nature portfolio

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

Supplementary information

Ultralow contact resistance between semimetal and monolayer semiconductors

In the format provided by the authors and unedited

nature research

Peer Review File

Manuscript Title: Ultralow contact resistance between semimetal and monolayer semiconductors

Editorial Notes:

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

Paper reports high drive current in a back-gated MoS2 FET through contact resistance reduction. The work ascribes the low-contact resistance to the use of a semi-metal Bi as the contact metal to TMD channel. The data reported in the paper show high currents and an enhanced linearity in the electrical characteristics.
I have the following questions --

1. The authors have proposed that the reason for the unpinning is the semi-metallic nature of an evaporated Bismuth. Outside of its use as a contact metal in MoS2 transistor, can you share if any other electrical testing was done to confirm the nature of the Bismuth? What is its resistivity? How does it respond to a gate field?
2. Can an evaporated Bi metal layer be represented by a band structure? Would the small grain size complicate the picture?
3. How does one ensure that the fermi-level of the semi-metal aligns with conduction band edge of the n-type semiconductor? Would one still be able to make a zero-barrier contact if this is not the case?
4. Arrhenius plots to extract R_contact -- The authors show "normal" Arrhenius behavior with expected gate voltage dependence for Ni contacts. Bismuth, however, shows an opposite slope at high temperatures. This anomalous behavior is attributed to channel resistance dependence on mobility.
a. When the nickel contact is made more "transparent" at higher VG does the channel resistance dependence on temperature show up?
b. If the Bi-MoS2 device is biased in its off-state, the mobility of the channel should cease to matter. I would expect to see a "normal" Arrhenius plot which barrier height determined by the top of barrier in the channel. I request the authors to add this to extended data Fig 3.
5. Can the authors show what the barrier height of Bi contact to WS2 and WSe2 is? Does it follow expected trends from electron affinity of the channel?
6. What is the role of SiN as the gate oxide for the study? What is the channel width used for the MoS2 1L device with Bi contacts?

Referee #2 (Remarks to the Author):

The authors demonstrated a record-low contact resistance (RC) of 123 Ω μm, and a record-high on-state current density (ION) of 802 μA μm-1 on monolayer MoS2 by achieving zero Schottky barrier height. They suggested and proved a new strategy for ohmic contact formation by suppressing the CB component of MIGS using semimetal-semiconductor contacts to avoid the GSP. The results were quite impressive and meaningful for next generation transistor technologies beyond Si. The experiment and simulation in manuscript logically described. However, in order to more clarify the suggested concept, the manuscript has following questions and issues that must be fully addressed.

1. Ti-MoS2 contact showed a different performance and barrier height compared to Bi-MoS2, despite having similar low work functions. I wonder if the experimental difference between Bi and Ti is due to surface deformation such as Ti and MoS2 bond formation as previously reported. If MoS2 formed

interface without damage, it would be better to add the simulation data to support the role of semi-metal more clearly as shown in Fig. 3e.
2. Figure 1 shows the main concept of this paper. However, since it is still before the concept is understood as a result, it must be clearly presented on the key points without any confusion.
i) Band diagram of semi-metal in Figure 1e should be modified as like Figure 1b. It would be better to understand intuitively by the schematic.
ii) Where is the origin of TB in Figure 1d and 1e? Does it mean vdW gap?
iii) The reviewer proposes to change the GSS to clearly show the phenomenon between Bi and MoS2 in Figure 1f.
3.i) How to control nD in Figure 2c?
ii) In Fig 2f, it is necessary to clearly explain why the positive slop exist in 200-300K.
iii) Theoretically, mobility decreases with temperature because more carriers are present and these carriers are more energetic at higher temperatures. Each of these facts results in an increased number of collisions and mobility decreases. Why does mobility behavior in Ti-MoS2 FET have the opposite phenomenon?
4.i) Based on Figure 3b and Extended data Figure. 1f, what crystallinity does Bi on defective CVD MoS2 have? Is it like Bi on amorphous carbon? Or does it have a rhombohedral structure like on intrinsic MOCVD MoS2?
ii) Based on Extended data Figure 8a and 8b, drain current is different each other. Here, defective CVD MoS2 showed significantly low performance. It is necessary to present film analysis data on how CVD MoS2 and MOCVD MoS2 are different. The authors need to explain how the meaning of defective is distinguished.
iii) The authors should show the cross-section TEM images of Bi on two types of MoS2. As mentioned in the previous studies, de-pinning of MoS2 begins with the no-bonding and no-damage between Metal and MoS2.
iv) The authors explained that there is charge transfer between Bi and MoS2 in Figure 3h. Is there any nature of BixSy bonding due to charge transfer?
5. The author showed diverse FET results of Bi-MoS2 according to various channel lengths in each Figure. They then compared different characteristics for each device. In terms of contact resistance, the world best record is important, but it makes sense to systematically show the mobility, contact resistance, Ion, and Ioff associated with each other according to length changes. I recommend summarizing FET characteristics according to the channel length scale.

Referee #3 (Remarks to the Author):

The paper addresses the very important topic of lower contact resistance to transistors where the channel is a 2D transition metal dichalcogenide. This class of materials has been put forth as having excellent properties to extend transistor gate length scaling beyond what can be implement with Si transistors. Despite a long slew of articles in high impact journals, relatively little has been demonstrated experimentally in terms of device performance, i.e. I_ON. This paper is trying to tackle this challenge by improving contact resistance to 2D channels. The solution investigated here is very simple, using Bi as contact material to the 2D channel and trying to prove that the contact thus made is ohmic.

We expect a paper on this topic and proving beyond doubt would have a large impact on the semiconductor industry and thus be a cornerstone for technology for years to come.

Before going into technical details, a note about readability: the paper would benefit from an extended format as about half of the figures described in the main text and key figure for the paper, are now relegated to Extended Data. While the abundance of data is needed to support the claims of the paper, the continuous back and forth between the data in the main body and the extended data makes for a cumbersome read.

Gauging the full achievement of the paper is difficult because of inconsistent data reporting plotting across figures. For example figure 2 in main body of the paper shows Id-Vg data at Vds= 1V. Data in

figure 4d (on current performance) seems to be reported at Vds=1.5V and data in extended figure 3c is plotted at Vds=0.5V. Very difficult to follow and compare. We propose keep one VDS throughout the paper 1V and include extended data at VDS=50mV.

Fig 2, panel a. Comparison of transfer characteristic for MoS2 with Bi, Ni or Ti contact. Current levels for Ni and Ti are lower than literature elsewhere (for example papers from Pop group at Stanford) which report ~ 10-20uA/um for similar device conditions with Au contact. This makes the comparison here look very good for Bi, but not clear if this stands when compared with best data out there.
Figure 2 panel c: contact resistance extraction is performed in a back-gated configuration at very high doping levels. Relevant data for transistor performance is normally done without overlap between gate and source/ drain. Please include data or extrapolation at zero back-gate voltage, or data from devices when the contacts are not gates. Otherwise, comparison with Si devices and the IRDS target is irrelevant.
The authors use TLM as the method to extract contact resistance. Several publications on 2D materials and SOI have proposed that the method has high inaccuracy for these types of thin channels. In the case of graphene, several report zero or negative contact resistance. This has been ascribed to this inaccuracy. Please compare TLM extracted contact resistance with that from 4-point probe measurements.
Please show series for Id-Vg data at different channel length at VDS=1V. Data from figure 2a is no included in the 2c plot. Why not? Can you please include?
The paper compares contact resistance with IRDS targets for 2024. This is irrelevant for the technological target. They should be derived from performance in a loaded ring oscillator from implications on delay considering the target drive current.

Probably most exciting part of the paper is now relegated to Fig 10 in extended data. Any kind of data from scaled devices especially showing channels scaled to 35nm should be prime and center in the paper itself. While Id-Vd data is shown for 35nm channel, Id-VG data is shown for 150nm channel length. To prove ohmic contacts, please include data from 35nm channel without Off current degradation, so include Id-Vg data for Lch=35 nm.

In the current form, I do not recommend the paper for publication in Nature. Addressing data consistency as described below and including crucial data Id-Vg at Lch<50nm could make it into the quality and value of reporting we expect from Nature.

Author Rebuttals to Initial Comments:

*The responses are shown in blue fonts.

Referee #1:

Paper reports high drive current in a back-gated MoS_2 FET through contact resistance reduction. The work ascribes the low-contact resistance to the use of a semi-metal Bi as the contact metal to TMD channel. The data reported in the paper show high currents and an enhanced linearity in the electrical characteristics.

I have the following questions --

1. The authors have proposed that the reason for the unpinning is the semi-metallic nature of an evaporated Bismuth. Outside of its use as a contact metal in MoS_2 transistor, can you share if any other electrical testing was done to confirm the nature of the Bismuth? What is its resistivity? How does it respond to a gate field?

Answer: To characterize the electrical properties of the Bismuth (Bi) contacts, 20 nm of Bi thin film was evaporated on monolayer MOCVD MoS_2 with 100-nm SiN_x and heavily doped silicon as the dielectric and

back-gate, respectively (inset of Figure R1a). The whole device architecture is the same as the Bi contacts used in the presented transistors in the manuscript.

As can be seen in Figure R1a, the Bi thin film (or the Bi contacts in this work) clearly shows no gate dependence over the entire range of gate voltages (-40 V ~ 40 V), confirming its metallic nature. The linearity of the output characteristic shown in Figure R1b again suggests the metallic nature of the Bi contact itself. The sheet resistance (R_{SH}) is estimated to be 0.46 kΩ/square, which is two orders of magnitude smaller than that of monolayer semiconducting MoS_2 (for example, R_{SH} ~ 17 kΩ/square for the our MoS_2 channel with a carrier density of 1.5 x 10^{13} cm^2). Therefore, the semi metallic Bi contacts can act well as electrical contacts to 2D semiconductors, as demonstrated in the manuscript. The electrical resistivity of the Bi thin film is estimated to be 9 x 10^{-6} Ω m.

