

# Impact of Device Scaling on the Electrical Properties of MoS<sub>2</sub> Field-effect Transistors

Goutham Arutchelvan (✉ [goutham.arutchelvan@imec.be](mailto:goutham.arutchelvan@imec.be))

Imec

Quentin Smets

Imec

Devin Verreck

Imec

Zubair Ahmed

Imec

Abhinav Gaur

KU Leuven

Surajit Sutar

Imec

Julien Jussot

Imec

Benjamin Groven

Imec

Marc Heyns

Imec

Dennis Lin

Imec

Inge Asselberghs

Imec

Iuliana Radu

Imec

---

## Research Article

**Keywords:** semiconducting materials, ultimate device scaling, MoS<sub>2</sub> material , capacitive effective oxide thickness (CET)

**Posted Date:** January 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-141737/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Scientific Reports on March 23rd, 2021. See the published version at <https://doi.org/10.1038/s41598-021-85968-y>.

# Impact of device scaling on the electrical properties of MoS<sub>2</sub> field-effect transistors

Goutham Arutchelvan<sup>1,2</sup>, Quentin Smets<sup>1</sup>, Devin Verreck<sup>1</sup>, Zubair Ahmed<sup>1</sup>, Abhinav Gaur<sup>2</sup>, Surajit Sutar<sup>1</sup>, Julien Jussot<sup>1</sup>, Benjamin Groven<sup>1</sup>, Marc Heyns<sup>1,2</sup>, Dennis Lin<sup>1</sup>, Inge Asselberghs<sup>1</sup> and Iuliana Radu<sup>1</sup>

<sup>1</sup> imec, Leuven, Belgium

<sup>2</sup> KU Leuven, Belgium

## I. Abstract

Two-dimensional semiconducting materials are considered as ideal candidates for ultimate device scaling. However, a systematic study on the performance and variability impact of scaling the different device dimensions is still lacking. Here we investigate the scaling behavior across 1300 devices fabricated on large-area grown MoS<sub>2</sub> material with channel length down to 30 nm, contact length down to 13 nm and capacitive effective oxide thickness (CET) down to 1.9 nm. These devices show best-in-class performance with transconductance of 185  $\mu\text{S}/\mu\text{m}$  and a minimum subthreshold swing (SS) of 86 mV/dec. We find that scaling the top-contact length has no impact on the contact resistance and electrostatics of three monolayers MoS<sub>2</sub> transistors, because edge injection is dominant. Further, we identify that SS degradation occurs at short channel length and can be mitigated by reducing the CET and lowering the Schottky barrier height. Finally, using a power performance area (PPA) analysis, we present a roadmap of material improvements to make 2D devices competitive with Silicon gate-all-around devices.

## II. INTRODUCTION

CMOS technology has advertently followed Moore's law of device scaling for the past 50 years to achieve higher transistor density, higher speed and power improvements. A significant part of this device scaling, especially for the planar Metal-Oxide-Semiconductor-Field-Effect-Transistor (MOSFET) was achieved by scaling the gate length<sup>1</sup>. This scaling is reaching its limits as short channel effects (SCE) significantly degrade the device performance. To partially overcome SCE, the tri-gate (FinFET) structure has been introduced<sup>2</sup>. For future technology nodes, the gate-all-around nanosheet FET, which sandwiches thin layers of Silicon channel between multiple gates, is expected to provide additional improvements. Both configurations enhance the electrostatic control over the channel and allow for further gate length scaling. However, it has been reported<sup>3</sup> that the required Silicon channel thickness scaling below 10 nm severely degrades the carrier mobility due to increased surface-roughness scattering. In this context, two-dimensional (2D) semiconducting materials such as transition metal-dichalcogenides (TMDs) are considered to be ideal candidates due to their naturally passivated surface and ultra-thin body (1 monolayer MoS<sub>2</sub> ~ 0.65 nm), providing excellent gate-control and enhanced transport<sup>4,5,6,7</sup>. However, since many studies are performed with manually exfoliated flakes and collecting large datasets is very labor-intensive, there has been a strong focus on only selecting top performing devices, at the cost of less device understanding. Until recently, only a few TMD studies<sup>8,9,10</sup> have focused on devices fabricated using large area grown films. Especially for device scaling<sup>11</sup>, a statistically significant set of data is still lacking.

Therefore, we carry out a study of the impact of geometrical scaling on an extensive data set of large-area grown tri-layer MoS<sub>2</sub> MOSFETs (1300 devices). We investigate the impact of scaling the channel length ( $L_{\text{ch}}$ ) and width ( $W_{\text{ch}}$ ), contact length ( $L_{\text{cont}}$ ) and effective oxide thickness (EOT) on various device performance metrics such as the on- and off-current ( $I_{\text{on}}$ ,  $I_{\text{off}}$ ), contact resistance ( $R_{\text{c}}$ ), subthreshold swing (SS), interface trap density ( $D_{\text{it}}$ ) and threshold voltage ( $V_{\text{T}}$ ). We demonstrate that scaling the contact length down to 13 nm has no impact on the device performance. This confirms that carrier injection occurs exclusively from the edge of the metal directly into the thin TMD channel, which is in line with our TCAD simulations. Further, using our large data set, we make a detailed assessment on the scaling trends of SS and  $V_{\text{T}}$  with device dimensions. We identify the variation in the number of MoS<sub>2</sub> layers in the channel and contact regions as a possible source for SS

degradation and  $V_T$  variability for ultra-scaled TMD MOSFETs. Such insights are crucial for device understanding and enables device architectures such as double-gate<sup>12</sup> or stacked TMD FETs to outperform Si FETs<sup>13</sup>. This article is an extension of our previous work presented at IEDM 2019<sup>10</sup>.

