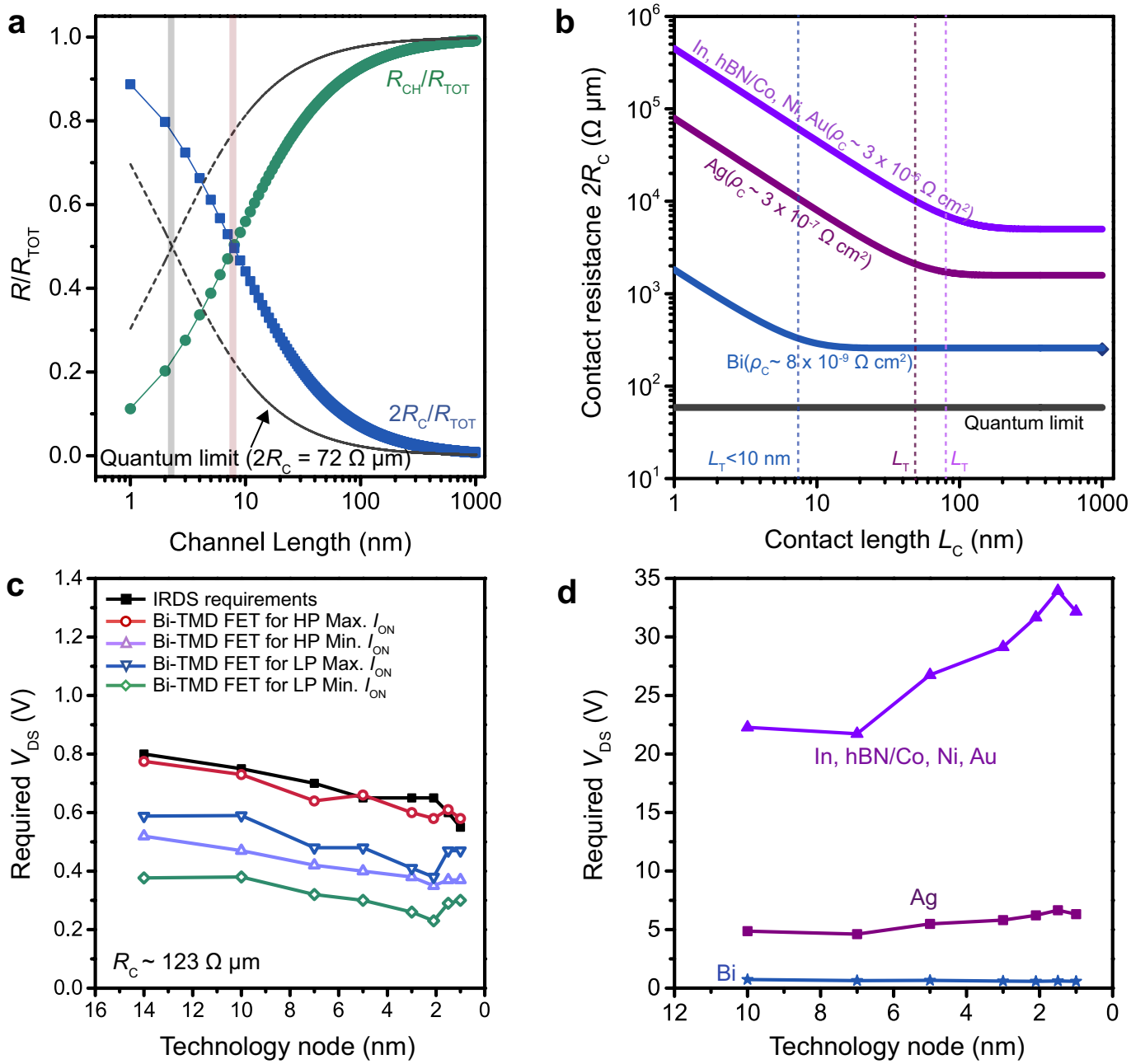
Extended Data Fig. 6 | Transfer characteristics of monolayer WS_2 and WSe_2 FETs with Bi contacts. a, **b**, Typical transfer characteristics of Bi–WS_2 (**a**) and Bi–WSe_2 (**b**) FETs on 100-nm SiN_x at room temperature. Both transistors exhibit n-type conduction with a high I_{ON}/I_{OFF} ratio of $>10^7$.