In the revised manuscript, we have added the following sentence in the "Device fabrication and characterization" section of Methods: "*The electrical resistivity of the evaporated bismuth film is measured to be 9×10^{-6} Ω·m.* "

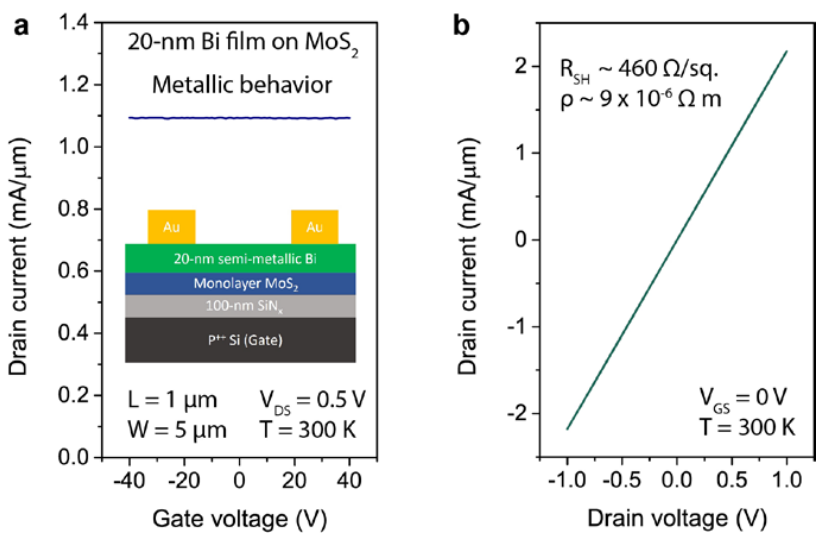

Figure R1. Electrical properties of a 20-nm Bi film evaporated on monolayer MoS_2.

2. Can an evaporated Bi metal layer be represented by a band structure? Would the small grain size complicate the picture?

Answer: The TEM SAED image (with aperture size of 1 μm) shows that the crystal orientation is highly aligned and the diffraction pattern of Bi can be clearly visualized, which is a strong evidence that Bi can be described as crystals well depicted by atomic models and first-principles calculations. Therefore, the grain boundaries, which is only a small fraction of the totally area, should not be a dominating factor in alternating contact properties.

3. How does one ensure that the fermi-level of the semi-metal aligns with conduction band edge of the n-type semiconductor? Would one still be able to make a zero-barrier contact if this is not the case?

Answer: From first-principles calculation, we have concluded that the following conditions need be met for an ohmic contact to be realized:

a. The electron hybridization between metal and semiconductor needs to be weak so the metal-induced gap states are minimized. Bi semimetal has two characteristics to ensure this: (1) The density of states (DOS) of semimetal around Fermi level is zero, so metal-induced gap states (MIGS) is minimal around the Fermi level. (2) The layered structure of Bi semimetal ensures that the electron bonds are completely saturated at the surface, excluding the possibility of having dangling bonds which may induce significant metal-induced gap state. This also requires the semiconductor to be free of dangling bonds, where MoS_2 fortunately is.

b. The work function of the semimetal (or metal) and the electron affinity of the semiconductor *before* contact is important, because if the Fermi level of (semi)metal is not aligned with the bands (either conduction or valence bands) of semiconductor in the first place, no ohmic contact can be formed. For example, it has been experimentally shown that graphene, which is also a semimetal, does not have as good contact with MoS_2, due to the fact that graphene itself has a work function of around 4.7 eV, larger than the electron affinity of MoS_2. We have also predicted in the main text that arsenic doesn't have a good contact with MoS_2, for the same reason. More details can be found in Fig. 3g.

4. Arrhenius plots to extract R_contact -- The authors show "normal" Arrhenius behavior with expected gate voltage dependence for Ni contacts. Bismuth, however, shows an opposite slope at high temperatures. This anomalous behavior is attributed to channel resistance dependence on mobility.

a. When the nickel contact is made more "transparent" at higher V_G does the channel resistance dependence on temperature show up?

Answer: Figure R1b shows the Arrhenius plots of the Ni-MoS_2 device at a higher gate voltage presented in the previous Extended Data Fig. 3. Indeed, when the Schottky barrier of Ni/MoS_2 interface becomes more transparent due to a higher electron doping level at the interface, the device is dominated more by the channel resistance and the similar positive slope also shows up. For Ni contacts, this positive slope only happens at a high gate voltage (60 V) and high carrier concentration in the channel (~4.3×10^{12} cm^{-2}), while a Bi-MoS_2 transistor shows such behavior with a much lower gate voltage and lower carrier concentration in the channel (~10^{11} cm^{-2}), as shown in Fig. 2f and Extended Data Fig. 2b in the revised manuscript.

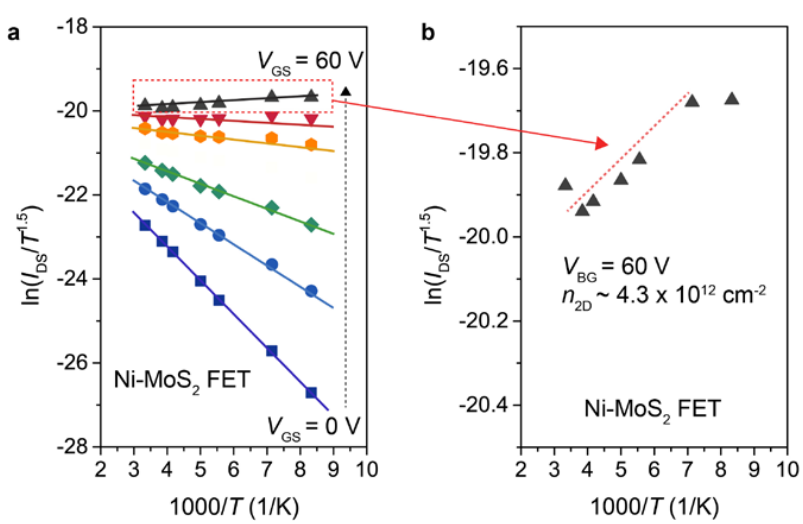

Figure R2. a. Arrhenius plot of a Ni-MoS_2 FET with different gate voltages (same as previous Extended Data Fig. 3). **b.** Zoom-in plot of **a** focusing on a high gate voltage of 60 V.

b. If the Bi-MoS_2 device is biased in its off-state, the mobility of the channel should cease to matter. I would expect to see a "normal" Arrhenius plot which barrier height determined by the top of barrier in the channel. I request the authors to add this to extended data Fig 3.

Answer: Thank you for pointing this out. We have plotted the Arrhenius plot of the Bi-MoS_2 device biased at a negative gate voltage of -60 V so that the device is in its off-state (the threshold voltage V_T is around 0 V). As can be seen in Figure R3, the device at this condition shows a negative slope in the Arrhenius plot and the effective barrier height is extracted to be ~ 130 meV. As the reviewer suggested, this barrier originates from the energy difference between the Fermi level of the degenerate MoS_2 underneath Bi and the CBM of the depleted MoS_2 channel. We have added this plot into Extended Data Fig.2b (light blue curve) in the revised manuscript.

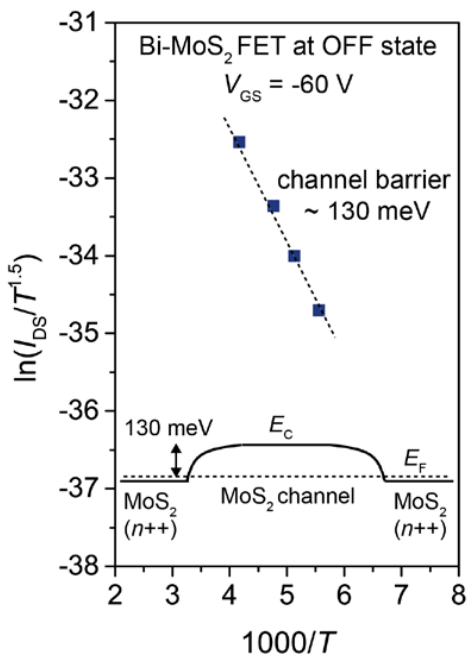

Figure R3. Arrhenius plot of a Bi-MoS_2 device at its OFF state. The device was fabricated on a 300-nm thick SiO_2 /Si substrate as the back gate. The data were extracted at V_{GS} = -60 V and V_{DS} = 1 V.

5. Can the authors show what the barrier height of Bi contact to WS_2 and WSe_2 is? Does it follow expected trends from electron affinity of the channel?

Answer: In this work we have done a systematic study on MoS_2, we have not measured the barrier heights for WS_2 or WSe_2 yet, which is in our plan for the investigation in the next step. However, following the answer in question 3 above, the increasing trend of R_C (extracted from a virtual source compact model, see Figure R4) between Bi-contacted MoS_2, WS_2, and WSe_2 follows the general trend of decreasing electron affinity of these TMDs [Y. Liu et al, *Sci. Adv.* 2, e1600069 (2016), ref. 15], which in turn implies the possibility of a small increase in Schottky barriers when Bi and these TMDs are in contact. On the other hand, the same trend can also be obtained from DFT calculation where the Fermi level in Bi-WS_2 is lower than Bi-MoS_2, although still above the CBM, as can be found in Extended Fig. 4c and Fig. 3e. This matches with our observation in Figure R4 and Fig. 4d.