### III. RESULTS AND DISCUSSION

We employ large area MoS<sub>2</sub> grown on a 2" c-plane sapphire template by metal-organic chemical vapor deposition (MOCVD) process using molybdenum hexacarbonyl and dihydrogen sulfide as the precursors. Atomic force microscopy (AFM) shows the MoS<sub>2</sub> is composed of 3 monolayers (ML) fully closed and continuous film, with nucleation of 4 ML and 5 ML island regions (Fig 1a). The average thickness is 3.6 ML, measured using Rutherford backscattering spectrometry (RBS). The device schematic is illustrated in Fig 1b and details of the fabrication process (Fig 1c) are discussed in the Methods section. Three different gate-oxides; (1) 50 nm SiO<sub>2</sub> (2) 12 nm HfO<sub>2</sub>, and (3) 4 nm HfO<sub>2</sub> are used. An optical image after contact deposition is shown in Fig 1d and cross-section TEM images of the final fabricated device are shown in Fig 1e,f.

Direct current measurements are performed in N<sub>2</sub> ambient to avoid any impact of ambient humidity. A total of 1300 devices with varying  $L_{ch}$  (30 nm to 5  $\mu$ m),  $L_{cont}$  (10 nm to 500 nm) and  $W_{ch}$  (200 nm to 10  $\mu$ m) are measured at two different drain-source bias ( $V_{DS} = 0.05$  V, 1 V). Back-gate leakage is low and below the tool noise range (<1 pA) for the 50 nm SiO<sub>2</sub> and 12 nm HfO<sub>2</sub>. Devices with 4 nm HfO<sub>2</sub> have higher gate-drain leakage at  $V_{DS} = 1$  V due to large contact pads. Therefore, the source current ( $I_S$ ), instead of the drain current ( $I_D$ ), is used in their analysis. Channel edge effects are negligible, as confirmed by the constant on-state current density for several  $W_{ch}$  (Fig. S1). Devices with short  $L_{ch}$ , wide  $W_{ch}$  and therefore high absolute current, show a large parasitic voltage drop over the source-drain metal probes, and are therefore ~~are~~ omitted from the analysis. The threshold voltage of the FETs for  $V_{DS} = 0.05$  V, 1 V is obtained by both the linear extrapolation from peak-transconductance ( $V_{T,LE}$ ) and constant-current method ( $V_{T,CC}$  extracted at  $I_D = 10$  nA \*  $W_{ch}/L_{ch}$ ). SS is reported either as  $SS_{min}$ , which is the minimum value across the entire swing, or as  $SS_{CC}$ , extracted at a current level of  $I_D = 1$  nA \*  $W_{ch}/L_{ch}$  for the stated  $V_{DS}$  bias.