Extended Data Fig. 7 | Monolayer MoS_2 transistors with very high I_{ON}. **a**, Transfer characteristics of a 35-nm L_{CH} Bi–MoS_2 FET. **b**, **c**, Transfer and output characteristics of a 50-nm L_{CH} Bi–MoS_2 FET. **d**, Output characteristics of a 120-nm L_{CH} Bi–MoS_2 FET. The excellent current-delivery capacities represent, to our knowledge, new records for monolayer MoS_2 at these device dimensions, outperform thicker TMD devices, and are comparable to three-dimensional semiconductor devices such as 90-nm node-strained Si and AlGaAs/InGaAs HEMT transistors with similar channel lengths[39,40,46]. Note that the required drain voltage for the ohmic Bi–monolayer MoS_2 FET to achieve a high I_{ON} is relatively small compared to previously reported high-performance TMD transistors (that is, typically V_{DS} > 2 V with a thicker channel thickness)[18,22,30–34,42–45]. Inset, optical microscopic image of the device. **e**, Semi-logarithmic plot of the transfer characteristic of a different Bi–MoS_2 FET showing an excellent I_{ON}/I_{OFF} ratio of 10^8. Insets, SEM image of a representative 150-nm L_{CH} Bi-contacted monolayer MoS_2 FET on 100-nm-thick SiN_x and its channel region. **f**, Output characteristics of the same Bi–MoS_2 transistor as in **e**. The drain current saturates at a V_{DS} of -1.5 V and scales linearly with the gate voltage, which suggests that the electrons travelling in the monolayer MoS_2 channel reach its saturation velocity. The gate dielectrics of devices presented in this figure are 100-nm SiN_x.



Extended Data Fig. 8 | Effects of TMD quality on the output characteristics. **a**, **b**, Sample-quality-dependent contact performance for the case of monolayer MoS_2. The room-temperature output characteristics of the Bi–MoS_2 transistors fabricated with a CVD-grown defective MoS_2 monolayer (**a**) and MOCVD-grown MoS_2 monolayer (**b**). Inset to **a**, optical image of a typical low-quality MoS_2 crystal with a non-clean surface and curved edges; scale bar, 5 μm. Inset to **b**, optical image of a typical high-quality MoS_2 crystal with a clean surface; scale bar, 10 μm. **c**, **d**, Output characteristics of Bi-contact transistors fabricated with fresh CVD-grown monolayer WS_2 (**c**) and monolayer WSe_2 (**d**) FETs, showing that the proposed gap-state-saturation-induced ohmic contact can also be formed on high-quality WS_2 and WSe_2 CVD samples. **e**, **f**, Room-temperature output characteristics (**e**) and transfer curves (**f**) of the Bi–WSe_2 transistors fabricated with an aged CVD-grown WSe_2 monolayer (low quality). Scale bar, 10 μm. **g**, **h**, Room-temperature output characteristics (**g**) and transfer curves (**h**) of the Bi–WSe_2 transistors fabricated with a fresh CVD-grown WSe_2 monolayer (medium quality). Scale bar, 10 μm. **i**, **j**, Room-temperature output characteristics (**i**) and transfer curves (**j**) of the Bi–WSe_2 transistors fabricated with a mechanically exfoliated WSe_2 monolayer (high quality). Scale bar, 5 μm. The results show a clear evolution from p-type conduction to enhanced n-type conduction with the sample quality improvement. These variations could be attributed to the gap-state pinning effect induced by the chalcogen vacancies (Extended Data Fig. 4b). Insets to **f**, **h** and **j** are the optical images of a typical low-quality CVD WSe_2 crystal with an obviously defective surface (**f**), a medium-quality CVD WSe_2 with an irregular crystal shape (**h**), and a high-quality, freshly exfoliated WSe_2 with a clean surface (**j**).