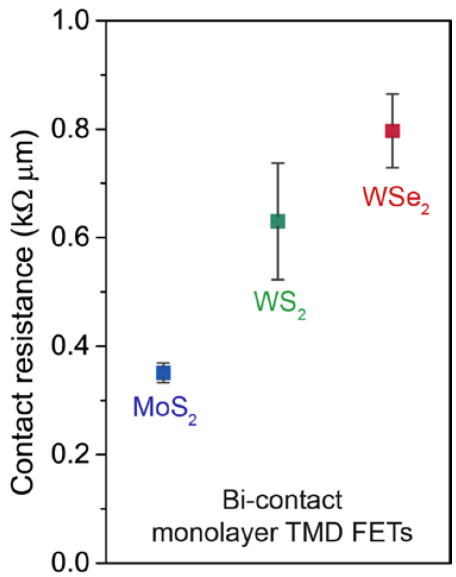

Figure R4. Contact resistance for three types of Bi-contact monolayer TMD devices extracting from device modeling.

6. What is the role of SiN as the gate oxide for the study? What is the channel width used for the MoS_2 1L device with Bi contacts?

Answer: In this work, we presented two different device structures: 1L TMD on 300 nm SiO_2 (Fig. 2) and 1L TMD on 100 nm SiN_x (Fig. 4a-c). SiO_2 is the most commonly used dielectrics for 2D-material-based device studies, so we performed the temperature-dependent measurements and the comparison study for Bi, Ni and Ti contacts on SiO_2, to make these results more consistent with previous studies. In both our experiment and literature (T. Liu et al. *Nat. Nanotech.* 14, 223-226 (2019).), it is observed that SiN_x is a better substrate because TMD tends to have better carrier mobility on SiN_x and the higher thermal conductivity of SiN_x (12 W/m/K for SiN_x and 1.3 W/m/K for SiO_2) can reduce the current degradation due to self-heating for high-performance transistors. We therefore selected SiN_x as the substrate to demonstrate the high-performance transistors.

It should be noted that the selection of substrates does not impact the electrical contact at the Bi-TMD interface. First, the crystallinity of the evaporated Bi on TMD, and thus the resulting interface, should not be affected in any way by the substrate. Second, since the MoS_2 underneath Bi is in a degenerate state as presented in our work, there is no Schottky barrier and no depletion region on the MoS_2 channel side of the contact. As a result, no barrier width modulation exists which normally comes from the different electrostatics determined by the dielectric constant and the thickness of the gate oxide, as is usually the case in previous work (D. Schulman et al. *Chem. Soc. Rev.* 47, 3037-3058 (2017)).

The channel width in this study is in the range of 2 to 10 µm. Since we presented the current density (drain current normalized by channel width) throughout the manuscript, we considered it unnecessary to show the channel width for each device. But based on the reviewer's question here, we have mentioned the range of channel width in our study in the "Device fabrication and characterization" section of Methods of the revised manuscript: *"The channel widths for the devices in this study are in the range of 2 to 10 µm."*

Referee #2:

The authors demonstrated a record-low contact resistance (R_C) of 123 Ω μm, and a record-high on-state current density (I_{ON}) of 802 μA μm-1 on monolayer MoS2 by achieving zero Schottky barrier height. They suggested and proved a new strategy for ohmic contact formation by suppressing the CB component of MIGS using semimetal-semiconductor contacts to avoid the GSP. The results were quite impressive and meaningful for next generation transistor technologies beyond Si. The experiment and simulation in manuscript logically described. However, in order to more clarify the suggested concept, the manuscript has following questions and issues that must be fully addressed.

1. Ti-MoS_2 contact showed a different performance and barrier height compared to Bi-MoS_2, despite having similar low work functions. I wonder if the experimental difference between Bi and Ti is due to surface deformation such as Ti and MoS_2 bond formation as previously reported. If MoS_2 formed interface without damage, it would be better to add the simulation data to support the role of semi-metal more clearly as shown in Fig. 3e.

Answer: Previous simulation between Ti and MoS_2 reported by H. Zhong et al. [*Sci. Rep.* 6, 21786 (2016), ref. 25] indeed shows that there is strong interaction (bonding) between Ti as mentioned by the reviewer – a pristine non-damaged MoS_2 leads to what they call "metallization" of MoS_2. Based on the understanding of our current work, this results in a very strong gap-state pinning (GSP) between Ti and MoS_2 contact and a large Schottky barrier. Because of the strong interaction, Ti-MoS_2 contact cannot be explained by the Schottky-Mott limit in that the barrier height of Ti-MoS_2 is not proportional to the work function of the metal. Therefore, even though Ti has similar work function as Bi, Ti-MoS_2 devices has large contact resistance.

For the question "If MoS_2 formed interface without damage, it would be better to…", we are unclear if the reviewer is referring to MoS_2-Bi or MoS_2-Ti interface, but most likely it is the former as we all agree earlier results already indicated MoS_2-Ti bond formation. As we have shown in the main text, the sample with poor quality might suffer from Gap State Pinning due to the pinning effect of defect states such as sulfur vacancy. However, we do not observe such pinning for a MOCVD MoS_2 samples, proving that the Bi deposition process is mild and free of defect creation.

2. Figure 1 shows the main concept of this paper. However, since it is still before the concept is understood as a result, it must be clearly presented on the key points without any confusion.

i) Band diagram of semi-metal in Figure 1e should be modified as like Figure 1b. It would be better to understand intuitively by the schematic.

Answer: Thank you very much for this suggestion, we would like to follow but realized Fig. 1e is plotted in real space while Fig. 1b is in reciprocal space. Overlapping the DOS (Figure 1b) with the band diagram (Figure 1e) may introduce additional confusion about the DOS in real space versus in reciprocal space, so we still keep the current layout to represent the concept.

ii) Where is the origin of TB in Figure 1d and 1e? Does it mean vdW gap?

Answer: TB means tunneling barrier. The origin of the tunneling barrier could be different for different contact technologies, such as (1) the vdW gap for non-covalently bonded metal-TMD interface (as in the case of In-MoS_2, Au-MoS_2, graphene-MoS_2 [Y. Wang et al. *Nature* 568, 70-74 (2019), ref. 13; C. D. English et al. *Nano Lett.* 16, 3824-3830 (2016), ref. 18; S. S. Chee et al. *Adv. Mater.* 31, 1804422 (2019), ref. 19], and in our case, Bi-MoS_2), (2) the small energy barriers formed at the covalently bonded metal-TMD interface (as in the case of Ti-MoS_2 [H. Zhong et al. *Sci. Rep.* 6, 21786 (2016), ref. 25]), or (3) the

tunneling barrier introduced by the metal/thin insulator/semiconductor structures (as in the case of Co-hBN-MoS_2 [X. Cui et al. *Nano Lett.* 17, 4781-4786 (2017), ref. 14]).

iii) The reviewer proposes to change the GSS to clearly show the phenomenon between Bi and MoS_2 in Figure 1f.

Answer: Thank you very much for this suggestion. We have modified accordingly in the revised manuscript, we have moved the "GSS" to the side and replaced it with "n^{++}" to clarify the degenerate state of MoS_2.

3.i) How to control n_D in Figure 2c?

Answer: n_{2D} is controlled by the back-gate voltages. The way we estimated n_{2D} is described in Method. Please see details in "Extraction of contact resistance through transfer length method (TLM)" in Methods section.

ii) In Fig 2f, it is necessary to clearly explain why the positive slop exist in 200-300K.

Answer: Because in the case of Bi-MoS_2 devices, the contact resistance (determined by the Schottky barrier) is much smaller than the channel resistance (determined by the carrier mobility), the I_{DS}-T trend in the Bi-MoS_2 device is dominated by the trend of mobility-*T* (**Fig. 2e**): mobility remains constant in the low temperature range (<200K), and decreases with temperature at higher temperature (200-300K). This explains the positive slope in the Arrhenius plot (Fig. 2f). Please find this explanation on page 6, paragraph 1of the revised manuscript: "*However, this analysis becomes invalid for Bi-MoS_2 FETs (Fig. 2f and Extended Data Fig. 3b). Instead, the saturation-like regime at lower temperatures (< 200 K) suggests a zero contact barrier height for electron transport through the conduction band of MoS_2, while the positive slope in the range of 200 – 300 K can be attributed to the negative correlation between mobility and temperature.*"

iii) Theoretically, mobility decreases with temperature because more carriers are present and these carriers are more energetic at higher temperatures. Each of these facts results in an increased number of collisions and mobility decreases. Why does mobility behavior in Ti-MoS_2 FET have the opposite phenomenon?

Answer: We agree with the reviewer about the relationship between the carrier mobility and temperature. This trend should stay the same for different contact metals and it has been shown in Fig. 2e. However, the drain current, or the output resistance, is affected by both the channel resistance (thus carrier mobility) and the contact resistance (thus Schottky barrier). The opposite temperature dependence seen in the Ni and Ti devices is because the drain current for these devices are dominated by the contact resistance or the Schottky barrier: at lower temperature the thermionic emission across the Schottky barrier is suppressed, giving rise to a lower drain current. This experimental observation is not in conflict with the fact that the carrier mobility in the MoS_2 channel is enhanced at lower temperature.

4.i) Based on Figure 3b and Extended data Figure. 1f, what crystallinity does Bi on defective CVD MoS_2 have? Is it like Bi on amorphous carbon? Or does it have a rhombohedral structure like on intrinsic MOCVD MoS_2?