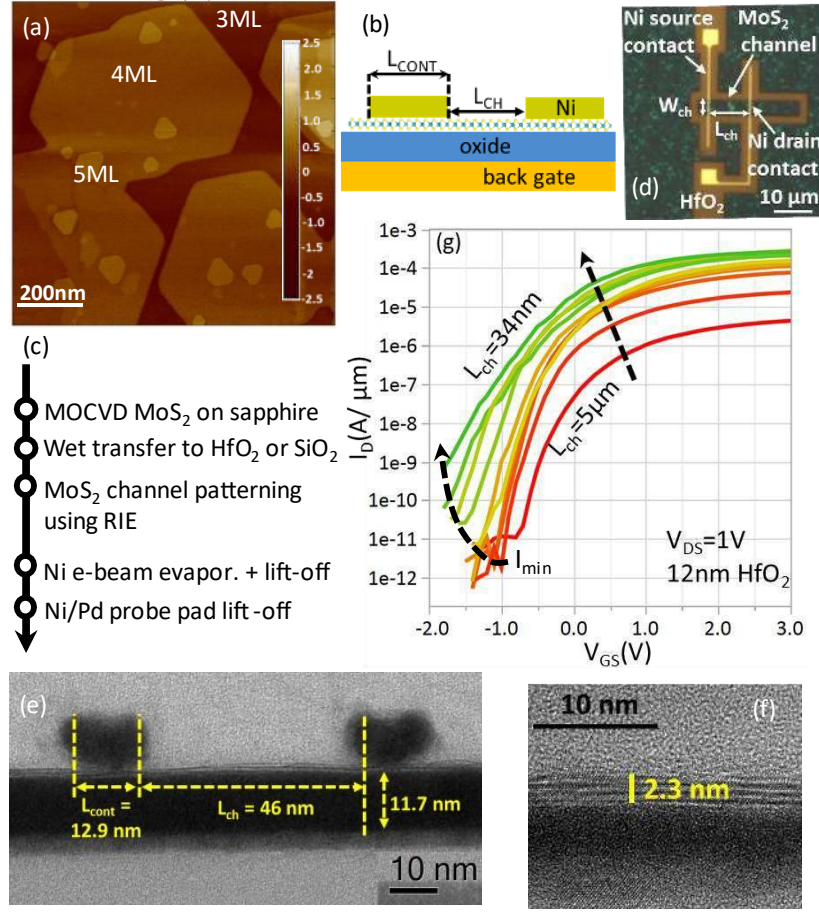


Fig 1. (a) Atomic force micrograph of CVD MoS<sub>2</sub> on sapphire template shows a closed 3ML layer with islands of 4 and 5 monolayers distributed randomly. (b) Device schematic with global back-gate and top source/drain contacts. (c) Fabrication flow for the back-gated devices. (d) Optical micrograph showing the patterned MoS<sub>2</sub> channel with 10 nm Ni contacts. (e) Cross-TEM shows a fabricated device with  $L_{\text{cont}}=13$  nm and  $L_{\text{ch}}=46$  nm on 12 nm HfO<sub>2</sub>. (f) Zoom-in of the channel region for another device showing 3 monolayer MoS<sub>2</sub> on nominal 4 nm HfO<sub>2</sub> (g) Transfer characteristics at a fixed  $V_{\text{DS}}=1$  V. Maximum drive current at  $V_{\text{GS}}=3$  V scales with  $L_{\text{ch}}$  saturating for short-channel devices. The plot shows  $L_{\text{ch}}=34$  nm, 44 nm, 50 nm, 70 nm, 100 nm, 200 nm, 300 nm, 500 nm, 1000 nm, 5000 nm.

## A. Scaling of on- and off- state currents

From the representative transfer characteristics in Fig 1g, we observe that the off-state current significantly increases as  $L_{\text{ch}}$  is scaled, as a result of a loss of gate control. Accordingly, we extract the minimum current in the entire back gate sweep ( $I_{\text{min}}$ ), and we observe that it is the same for both oxides and lower than the noise floor of the tool ( $<1$  pA). However, when comparing the  $I_{\text{off}}$  in the scatterplot Fig 2a, which is extracted at a fixed displacement field of 0.4 V/nm below  $V_{\text{T,CC}}$ , we note that the HfO<sub>2</sub> sample exhibits higher  $I_{\text{off}}$  compared to the SiO<sub>2</sub> sample. This suggests that the subthreshold swing is limited by the high interface trap density (see Section D). We also note that for both oxides,  $I_{\text{off}}$  degrades with smaller  $L_{\text{ch}}$ . This is mainly due to SS degradation observed for short  $L_{\text{ch}}$  devices, and will be further discussed in Section E.