Extended Data Fig. 9 | Performance projection of Bi–monolayer TMD technology. a, Fraction of channel resistance (R_{CH}, green line) and total contact resistance ($2R_C$, blue line) with respect to the total device resistance ($R_{TOT} = R_{CH} + 2R_C$) in Bi–MoS_2 FETs as a function of the channel length (L_{CH}) at room temperature based on the device and material parameters extracted from Fig. 2c. The dashed lines show the quantum limit, representing the minimum R_C that can be achieved in a transistor. The quantum limit R_C is $\pi h/(4q^2k_F) \approx 0.036(n_{2D})^{-0.5}$ kΩ μm, which is determined by the quantum resistance ($h/2q^2 \approx 12.9$ kΩ) and the number of conducting modes per channel width (k_F/π), which is related to the 2D sheet carrier density (n_{2D}, in units of 10^{13} cm^{-2})[2]. **b**, Projection of $2R_C$ as a function of the contact length (L_C) in monolayer TMD transistors based on the transmission line model with various metal contacts at room temperature. The vertical dashed line represents the current transfer length (L_T) for each metal contact. The results are calculated based on the data extracted from previously reported TLM results[13,47]. As can be seen, R_C increases as L_C becomes comparable to L_T, owing to the current-crowding effect (equation (4))[18]. Note that In, hexagonal boron nitride (hBN)/Co, Ni and high-vacuum Au contacts to monolayer MoS_2 exhibit similar values of R_C (~3–6 kΩ μm) and ρ_C (~10^{-6}–10^{-5} Ω cm²)[13,14,18,50]. **c**, Required minimum V_{DS} for Bi-contacted monolayer TMD transistors to work in the velocity saturation regime using our best R_C of 123 Ω μm and a theoretical F_C of 1.15×10^5 V cm^{-1}. The V_{DD} required by IRDS is also plotted. **d**, The required V_{DS} to bias monolayer MoS_2 transistors in the velocity-saturation regime for different contact technologies.

Extended Data Table 1 | Key performance metrics of representative devices

Channel	Synthesis method	Contact	Gate oxide	L (nm)	$\mu_{FE,2t}$ ($cm^2/V/s$)	I_{ON} ($\mu A/\mu m$) /V_{DS} (V)	I_{ON}/I_{OFF}
1L MoS_2	MOCVD	Bi	100 nm SiN_x	120	21	560/1.5	10^7
				150	21	378/1.5	10^8
				500	17	150/1.5	10^7
			300 nm SiO_2	1000	30	28/1	10^8
		Ni		1000	3	2/1	10^6
		Ti		1000	0.03	0.02	10^4
	CVD, high quality	Bi	100 nm SiN_x	35	22	1135/1.5	10^6
				50	25	1005/1.5	10^7
				100	16	434/1.5	10^7
				200	15	339/1.5	10^7
	CVD, low quality		300 nm SiO_2	500	0.2	0.2/1	10^3
1L WS_2	exfoliated, high quality	Bi	100 nm SiN_x	120	19	350/1.5	10^7
	CVD, high quality			150	21	100/1	10^7
1L WSe_2	exfoliated, high quality	Bi	100 nm SiN_x	120	12	321/1.5	10^8
	CVD, high quality			1000	17	14/1	10^8
	CVD, medium quality			1000	4	3.9/1	10^6
	CVD, low quality (aged)		300 nm SiO_2	300	0.02	0.06/1	10^4

The field-effect mobility, $\mu_{FE,2t}$, is extracted by two-terminal configurations in which the effect of contact resistance is included (see Methods for details). 1L, monolayer.