Answer: Bi still preserves the same type of diffraction patterns on CVD MoS_2 as in MOCVD MoS_2 as shown in our experiments (Please note that the original Extended Data Fig. 1 has been changed to Extended Data Fig. 3.)

ii) Based on Extended data Figure 8a and 8b, drain current is different each other. Here, defective CVD MoS_2 showed significantly low performance. It is necessary to present film analysis data on how CVD

MoS_2 and MOCVD MoS_2 are different. The authors need to explain how the meaning of defective is distinguished.

Answer: (Please note that the original Extended Data Figs. 8 and 9 have been merged to Extended Data Fig. 8.) Thank you for this suggestion. The sample quality can be distinguished visually by the morphology of individual MoS_2 domains as shown in the insets of Extended. Data Fig. 8a,f,h, and more rigorously from Raman spectra, as shown in Figure R5. The trends are also summarized in Table R1. Generally, the full width at half maximum (FWHM) and the shift of Raman peaks can be used to identify the crystal quality due to the activation of new vibrational modes from phonon scattering by defects [S. Mignuzz et al. *Phys. Rev. B* 91, 195411 (2015)] – The higher the FWHM is, the worse the quality is. We noted that for some CVD samples, a much broader FWHM of E_{2g} peak than that of MOCVD-grown samples can be observed. We thus consider this type of CVD sample low-quality CVD MoS_2. For those CVD samples exhibiting a similar E_{2g} FWHM as MOCVD are grouped as high-quality CVD MoS_2, as shown in Figure R5.

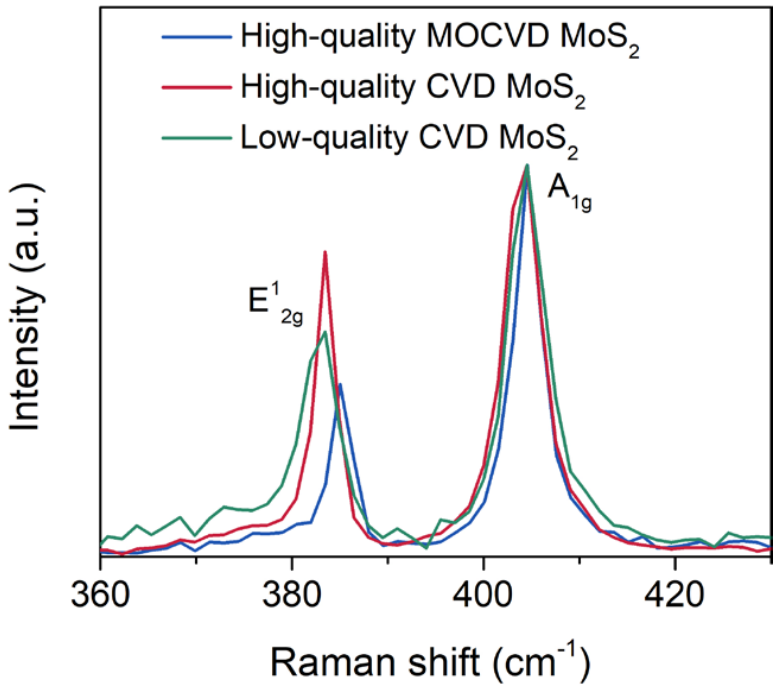

Figure R5. Comparison of Raman spectra for monolayer MoS_2 on SiO_2 prepared by CVD and MOCVD methods.

Table R1. A summary of Raman spectroscopy results and morphologies for different MoS_2 samples.

	E_{2g} (cm^{-1})	E_{2g} FWHM (cm^{-1})	A_{1g} (cm^{-1})	A_{1g} FWHM (cm^{-1})	Morphology
high-quality MOCVD MoS_2	384.9	2.9	404.5	4.8	Perfect triangles with flat edges and clean surfaces
high-quality CVD MoS_2	383.5	3.0	404.4	4.9	Perfect triangles with flat edges and clean surfaces
low-quality CVD MoS_2	383.4	5.1	404.5	5.0	Triangles with curved edges and surface contaminants

iii) The authors should show the cross-section TEM images of Bi on two types of MoS_2. As mentioned in the previous studies, de-pinning of MoS_2 begins with the no-bonding and no-damage between metal and MoS_2.

Answer: Thank you for this suggestion, and we have indeed performed cross sectional TEM for Bi-MoS_2. However, atomic cross sectional STEM is particularly hard to be performed on Bi-MoS_2 boundary, due to the low melting point of Bi. As shown in Figure R6, the crystal structure of Bi-MoS_2 is completely damaged by the ion beam during FIB. The clustering of Bi and Au particles shows a polycrystalline state of metals, different from the status demonstrated by Fig. 3a. Bi crystal is not stable under electron beam exposure in STEM either – a low dose of electron irradiation with a few frames of scan can get Bi amorphized.

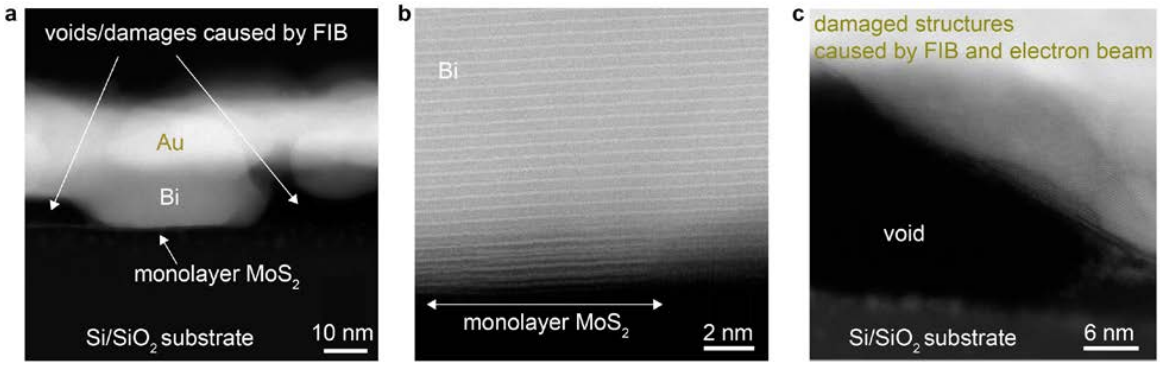

Figure R6. Scanning transmission electron microscopy (STEM) cross-sectional images of Bi/MoS_2. a. Overall view of sample structure (15-nm Au/20-nm Bi/MoS_2). **b.** STEM image of a Bi/MoS_2 structure focusing on a likely non-damaged region. **c.** STEM image of a damaged region.

We would like to point out that the ultralow contact resistance observed at Bi-TMD interfaces is not due to the de-pinning mechanism as reported previously, where no-bonding and no-damage is critical for reducing the defect-induced gap state pinning.

iv) The authors explained that there is charge transfer between Bi and MoS_2 in Figure 3h. Is there any nature of Bi_xS_y bonding due to charge transfer?

Answer: First, we would like to clarify that the charge transfer between Bi and MoS_2 is very small. The electron transfer from Bi to MoS_2, according to the Bader charge analysis performed by DFT, is at 4×10^{11} cm^{-2} which is a very small amount (shown by the differential charge analysis in Fig. 3c). Furthermore, Fig. 3h shows that MoS_2 underneath Bi is heavily doped, not because of charge transfer between Bi and MoS_2, but because of the Gap State Saturation (GSS) mechanism (i.e., due to MoS_2 contacting with Bi, the reduction in the number of valence band (VB) states in MoS_2 is more than the number of increased MIGS, thus those electrons in VB before now are filled in the MIGS and into the conduction band. Therefore, the Fermi level of MoS_2 moves into conduction band, and as a result, the free electron concentration in MoS_2 increase significantly).

Second, we do not think there is Bi_xS_y bonding formed in MoS_2. If Bi_xS_y exists at the Bi and MoS_2 interface, XPS characteristic peaks of the interfacial Bi_xS_y and the adjacent Bi layer should be both observed. However, as can be seen in Extended Data Fig. 5c, only two prominent peaks of pure Bi were observed (157.3 and 162.6 eV), confirming the lack of Bi_xS_y bonding formed at the interface. Our XPS spectra of Bi is also in good agreement with the thermo scientific XPS database for Bi (https://xpssimplified.com/elements/bismuth.php). In addition, previous studies on Bi_2S_3 have shown that Bi_xS_y exhibits XPS peaks at ~ 158.9 and 164.2 eV for Bi $4f_{7/2}$ and Bi $4f_{5/2}$, respectively [F. P. Ramanery et al. *Nanoscale Res. Lett.* 11, 187 (2016); S. V. P. Vattikuti et al. *Sci. Rep.* 8, 1-16 (2018).]. This significant difference in peak position of our Bi compared with the Bi_xS_y XPS spectra also supports the lack of Bi_xS_y formation at the contact interfaces in our Bi-MoS_2 devices.

5. The author showed diverse FET results of Bi-MoS_2 according to various channel lengths in each Figure. They then compared different characteristics for each device. In terms of contact resistance, the world best record is important, but it makes sense to systematically show the mobility, contact resistance, I_{on}, and I_{off} associated with each other according to length changes. I recommend summarizing FET characteristics according to the channel length scale.

Answer: We acknowledge the suggestion. Per the reviewer's request, we made a table to summarize the key performance metrics of different devices (including different materials, different gate oxides, and different channel lengths). The field-effect mobility μ_{FE} is extracted by 2-terminal configuration in which the effect of contact resistance is included. Since the extraction of R_c requires special device structures (TLM or 4-terminal device), we only have data for MOCVD 1L MoS_2 on 100 nm SiN_x, as shown in Fig. 2c. We expect the R_C values to be similar for different substrates. The key performance metrics of different devices in this study are summarized as Table R2 below and have been also added into the revised manuscript (Please see Extended Data Table 1).

Table R2. Key performance metrics of representative devices.