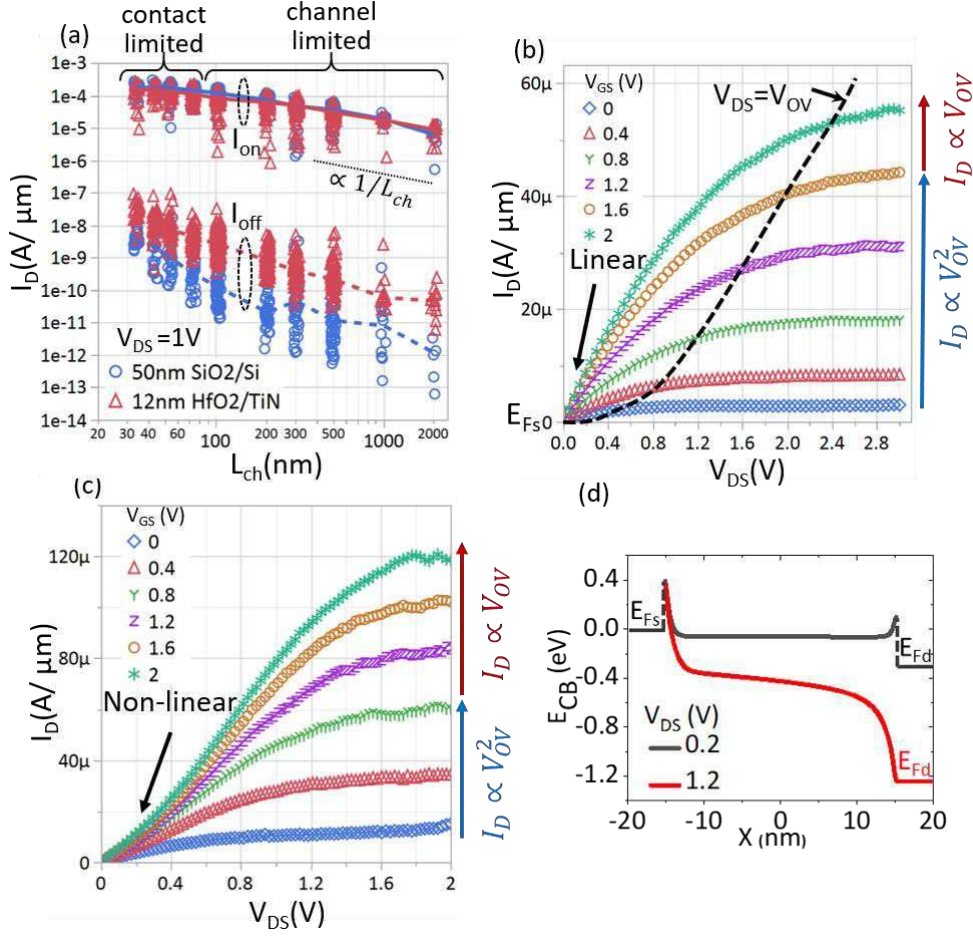


Fig 2. (a) Scatter plot (with median line) showing  $I_{on}$  extracted at  $n_s = 1e13$  cm<sup>-2</sup> and  $I_{off}$  at a fixed displacement field of 0.4 V/nm below  $V_{T,CC}$  for  $V_{DS} = 1$  V.  $I_{on}$  for the 50 nm SiO<sub>2</sub> and 12 nm HfO<sub>2</sub> devices overlap indicating no impact on low-field mobility and contact barrier.  $I_{on}$  roughly scales as  $1/L_{ch}$  for  $L_{ch} > 500$  nm and saturates for  $L_{ch} < 50$  nm.  $I_{off}$  is higher for HfO<sub>2</sub> compared to SiO<sub>2</sub> due to high interface trap density. (b)  $I_D$ - $V_{DS}$  for  $L_{ch} = 500$  nm shows linear triode regime and saturation at high  $V_{DS}$ . The dashed line follows the current at  $V_{DS} = V_{OV}$ . While the onset of saturation follows the  $V_{OV}$  at low  $V_{GS}$ , it saturates at  $V_{DS} = 2.4$  V for high  $V_{GS}$ . The saturation current roughly scales  $V_{OV}^2$  and  $V_{OV}^1$ , at low and high  $V_{GS}$ , respectively. (c)  $I_D$ - $V_{DS}$  for  $L_{ch} = 30$  nm shows non-linear triode regime due to Schottky contacts and saturation at high  $V_{DS}$ . Saturation current follows a similar trend as  $L_{ch} = 500$  nm but  $V_{DS}$  at onset of velocity saturation is reduced to 1.4 V. (d) Conduction band profile for  $L_{ch} = 30$  nm device with Schottky contacts shown for low and high  $V_{DS}$ . The Fermi-level at the source and drain are indicated by  $E_{FS}$  and  $E_{FD}$ , respectively. The Schottky barrier is shown as the abrupt potential change at the contact-channel interface. At low  $V_{DS}$ ,  $I_D$  is determined by Schottky contacts. At high  $V_{DS}$ , though the potential drops significantly at the source contact, velocity saturation near the drain determines the  $I_D$  characteristics.

1           Next, we evaluate the  $I_{on}$  at a fixed charge density ( $n_s$ ) of  $10^{13}$  cm<sup>-2</sup> and do not observe any difference  
2   between the 50 nm SiO<sub>2</sub> and 12 nm HfO<sub>2</sub> samples (Fig 2a). This indicates that the carrier transport in the MoS<sub>2</sub>  
3   channel is predominantly limited by charged impurities<sup>14</sup> in the MoS<sub>2</sub> or at the interfaces, and not by remote  
4   phonons<sup>15</sup> in the gate oxide.

5           For the  $I_{on}$ , two distinct channel length scaling regimes can be identified in Fig 2a. In the long-channel  
6   limit ( $\sim L_{ch} > 500$  nm), the  $I_{on}$  increases roughly proportional to  $1/L_{ch}$  and the device operates in the triode region

(illustrated in Fig 2b for the 12 nm HfO<sub>2</sub> sample and  $L_{ch} = 500$  nm) i.e. gate-overdrive ( $V_{OV} = V_{GS} - V_T$ )  $> V_{DS}$  for both oxides. The drain current also exhibits strongly linear dependence with  $V_{DS}$  in the triode region (Fig 2b), suggesting that the channel resistance is dominant for this  $L_{ch}$  and beyond. We also extract a low-field-effect mobility of  $\sim 15$  cm<sup>2</sup>/V.s (inset of Fig 3c) using the transfer length method (TLM) for both the samples with 12 nm HfO<sub>2</sub> and 50 nm SiO<sub>2</sub>. At higher lateral electric field (higher  $V_{DS}$ ),  $I_D$  saturates (Fig 2b), and the saturation current scales quadratically with  $V_{OV}$  (here  $V_{T,CC} = -0.4$  V) due to channel pinch-off near the drain. However, for the highest  $V_{OV}$  ( $\sim 2 - 2.4$  V), the saturation current scales roughly linear with  $V_{OV}$ , indicating that it is limited by saturation of drift velocity at high lateral-field<sup>16</sup> ( $F_{LATERAL} > 5$  V/ $\mu$ m).

In the short-channel limit ( $\sim L_{ch} < 50$  nm), the dependence of  $I_{on}$  on  $L_{ch}$  saturates (Fig 2a). Accordingly, in the output characteristics for  $L_{ch} = 30$  nm (Fig 2c), we make two observations; (1) super-linear  $I_D$  for  $V_{DS} < 0.4$  V and (2) saturation of  $I_D$  for  $V_{DS} > 1.4$  V. The distinct super-linear dependence of  $I_D$  with  $V_{DS}$  (Fig 2c) suggests that the Schottky contacts at the metal-MoS<sub>2</sub> interface limit the current even though the bias conditions ( $V_{OV} > V_{DS}$ , here  $V_{T,CC} = -0.3$  V) ensure that the channel is continuously accumulated with electrons. At higher  $V_{DS}$ ,  $I_D$  saturates similarly to the  $L_{ch} = 500$  nm device. The current at the onset of saturation is roughly proportional to  $V_{OV}^{1.5-1.7}$  and  $V_{OV}^{0.8-0.9}$  for low and high  $V_{OV}$ , respectively, closely following the long-channel characteristics. This indicates that while contact resistance dominates at low  $V_{DS}$ , velocity saturation near the drain ~~likely~~ determines the current at high  $V_{DS}$ .

We can further understand both these observations from the simulated conduction band profile of  $L_{ch} = 30$  nm device (Fig 2d) for low and high  $V_{DS}$ . In the linear regime ( $V_{DS} = 0.2$  V and  $V_{OV} > V_{DS}$ ), the drain-source potential is predominantly dropped across the reverse-biased source and forward-biased drain Schottky contacts. With increasing  $V_{DS}$  (higher lateral field), the transmission probability across the Schottky contacts increases rapidly, especially across the reverse-biased source, giving rise to the super-linear dependence of  $I_D$  with  $V_{DS}$ . At even higher  $V_{DS}$  ( $V_{DS} = 1.2$  V), the electric field in the channel near the drain is large enough that the drift velocity saturates, which saturates the current.

## B. Contact length scaling

Fig. 3a shows that  $I_{on}$  (@  $n_s = 10^{13}$  cm<sup>-2</sup>) does not degrade as  $L_{cont}$  is scaled down to 13 nm. This agrees with TCAD simulations<sup>10,17,18</sup> that predict contact edge injection of carriers for 1-3 layers of MoS<sub>2</sub> channel. This