Channel	Synthesis method	Contact	Gate oxide	L (nm)	$\mu_{FE,2t}$ (cm^2/Vs)	I_{ON} (μA/μm)/V_{DS} (V)	I_{ON}/I_{OFF}
1L MoS_2	MOCVD	Bi	100 nm SiN_x	120	21	560/1.5	10^7
			100 nm SiN_x	150	21	378/1.5	10^8
			100 nm SiN_x	500	17	150/1.5	10^7
			300 nm SiO_2	1000	30	28/1	10^8
		Ni	300 nm SiO_2	1000	3	2/1	10^6
		Ti	300 nm SiO_2	1000	0.03	0.02	10^4
	CVD, high quality	Bi	100 nm SiN_x	35	22	1135/1.5	10^6
	CVD, high quality	Bi	100 nm SiN_x	50	25	1005/1.5	10^7
	CVD, high quality	Bi	100 nm SiN_x	100	16	434/1.5	10^7
	CVD, high quality	Bi	100 nm SiN_x	200	15	339/1.5	10^7
	CVD, low quality	Bi	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	Bi	100 nm SiN_x	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	Bi	100 nm SiN_x	1000	17	14/1	10^8

	CVD, medium quality	Bi	100 nm SiN_x	1000	4	3.9/1	10^6
	CVD, low quality (aged)	Bi	300 nm SiO_2	300	0.02	0.06/1	10^4

Referee #3:

The paper addresses the very important topic of lower contact resistance to transistors where the channel is a 2D transition metal dichalcogenide. This class of materials has been put forth as having excellent properties to extend transistor gate length scaling beyond what can be implement with Si transistors. Despite a long slew of articles in high impact journals, relatively little has been demonstrated experimentally in terms of device performance, i.e. I_ON. This paper is trying to tackle this challenge by improving contact resistance to 2D channels. The solution investigated here is very simple, using Bi as contact material to the 2D channel and trying to prove that the contact thus made is ohmic.

We expect a paper on this topic and proving beyond doubt would have a large impact on the semiconductor industry and thus be a cornerstone for technology for years to come.

Before going into technical details, a note about readability: the paper would benefit from an extended format as about half of the figures described in the main text and key figure for the paper, are now relegated to Extended Data. While the abundance of data is needed to support the claims of the paper, the continuous back and forth between the data in the main body and the extended data makes for a cumbersome read.

Answer: We would like to thank the referee for pointing this out, and we apologize for the hassle created by back-and-forth referral of our data. We have revised the order of all the Extended Data Figures with respect to their appearing order in the main text, so hopefully the readability is improved in this version. Due to the editorial constraint of Nature and the large amount of data we want to present in order to comprehensively prove our claims, we can only pick up the most essential and representative data in the main text. Even in the current status, we still need to shorten the main text by 1000 words. We only managed to move the original Extended Data Fig. 6e (now Extended Data Fig. 9) to the main body (Fig. 4e) in the revised manuscript. We feel deeply sorry for the compromise in the readability due to the editorial limitation.

Gauging the full achievement of the paper is difficult because of inconsistent data reporting plotting across figures. For example figure 2 in main body of the paper shows I_d-V_g data at V_{ds}= 1V. Data in figure 4d (on current performance) seems to be reported at V_{ds}=1.5V and data in extended figure 3c is plotted at V_{ds}=0.5V. Very difficult to follow and compare. We propose keep one V_{DS} throughout the paper 1V and include extended data at V_{DS}=50mV.

Answer: We are sorry for the confusion. In fact, each voltage was chosen for a reason: the interest regimes of device operation are either linear or velocity saturation depending on the channel length and V_{DS}. In the following, we provide the reason for each of them.

a. In Fig. 4d, a V_{DS} at 1.5 V was chosen because we need the devices to work at velocity saturation to demonstrate the current delivery capability and compared with literature; whereas V_{DS} at 1 V for devices with similar dimensions (L_{CH} = 100-150 nm) typically corresponds to the transition regime between linear and velocity saturation (see Fig 4a-c), making the comparison less informative.

b. a V_{DS} at 0.5 V was only used in Extended Data Fig. 2c for the TLM devices because this value allows all the devices with different channel lengths (from 100 nm to 500 nm) to work in linear regime and at the same time to extract as much current as possible so that we were able to extract the contact resistance more accurately.
c. Finally, V_{DS} = 1V was used for the rest of the manuscript (Fig.2a,d, Extended Data Fig. 1, 6, 7, 8) to consistently demonstrate the transfer characteristics of different devices. Among them, a special case is that extraction of Schottky barriers from the Arrhenius plots typically require a sufficiently large V_{DS} while maintaining the device in linear regime (see Equation 1 to 3 in Methods). A V_{DS} of 1V is widely used for such Schottky barrier extraction [Kim, Changsik, et al. *ACS Nano* 11, 1588-1596 (2017).].

To further address the reviewer's concern, we have included I_{DS}-V_{GS} curves with V_{DS}=0.05 and 1 V for the short channel devices (L_{CH}=35 nm and 50 nm) in Extended Data Fig. 7, and have added a table that summarizes the performance of different devices in our study (Table. R2).

Fig 2, panel a. Comparison of transfer characteristic for MoS_2 with Bi, Ni or Ti contact. Current levels for Ni and Ti are lower than literature elsewhere (for example papers from Pop group at Stanford) which report ~ 10-20 uA/um for similar device conditions with Au contact. This makes the comparison here look very good for Bi, but not clear if this stands when compared with best data out there.

Answer: Thank you for pointing this out. We have tabulated several representative papers including Prof. Pop group's work on Au contacts as the reviewer mentioned. We noted that the device performance reported from the representative literature employed much thinner gate dielectrics, which can be translated into a much higher carrier density, thus increasing the device drain current. In Table R3, we summarized I_{DS} in our results in comparison with Prof. Pop's work with the same level of carrier density. The drain current of our Ni-MoS_2 devices is comparable to the values reported by other representative works. Another important factor is that one of Prof. Pop's work was done on multilayer MoS_2 samples, which are much easier to form good contact as discussed in our manuscript as well as in C. D. English et al. *Nano Lett.* 16, 3824-3830 (2016), ref. 18.

We chose Ni contacts as the comparison group because it is the most widely used metal contact to MoS_2 with a good balance between performance and consistency among different literatures. In addition, Ni deposition does not require special instrumentation such as the UHV system in Prof. Pop's Au contact [C. D. English et al. *Nano Lett.* 16, 3824-3830 (2016), ref. 18], and special transfer technique of Prof. Xiangfeng Duan(UCLA)'s transferred metal technique [Y. Liu et al. *Nature* 557, 696-700 (2018), ref. 16] or Prof. James Hone(Columbia)'s hBN interfacial layer technique [X. Cui et al. *Nano Lett.* 17, 4781-4786 (2017), ref. 14]. In fact, we have taken these works into consideration when we benchmark our work (Fig. 4). Therefore, we believe Ni is a valid choice as the comparison group.

Table R3. Comparison between the key metrics across different works.

	Channel	Contact	Gate oxide	C_{ox}	L	I_{DS} @ (V_{DS}=1V, n_{2D} = 5X10^{12} cm^{-2})

This work	Monolayer MoS_2 (MOCVD)	Bi	300 nm SiO_2	11.5 nF/cm^2	1 μm	28 μA/μm
	Monolayer MoS_2 (MOCVD)	Bi	100 nm SiN_x	60 nF/cm^2	500 nm	34 μA/μm
	Monolayer MoS_2 (MOCVD)	Ni	300 nm SiO_2	11.5 nF/cm^2	1 μm	2 μA/μm
Smets et al. In IEEE *International Electron Devices Meeting* (IEDM) 23.2.1-23.2.4 (2019).	3~4 layer MoS_2 (MOCVD)	Ni	50 nm SiO_2	69 nF/cm^2	1 μm	10 μA/μm
English et al. *Nano Lett.* 16, 3824 (2016)	4.5 nm MoS_2 (exfoliated)	UHV Au	90 nm SiO_2	38.4 nF/cm^2	1 μm	20 μA/μm
Smithe et al. *2D Mater.* 4, 011009 (2017)	Monolayer MoS_2 (CVD)	UHV Au	30 nm SiO_2	115 nF/cm^2	1.2 μm	6 μA/μm
Smithe et al. *ACS Nano* 11, 8456 (2017)	Monolayer MoS_2 (CVD)	Ag	30 nm SiO_2	115 nF/cm^2	5 μm	4 μA/μm

Figure 2 panel c: contact resistance extraction is performed in a back-gated configuration at very high doping levels. Relevant data for transistor performance is normally done without overlap between gate and source/ drain. Please include data or extrapolation at zero back-gate voltage, or data from devices when the contacts are not gates. Otherwise, comparison with Si devices and the IRDS target is irrelevant.

Answer: Thank you for this suggestion. We have extracted the contact resistance to be 167 Ω μm at zero back-gate voltage (Figure R7). The contact resistance of Bi devices is almost independent of the gate-induced carrier density in the channel, which is also shown in Fig. 4g. In the revised manuscript, we added this data point to Fig. 4g.

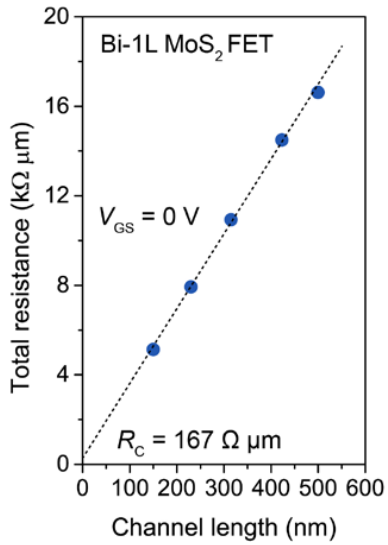

Figure R7. Contact resistance extraction for Bi-monolayer MoS_2 devices at zero back-gate voltage.