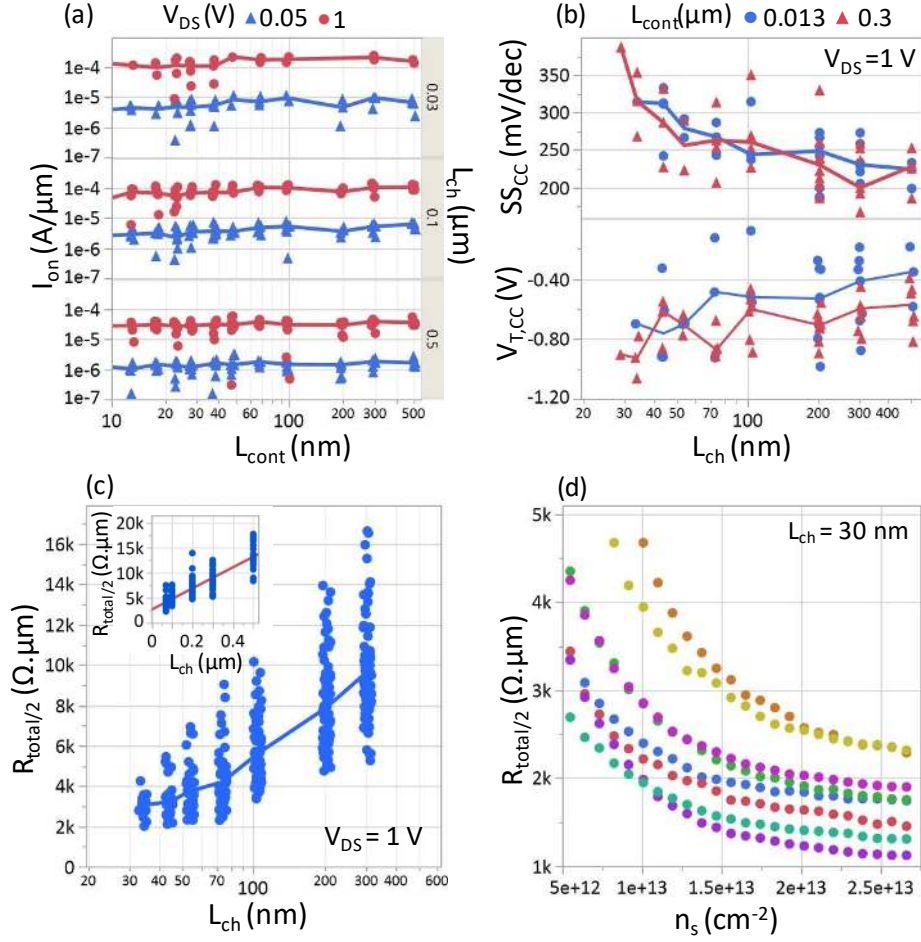


Fig 3. Scatter plot (with median line) of (a)  $I_{on}$  (at  $n_s = 1e13 \text{ cm}^{-2}$ ) versus  $L_{cont}$  for  $L_{ch} = 30 \text{ nm}$  (contact-limited),  $100 \text{ nm}$  (intermediate regime), and  $500 \text{ nm}$  (mobility limited). No dependence on  $L_{cont}$  down to  $13 \text{ nm}$  indicates carrier injection from the edge of the metal directly into the  $\text{MoS}_2$  channel with  $L_T < 13 \text{ nm}$ . (b)  $SS_{CC}$  and  $V_{T,CC}$  versus  $L_{ch}$  for  $L_{cont} = 13 \text{ nm}$  and  $300 \text{ nm}$ . No systematic deviation with  $L_{cont}$  indicates identical electrostatics in both cases. (c)  $R_{total/2}$  (at  $n_s = 1e13 \text{ cm}^{-2}$ ) versus  $L_{ch}$  show saturation below  $L_{ch} = 50 \text{ nm}$  due to contact resistance. Upper limit for  $R_C$  is obtained as median  $R_{total/2}$  for  $L_{ch} = 30 \text{ nm}$ . Median  $R_C$  values of  $3 \text{ k}\Omega \cdot \mu\text{m}$  with best performers at  $2 \text{ k}\Omega \cdot \mu\text{m}$  are obtained. (inset) TLM fit of  $R_{total/2}$  (at  $n_s = 1e13 \text{ cm}^{-2}$ ) versus  $L_{ch}$  gives  $R_C = 2.7 \text{ k}\Omega \cdot \mu\text{m}$  and field-effect mobility  $= 15 \text{ cm}^2/\text{V}\cdot\text{s}$ . (d)  $R_{total/2}$  versus  $n_s$  for  $L_{ch} = 30 \text{ nm}$  at  $V_{DS} = 1 \text{ V}$  of 8 devices.  $R_C$  significantly reduces at  $n_s = 2e13 \text{ cm}^{-2}$  due to better carrier injection into the accumulated channel.

1 observation holds true for three different  $L_{ch}$  ( $30 \text{ nm}$ ,  $100 \text{ nm}$ ,  $500 \text{ nm}$ ) over a wide range of  $L_{cont}$  ( $500 \text{ nm}$  to  
2  $13 \text{ nm}$ ) and for varying lateral field ( $V_{DS} = 0.05 \text{ V}$ ,  $1 \text{ V}$ ). In all three cases, as predicted, we do not observe any  
3 systematic degradation of  $I_{on}$  by scaling down  $L_{cont}$  from  $500 \text{ nm}$  to  $13 \text{ nm}$ . Even for the shortest  $L_{ch} = 30 \text{ nm}$ ,  
4 where the channel resistance is negligible and the device is Schottky contact limited ( $I_D$ - $V_{DS}$  is super-linear at  
5 low  $V_{DS}$  in Fig 2c), the contact resistance is independent of  $L_{cont}$ . Moreover, the electrostatic properties of the  
6 device are also unaffected by scaling down  $L_{cont}$  as can be seen in Fig 3b from the trend of  $SS_{CC}$  and  $V_{T,CC}$  ( $@ V_{DS}$   
7  $= 1 \text{ V}$ ) with  $L_{ch}$  for two extreme contact lengths. The SS degradation and  $V_T$  roll-off with shorter  $L_{ch}$ , are  
8 independent of the contact length. The insensitivity to  $L_{cont}$  scaling also holds for other gate-oxides and charge

densities (plots not shown). In summary, for 3 ML MoS<sub>2</sub>, the active region of MoS<sub>2</sub> under the metal contact where most of the electrons get injected (called the transfer length  $L_T$ ) is at least below 13 nm.

These results agree very well with our previous TCAD simulations. For thin MoS<sub>2</sub> (1-3 ML), these predict  $L_T$  smaller than the minimum simulated  $L_{\text{cont}}$  of 2 nm (Fig. S2). This is caused by the Schottky barrier (SB) at metal-MoS<sub>2</sub> interface, which depletes the MoS<sub>2</sub> underneath and prevents vertical electron injection. Therefore, injection is only allowed from the edge of the metal contact directly into the carrier-rich channel, which is also predicted in other work<sup>18,19</sup>. For thicker MoS<sub>2</sub> (more than 5 ML), the MoS<sub>2</sub> region underneath the contact is no longer depleted and a longer section of the contact contributes to carrier injection<sup>20,21,22</sup>.

In other work<sup>21,23,24,25</sup>, transfer lengths of 80 nm to 630 nm have been calculated using the transfer length method (x-axis intercept), but those values are in contradiction with our results. As argued elsewhere<sup>26</sup>, this method should not be used for thin TMD layers and Schottky contacts. The Schottky barrier fully depletes the TMD below, therefore the sheet resistance below the contact and in the channel are not the same, which is a requirement of the transfer length method. However, the transfer length method can still be reliably used for mobility calculation, because it does not have this requirement of identical TMD sheet resistance in the channel and below the metal.

### C. Contact resistance extraction

As we found in *Section A* that devices become more contact dominated as  $L_{\text{ch}}$  is scaled, we now take a closer look at the value of the contact resistance. We extract the contact resistance ( $R_c$ ) directly as half of the total device resistance ( $R_{\text{tot}}/2$ ) for devices with the shortest  $L_{\text{ch}} = 30$  nm, without any need for extrapolation like in the TLM method. By considering  $R_c \sim R_{\text{tot}}/2$ , an upper limit is obtained for  $R_c$ , as it assumes negligible channel resistance. Fig 3c shows a plot of  $R_{\text{tot}}/2$  at a charge density of  $10^{13} \text{ cm}^{-2}$  vs  $L_{\text{ch}}$ . For  $L_{\text{ch}} < 50$  nm, the  $R_{\text{tot}}/2$  saturates, and we obtain a median Nickel-MoS<sub>2</sub>  $R_c \sim 3 \text{ k}\Omega\cdot\mu\text{m}$  (at  $n_s = 10^{13} \text{ cm}^{-2}$ ), which is in good agreement with  $R_c$  extracted using TLM (inset of Fig 3c). Our  $R_c$  values are comparable to the state-of-the-art devices which have been demonstrated with Au<sup>20</sup> or Indium<sup>27</sup> contact metals. For increased  $V_{\text{OV}}$ , the contact resistance further drops due to better carrier injection into the accumulated channel, and we obtain  $R_c \sim 1.2 - 2 \text{ k}\Omega\cdot\mu\text{m}$  @  $n_s = 2 \times 10^{13} \text{ cm}^{-2}$  (Fig 3d). For even higher carrier densities (compare  $n_s = 2 \times 10^{13} \text{ cm}^{-2}$  to  $2.7 \times 10^{13} \text{ cm}^{-2}$ ),  $R_c$  no longer improves significantly. Significant device-to-device variation in contact resistance is

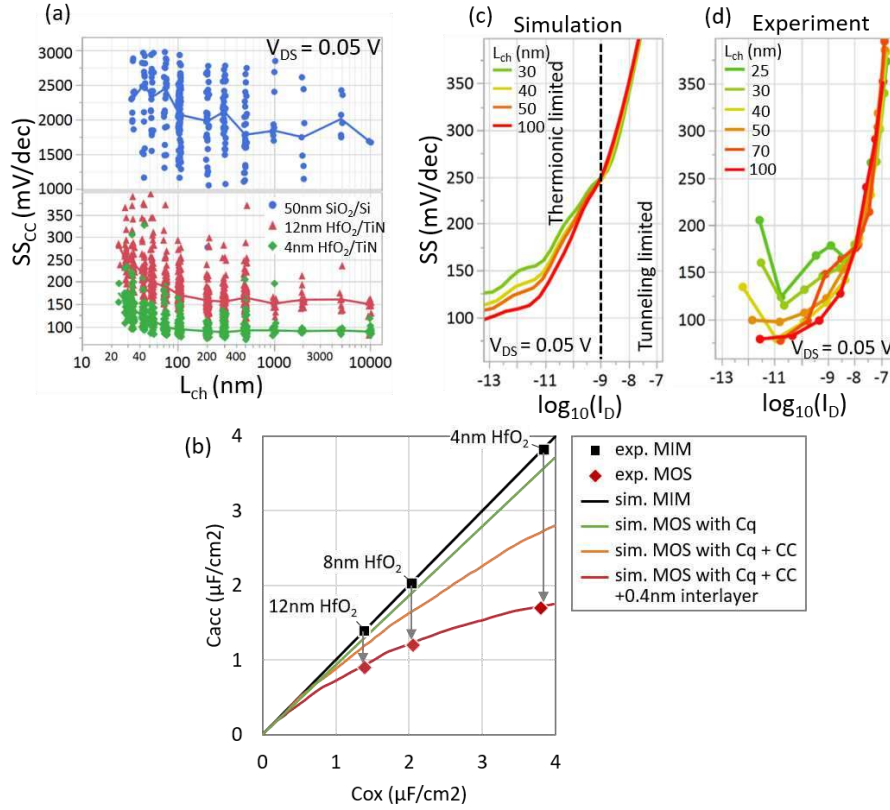


Fig 4. (a) Scatter plot (with median line) of  $SS_{CC}$  versus  $L_{ch}$  for the three different oxides. While  $SS$  improves with lower EOT, the degradation and scatter for short channel devices are attributed to electrostatic potential fluctuations caused by non-uniform thickness of MoS<sub>2</sub> in the contact and channel regions. (b) Experimentally measured maximum accumulation capacitance from MOScap ( $C_{max}$ ) versus MIMcap capacitance ( $C_{ox}$ ). Systematically, the  $C_{max}$  is lower than  $C_{ox}$  corresponding to an additional 1 nm CET over the measured EOT. Simulations show this is caused by the quantum capacitance  $C_q$  (MoS<sub>2</sub> having lower DOS than metal), the impact of the charge centroid (CC) further away in MOS than MIM, and additionally due to 0.4 nm of water or carbon residues stuck at the HfO<sub>2</sub>/MoS<sub>2</sub> interface during transfer. Qualitative comparison between (c) simulated and (d) experimental  $SS$  versus  $\log(I_D)$  for different  $L_{ch}$ . The simulated  $SS$  is for a uniform 3 monolayers MoS<sub>2</sub> with SBH = 0.45 eV. Two transport regimes at the contacts—thermionic emission and tunneling through the SB are identified. In the thermionic regime, the relative increase of field in the channel from the source/drain Schottky contacts degrades gate control for short  $L_{ch}$  devices. In the tunneling regime, the nearly equal tunneling lengths for the different  $L_{ch}$  results in a similar but degraded  $SS$  compared to the thermionic regime.

observed, possible due to polymer residues between the contact metal and the MoS<sub>2</sub>, which were not completely removed after the transfer and contact lithography steps of the fabrication flow.