The authors use TLM as the method to extract contact resistance. Several publications on 2D materials and SOI have proposed that the method has high inaccuracy for these types of thin channels. In the case of graphene, several report zero or negative contact resistance. This has been ascribed to this inaccuracy. Please compare TLM extracted contact resistance with that from 4-point probe measurements.

Answer: Thank you for this suggestion. We used TLM for the R_C extraction of Bi-MoS_2 devices because it has been considered a more accurate model than 4-probe measurement [C. D. English et al. *Nano Lett.* 16, 3824-3830 (2016), ref. 18]. It has been found both from this literature and our experimental results that the 4-probe measurement could underestimate R_C by more than a factor of 10 due to shunted current path at the metal/MoS_2 interfaces of the inner electrodes.

As the reviewer mentioned, TLM may also exhibit certain inaccuracy if one does not take caution with the critical factors: high channel uniformity and a couple of short channels for fitting are required for the extraction. Therefore, to improve the accuracy of the contact resistance extraction, we employed a more efficient gate dielectric (100-nm SiN_x) and fabricated five short channels (100 ~ 500 nm) for the TLM devices so that both the slope and the fitting errors for the intersection can be much smaller. In addition, the MOCVD MoS_2 channels exhibit high uniformity: the threshold voltage and electron mobility are the same for different TLM devices (Extended Data Fig. 2c and Fig. R8b). These have been considered the critical factors to make the Rc extraction robust and accurate [C. D. English et al. *Nano Lett.* 16, 3824-3830 (2016), ref. 18]. Based on the simple linear regression model [https://en.wikipedia.org/wiki/Simple_linear_regression], we have calculated the upper and lower bounds of the R_C extraction, as shown in Figure R8a. The R_C of our best Bi-MoS_2 devices lies in the range of 123 ± 63 Ω μm.

In the revised manuscript, we added the following analysis in the Section "Extraction of contact resistance through transfer length method(TLM)" in Methods: "*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. ”

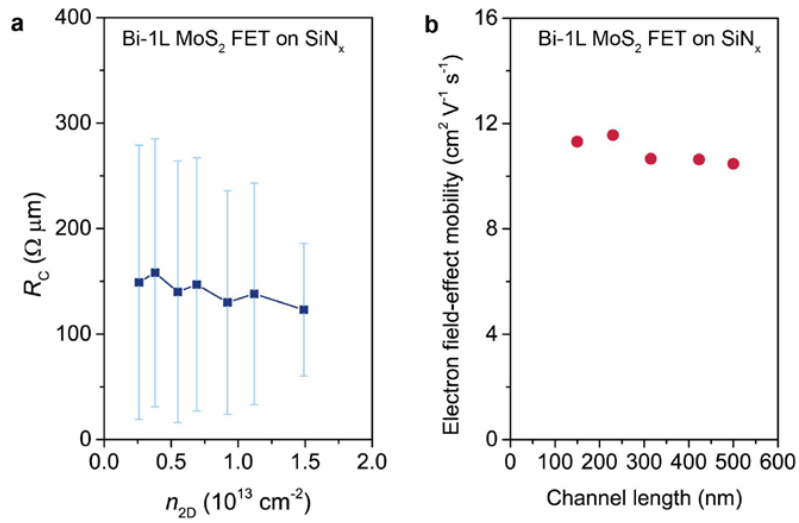

Figure R8. **a.** Contact resistance (R_C) vs carrier density (n_{2D}) induced in the MoS_2. **b.** Field-effect mobility of Bi-MoS_2 TLM devices, showing the high uniformity of the MoS_2 channels.

Please show series for I_d-V_g data at different channel length at V_{DS}=1V. Data from figure 2a is no included in the 2c plot. Why not? Can you please include?

Answer: I_{DS}-V_{GS} curves at different channel lengths have been provided in Extended Data Fig. 2c. The reason why we chose V_{DS}=0.5 V instead of 1 V has been discussed above. The main purpose for Fig. 2a is to provide a rough idea how different metal can impact the device performance. For this purpose, we showed devices fabricated on 300-nm SiO_2 since this is the most commonly used substrate. However, we noted that R_c extraction could be inaccurate for those devices using 300-nm SiO_2 dielectrics since the total device resistance in this case (typically in the order of 100 kΩ μm) can be several orders of magnitude higher than the contact resistance, and as a result, the total resistance-channel length plot (Fig. 2c and Extended Data Fig. 2d) will have a much higher slope and very inaccurate intersection ($2R_C$). To improve the accuracy of the contact resistance extraction, we employed a more efficient gate dielectric (100-nm SiN_x) for the TLM devices so that both the slope and the fitting errors for the intersection can be much smaller.

The paper compares contact resistance with IRDS targets for 2024. This is irrelevant for the technological target. They should be derived from performance in a loaded ring oscillator from implications on delay considering the target drive current.

Answer: We apologize for this mistake. After carefully reviewing the IRDS reports, we agree that the R_c for silicon is irrelevant. We have deleted this in the revised manuscript.

Probably most exciting part of the paper is now relegated to Fig 10 in extended data. Any kind of data from scaled devices especially showing channels scaled to 35nm should be prime and center in the paper itself.

While I_d-V_d data is shown for 35nm channel, I_d-V_G data is shown for 150nm channel length. To prove ohmic contacts, please include data from 35nm channel without Off current degradation, so include I_d-V_g data for L_{ch}=35 nm.

Answer: (Please note that the original Extended Data Fig. 10 has been combined to Extended Data Fig. 7.) Thank you very much for this suggestion. We have included data for 35-nm and 50-nm L_{CH} monolayer MoS_2 transistor here (Figure R9) and in the revised manuscript (Please see Extended Data Fig. 7 a-d). The device exhibits both good on- and off-state performance with an on/off ratio of > 10^6 and I_{ON} up to 1135 μA/μm for L_{CH} = 35 nm.

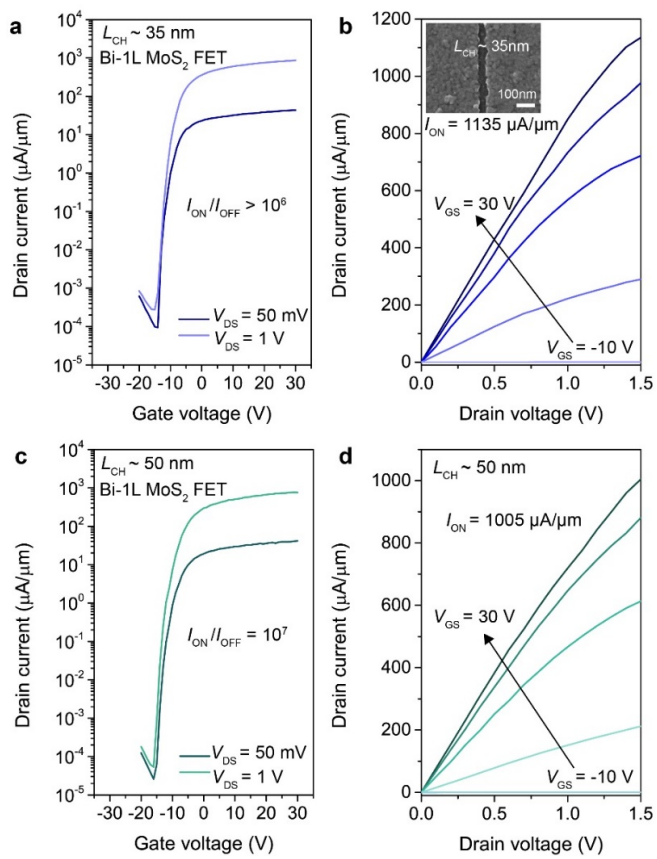

Figure R9. Characteristics of short channel transistors. a,b, Transfer and output characteristics of a 35-nm L_{CH} Bi-MoS_2 FET. Inset of b: SEM image of the 35-nm L_{CH} device. **c,d,** Transfer and output characteristics of a 50-nm L_{CH} Bi-MoS_2 FET.

In the current form, I do not recommend the paper for publication in Nature. Addressing data consistency as described below and including crucial data I_d-V_g at L_{ch}<50nm could make it into the quality and value of reporting we expect from Nature.

We really thank the referee for pointing out these two issues. We hope the revised version of our manuscript has solved all the problems.

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

Thank you for responding to my questions and adding additional information to the paper. Please see file attached for my response to the rebuttal

Response to Rebuttal in different color.

Paper reports high drive current in a back-gated MoS_2 FET through contact resistance reduction. The work ascribes the low-contact resistance to the use of a semi-metal Bi as the contact metal to TMD channel. The data reported in the paper show high currents and an enhanced linearity in the electrical characteristics.

I have the following questions --

1. The authors have proposed that the reason for the unpinning is the semi-metallic nature of an evaporated Bismuth. Outside of its use as a contact metal in MoS_2 transistor, can you share if any other electrical testing was done to confirm the nature of the Bismuth? What is its resistivity? How does it respond to a gate field?

Answer: To characterize the electrical properties of the Bismuth (Bi) contacts, 20 nm of Bi thin film was evaporated on monolayer MOCVD MoS_2 with 100-nm SiN_x and heavily doped silicon as the dielectric and back-gate, respectively (inset of Figure R1a). The whole device architecture is the same as the Bi contacts used in the presented transistors in the manuscript.

As can be seen in Figure R1a, the Bi thin film (or the Bi contacts in this work) clearly shows no gate dependence over the entire range of gate voltages (-40 V ~ 40 V), confirming its metallic nature. The linearity of the output characteristic shown in Figure R1b again suggests the metallic nature of the Bi contact itself. The sheet resistance (R_{SH}) is estimated to be 0.46 kΩ/square, which is two orders of magnitude smaller than that of monolayer semiconducting MoS_2 (for example, R_{SH} ~ 17 kΩ/square for the our MoS_2 channel with a carrier density of 1.5 x 10^{13} cm^2). Therefore, the semi metallic Bi contacts can act well as electrical contacts to 2D semiconductors, as demonstrated in the manuscript. The electrical resistivity of the Bi thin film is estimated to be 9 x 10^{-6} Ω m.

In the revised manuscript, we have added the following sentence in the "Device fabrication and characterization" section of Methods: "*The electrical resistivity of the evaporated bismuth film is measured to be 9×10^{-6} Ω·m.* "

Thank you for providing data about Bismuth material.

2. Can an evaporated Bi metal layer be represented by a band structure? Would the small grain size complicate the picture?

Answer: The TEM SAED image (with aperture size of 1 μm) shows that the crystal orientation is highly aligned and the diffraction pattern of Bi can be clearly visualized, which is a strong evidence that Bi can be described as crystals well depicted by atomic models and first-principles calculations. Therefore, the grain boundaries, which is only a small fraction of the totally area, should not be a dominating factor in alternating contact properties.

Data does point to metal being crystalline. In Extended Data Fig. 3 also suggests that Bi seems to be templating off underlying film. Does the electrical resistivity change if Bi is deposited on SiNx or amorphous substrate? **[Response required]**

3. How does one ensure that the fermi-level of the semi-metal aligns with conduction band edge of the ntype semiconductor? Would one still be able to make a zero-barrier contact if this is not the case?

Answer: From first-principles calculation, we have concluded that the following conditions need be met for an ohmic contact to be realized:

a. The electron hybridization between metal and semiconductor needs to be weak so the metalinduced

gap states are minimized. Bi semimetal has two characteristics to ensure this: (1) The density of states (DOS) of semimetal around Fermi level is zero, so metal-induced gap states (MIGS) is minimal around the Fermi level. (2) The layered structure of Bi semimetal ensures that the electron bonds are completely saturated at the surface, excluding the possibility of having dangling bonds which may induce significant metal-induced gap state. This also requires the semiconductor to be free of dangling bonds, where MoS_2 fortunately is.
b. The work function of the semimetal (or metal) and the electron affinity of the semiconductor *before* contact is important, because if the Fermi level of (semi)metal is not aligned with the bands (either conduction or valence bands) of semiconductor in the first place, no ohmic contact can be formed. For example, it has been experimentally shown that graphene, which is also a semimetal, does not have as good contact with MoS_2, due to the fact that graphene itself has a work function of around 4.7 eV, larger than the electron affinity of MoS_2. We have also predicted in the main text that arsenic doesn't have a good contact with MoS_2, for the same reason. More details can be found in Fig. 3g

Thank you for clarifying this point. I see that the line-up between SM and 2D material is described in the main paper.

4. Arrhenius plots to extract R_contact -- The authors show "normal" Arrhenius behavior with expected gate voltage dependence for Ni contacts. Bismuth, however, shows an opposite slope at high temperatures. This anomalous behavior is attributed to channel resistance dependence on mobility.
a. When the nickel contact is made more "transparent" at higher V_G does the channel resistance dependence on temperature show up?
Answer: Figure R1b shows the Arrhenius plots of the Ni-MoS_2 device at a higher gate voltage presented in the previous Extended Data Fig. 3. Indeed, when the Schottky barrier of Ni/MoS_2 interface becomes more transparent due to a higher electron doping level at the interface, the device is dominated more by the channel resistance and the similar positive slope also shows up. For Ni contacts, this positive slope only happens at a high gate voltage (60 V) and high carrier concentration in the channel ($\sim 4.3\times10^{12}$ cm^{-2}), while a Bi-MoS_2 transistor shows such behavior with a much lower gate voltage and lower carrier concentration in the channel ($\sim 10^{11}$ cm^{-2}), as shown in Fig. 2f and Extended Data Fig. 2b in the revised manuscript.
Figure R2. a. Arrhenius plot of a Ni-MoS_2 FET with different gate voltages (same as previous Extended Data Fig. 3). **b.** Zoom-in plot of **a** focusing on a high gate voltage of 60 V.
b. If the Bi-MoS_2 device is biased in its off-state, the mobility of the channel should cease to matter. I would expect to see a "normal" Arrhenius plot which barrier height determined by the top of barrier in the channel. I request the authors to add this to extended data Fig 3.
Answer: Thank you for pointing this out. We have plotted the Arrhenius plot of the Bi-MoS_2 device biased at a negative gate voltage of -60 V so that the device is in its off-state (the threshold voltage V_T is around 0 V). As can be seen in Figure R3, the device at this condition shows a negative slope in the Arrhenius plot and the effective barrier height is extracted to be ~ 130 meV. As the reviewer suggested, this barrier originates from the energy difference between the Fermi level of the degenerate MoS_2 underneath Bi and the CBM of the depleted MoS_2 channel. We have added this plot into Extended Data Fig.2b (light blue curve) in the revised manuscript.

[Response required] Thank you for showing the rebuttal plots R2 and R3 that confirms my point. In the revised version. The fact that the Nickel shows similar behavior at high VG as the Bi contact shows at lower VG does not necessarily mean the Bi is doing something unexpected. Its just a sign that the authors need to complete the plot in *extended data fig 2b* for all voltages between VG = -60V to 0V to extract an effective barrier height over the entire VG range. This will allow for an extraction of a true SB

height just like in the Nickel or Titanium case. The need for doing the SB height extraction at lower VG than Ni and Ti is also evident from 2.a. This number is critical to prove that the SB is negative / negligible.

5. Can the authors show what the barrier height of Bi contact to WS_2 and WSe_2 is? Does it follow expected trends from electron affinity of the channel?
Answer: In this work we have done a systematic study on MoS_2, we have not measured the barrier heights for WS_2 or WSe_2 yet, which is in our plan for the investigation in the next step. However, following the answer in question 3 above, the increasing trend of R_C (extracted from a virtual source compact model, see Figure R4) between Bi-contacted MoS_2, WS_2, and WSe_2 follows the general trend of decreasing electron affinity of these TMDs [Y. Liu et al, *Sci. Adv.* 2, e1600069 (2016), ref. 15], which in turn implies the possibility of a small increase in Schottky barriers when Bi and these TMDs are in contact. On the other hand, the same trend can also be obtained from DFT calculation where the Fermi level in Bi-WS_2 is lower than Bi-MoS_2, although still above the CBM, as can be found in Extended Fig. 4c and Fig. 3e. This matches with our observation in Figure R4 and Fig. 4d.

6. What is the role of SiN as the gate oxide for the study? What is the channel width used for the MoS_2 1L device with Bi contacts?
Answer: In this work, we presented two different device structures: 1L TMD on 300 nm SiO_2 (Fig. 2) and 1L TMD on 100 nm SiN_x (Fig. 4a-c). SiO_2 is the most commonly used dielectrics for 2D-material-based device studies, so we performed the temperature-dependent measurements and the comparison study for Bi, Ni and Ti contacts on SiO_2, to make these results more consistent with previous studies. In both our experiment and literature (T. Liu et al. *Nat. Nanotech.* 14, 223-226 (2019).), it is observed that SiN_x is a better substrate because TMD tends to have better carrier mobility on SiN_x and the higher thermal conductivity of SiN_x (12 W/m/K for SiN_x and 1.3 W/m/K for SiO_2) can reduce the current degradation due to self-heating for high-performance transistors. We therefore selected SiN_x as the substrate to demonstrate the high-performance transistors.
It should be noted that the selection of substrates does not impact the electrical contact at the Bi-TMD interface. First, the crystallinity of the evaporated Bi on TMD, and thus the resulting interface, should not be affected in any way by the substrate. Second, since the MoS_2 underneath Bi is in a degenerate state as presented in our work, there is no Schottky barrier and no depletion region on the MoS_2 channel side of the contact. As a result, no barrier width modulation exists which normally comes from the different electrostatics determined by the dielectric constant and the thickness of the gate oxide, as is usually the case in previous work (D. Schulman et al. *Chem. Soc. Rev.* 47, 3037-3058 (2017)).
The channel width in this study is in the range of 2 to 10 μm. Since we presented the current density (drain current normalized by channel width) throughout the manuscript, we considered it unnecessary to show the channel width for each device. But based on the reviewer's question here, we have mentioned

the range of channel width in our study in the "Device fabrication and characterization" section of Methods of the revised manuscript: *"The channel widths for the devices in this study are in the range of 2 to 10 μm."*

Thank you for sharing your thoughts on the SiN and channel width.

Another comment, I would recommend including the high drive current data into the main manuscript if possible.

From the data presented in the paper so far I think more evidence is needed to support your argument about Bismuth forming a zero or negative barrier contact. One path to providing convincing evidence is

to show Arrhenius plots below VT and constructing a phisB vs VG plot. The opposite temperature seen in Bi is not unique to Bi but can be accessed in other metals when contact is made transparent (seen in Nickel data presented in rebuttal figures).

Referee #2 (Remarks to the Author):

The authors clearly responded to the reviewer's comments. And, the revised manuscript is well organized and more clarified. Although it is necessary to find the best semimetal species according to semiconductor (TMDs) species to generate gap-state saturation by band alignment between the conduction band of the semiconductor and the Fermi level of the semimetal, this new discovery will be a stepping stone to overcome contact technology. I recommend a publication in Nature for the paper in its present form.

Author Rebuttals to First Revision:

*The responses are shown in blue fonts.

Response to Referee #1:

2. Data does point to metal being crystalline. In Extended Data Fig. 3 also suggests that Bi seems to be templating off underlying film. Does the electrical resistivity change if Bi is deposited on SiNx or amorphous substrate? **[Response required]**

Answer: We fabricated devices of 20-nm Bi film that are directly deposited onto 100-nm SiN_x with heavily doped silicon back-gate. As shown in Figure R1a and b, the Bi film shows metallic characteristics with a similar (slightly higher) resistivity as Bi deposited on MoS_2.

In the revised manuscript, the following sentence has been added to the Methods section:

"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."

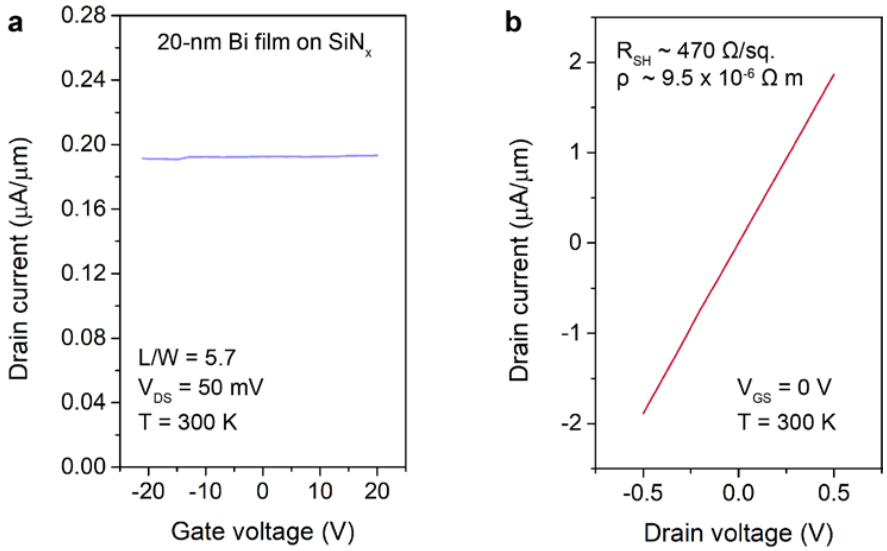

Figure R1. Electrical properties of a 20-nm Bi film evaporated on 100-nm SiNx (without monolayer MoS2).

4. [Response required] Thank you for showing the rebuttal plots R2 and R3 that confirms my point. In the revised version. The fact that the Nickel shows similar behavior at high VG as the Bi contact shows at lower VG does not necessarily mean the Bi is doing something unexpected. Its just a sign that the authors need to complete the plot in extended data fig 2b for all voltages between VG = -60V to 0V to extract an effective barrier height over the entire VG range. This will allow for an extraction of a true SB height just like in the Nickel or Titanium case. The need for doing the SB height extraction at lower VG than Ni and Ti is also evident from 2.a. This number is critical to prove that the SB is negative / negligible.

Answer: Thank you very much for the suggestion. We have included a wider gate voltage range of -50V to -10 V to extract an effective barrier height, as shown in Figure R2 (Other voltages can be found in Extended Data Fig. 2b.). The result is consistent with the key conclusion made in our manuscript. The Schottky barrier height (Φ_{SB}) of Bi-MoS2 FETs is negligible for electron injection. Figure R2b has been included into manuscript as Extended Data Fig. 1c, and the original Extended Data Fig. 1c is moved to the inset.

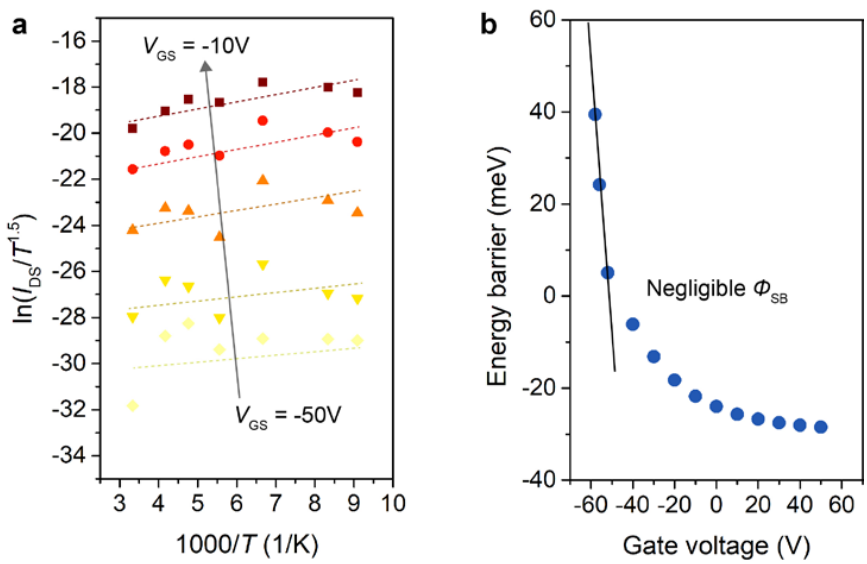

Figure R2. **a.** Arrhenius plot of a Bi-MoS2 FET (same as Extended Data Fig. 2) with gate voltages ranging from -50 V to -10 V. **b.** Schottky barrier height (Φ_{SB}) extraction for the Bi-MoS2 FET, showing a negligible contact barrier.

6. Thank you for sharing your thoughts on the SiN and channel width.

Another comment, I would recommend including the high drive current data into the main manuscript if possible.

From the data presented in the paper so far I think more evidence is needed to support your argument about Bismuth forming a zero or negative barrier contact. One path to providing convincing evidence is to show Arrhenius plots below VT and constructing a phisB vs VG plot. The

opposite temperature seen in Bi is not unique to Bi but can be accessed in other metals when contact is made transparent (seen in Nickel data presented in rebuttal figures).

Answer: Thank you for the suggestions. We agree with your viewpoint and have moved the high drive current data of the 35-nm MoS_2 device into the main manuscript (Fig. 4d).

In addition, as the reviewer suggested, we've plotted Schottky barrier height (Φ_{SB}) versus V_G for Bi contacts to show the negligible barrier for Bi-MoS2 FETs. Please see Figure R2.

Reviewer Reports on the Second Revision:

Referee #1 (Remarks to the Author):

Thank you for answering my questions regarding the Bismuth contacts and sharing the data I requested.

Based on the figure R2 and extended data fig 2, the manuscript states that "nearly saturated slopes at low temperatures observed in the Bi-MoS2 FET indicate the disappearance of an energy barrier for electron injection".

I disagree with this conclusion. It is indeed true that having a negligible barrier would result in weak temperature dependence when VG is close to VT of the transistor. However at very low VG the carriers still need to be emitted over the top of the channel potential controlled by the gate. This portion will exhibit a slope that is governed by the Fermi-tail at the operating temperature. I can see from the Arrhenius data you shared that this is not the case. Do you have a model to explain why in Bi-MoS2 FET the carriers are never blocked off by the barrier in the channel?

The authors have shown exciting results in terms of achieving high currents in 2D material FET. But most of the arguments about linearity are being made a deep in the on-state where the contact can look ohmic if barrier height is small enough. None of this is sufficient proof that Bismuth has a "negative" schottky barrier.

Author Rebuttals to Second Revision:

Referee #1 (Remarks to the Author):

Thank you for answering my questions regarding the Bismuth contacts and sharing the data I requested.

Based on the figure R2 and extended data fig 2, the manuscript states that "nearly saturated slopes at low temperatures observed in the Bi-MoS2 FET indicate the disappearance of an energy barrier for electron injection".

I disagree with this conclusion. It is indeed true that having a negligible barrier would result in weak temperature dependence when VG is close to VT of the transistor. However at very low VG the carriers still need to be emitted over the top of the channel potential controlled by the gate. This portion will exhibit a slope that is governed by the Fermi-tail at the operating temperature. I can see from the Arrhenius data you shared that this is not the case. Do you have a model to explain why in Bi-MoS2 FET the carriers are never blocked off by the

barrier in the channel?

The authors have shown exciting results in terms of achieving high currents in 2D material FET. But most of the arguments about linearity are being made a deep in the on-state where the contact can look ohmic if barrier height is small enough. None of this is sufficient proof that Bismuth has a "negative" Schottky barrier.

Answer: Thank you very much. In fact, we did not intend to claim a “negative” Schottky barrier throughout the whole text. To make it clearer, we have modified relevant sentences in the revised manuscript to clearly state that near-zero Schottky barrier is achieved only when the device is turned on. Below is detailed discussion about the reviewer’s questions.

We agree with the reviewer that the Fermi-tail would contribute to a temperature dependence and an effective “Schottky barrier” when the device is turned off. This is still in agreement with our experimental findings as shown in Figure R2 and Extended Data Figure 2. The positive effective Schottky barrier when V_G is smaller than -50 V is exactly what the reviewer suggested. In Extended Data Figure 2, we showed that the extracted barrier height is 130 meV when V_G is -60 V, and explained that “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.”

As for the extracted negative values of the Schottky barrier height (Φ_{SB}) when V_G is larger, we think this is because the conventional model of thermionic emission over a Schottky barrier used here is inaccurate (Equation 3), given that the model ignored the contribution of the MoS_2 channel resistance which starts to dominate the temperature dependence in our case. Therefore we did not include Figure R2 in main text but leave it in extended figure to avoid readers’ immediate confusion.

On page 5, 2nd paragraph of the revised manuscript, we have changed “*However, this analysis becomes invalid for Bi-MoS_2 FETs*” to “*However, this analysis becomes invalid for Bi-MoS_2 FETs when the device is turned on (V_G>-30 V).*”