#### D. Long channel electrostatics and $D_{it}$ extraction

Fig 4a shows that the subthreshold swing  $SS_{CC}$  obtained at  $V_{DS} = 0.05$  V for different  $L_{ch}$ , improves with thinner back-gate oxide due to better gate control of the charge in the channel. Consequently, we achieve the best subthreshold swing for the devices on 4 nm HfO<sub>2</sub> substrate (Fig. S3) with median  $SS_{min} = 90$  mV/dec and 110 mV/dec (at  $V_{DS} = 0.05$  V) for  $L_{ch} = 50$  nm and 30 nm, respectively.

In the long-channel limit i.e.  $L_{ch} > 1 \mu m$ ,  $SS_{CC}$  saturates to a constant median value of 80 mV/dec, 150 mV/dec, 1800 mV/dec for 4 nm  $HfO_2$ , 12 nm  $HfO_2$ , and 50 nm  $SiO_2$  respectively. This is determined by the charging of  $MoS_2$ /oxide interface and channel defects ( $60^\circ$  grain boundaries<sup>28</sup>, and point defects<sup>29</sup>), for which we calculate a trap density ( $D_{it,min}$ ) of  $4.5 - 7 \times 10^{12} \text{ cm}^{-2}\text{eV}^{-1}$  from  $SS_{min}$ . We also confirm this  $D_{it}$  value using multi-frequency C-V measurements of  $TiN/HfO_2/MoS_2$  MOScap<sup>30</sup>, where we obtain an acceptor-type trap density of  $3.2 - 6 \times 10^{12} \text{ cm}^{-2}\text{eV}^{-1}$  with energy levels near the midgap.

From C-V measurements, we find that the MOS capacitance is systematically lower than the target oxide capacitance due to exposure to water and/or atmospheric carbon during the wet transfer process from the sapphire template to the target substrates. Fig 4b shows how the maximum accumulation capacitance ( $C_{max}$ ) measured from  $TiN/HfO_2/MoS_2$  MOScap (shown as the red diamonds) is lower than the value of  $C_{ox}$  measured from  $TiN/HfO_2/TiN$  MIMcap (without  $MoS_2$ , shown as the black line). Equivalently, the capacitance equivalent oxide thickness (CET) values for MOScap (1.9 nm, 2.7 nm, and 3.8 nm) are systematically 1 nm higher than the EOT values of the MIMcaps (0.9 nm, 1.7 nm, and 2.6 nm). We calculate that the effect of quantum capacitance due to the limited density of states in  $MoS_2$ , and the effect of charge centroid being a few angstrom away from the interface, are insufficient to account for this 1 nm difference. As the MIMcaps are not exposed to water or polymer during the fabrication, Fig 4b shows the difference between the CET and EOT values can be explained by a 0.4nm thick layer of water or hydrocarbons adsorbed from the ambient, or a combination thereof. In the future, we expect dry transfer in a controlled ambient will lower the CET, closer to the nominal EOT.

## **E. Short channel electrostatic degradation and variability**

In the short-channel limit, i.e.  $L_{ch} < 100 \text{ nm}$ , Fig 4a shows a degradation of median  $SS_{CC}$  but also increased scatter ( $SS_{CC}$  at  $V_{DS} = 1 \text{ V}$  in Fig. S4). A similar trend is also seen for  $SS_{min}$  (Fig. S3). We hypothesize that the increased median and scatter could both be caused by the Schottky contacts, where the median  $SS$  degradation with shorter  $L_{ch}$  is related to the relative increase of depletion regions from the Schottky contacts, while the scatter could be due to the variation in Schottky barrier height<sup>31</sup> (SBH) induced by the non-uniform thickness of the  $MoS_2$ , seen in the AFM image in Fig. 1a.

1        We first verify the hypothesis of degraded median SS for shorter  $L_{ch}$  by comparing representative  
2        experimental SS versus  $I_D$  curves to simulations in Fig 4c. We consider full SS -  $I_D$  curves instead of extracting  
3        SS at a single current level to understand the injection mechanism in a wider operation range. The simulations  
4        are performed for a SBH = 0.45 eV and uniform 3 ML MoS<sub>2</sub> channel. We observe two different regimes for SS  
5        for both the simulated and experimental data. In the first low-current regime ( $I_D < 1e-9$  A/ $\mu$ m), the current is  
6        limited by the thermionic emission of carriers from the metal into the channel. Here, the barrier for electrons  
7        consists of the highest position of conduction band edge inside the channel determined by the gate-bias. In this  
8        low-current regime, SS is determined by the change in the conduction band edge with gate-bias. As discussed  
9        in section D, the lower limit for SS (which corresponds to SS<sub>min</sub> in fig.S3), is defined by the interface trap  
10       density. The degradation of SS<sub>min</sub> for short- $L_{ch}$  devices is due to the electrostatic potential of the source and  
11       drain metallurgical junctions influencing the channel potential and degrading the gate control. This is illustrated

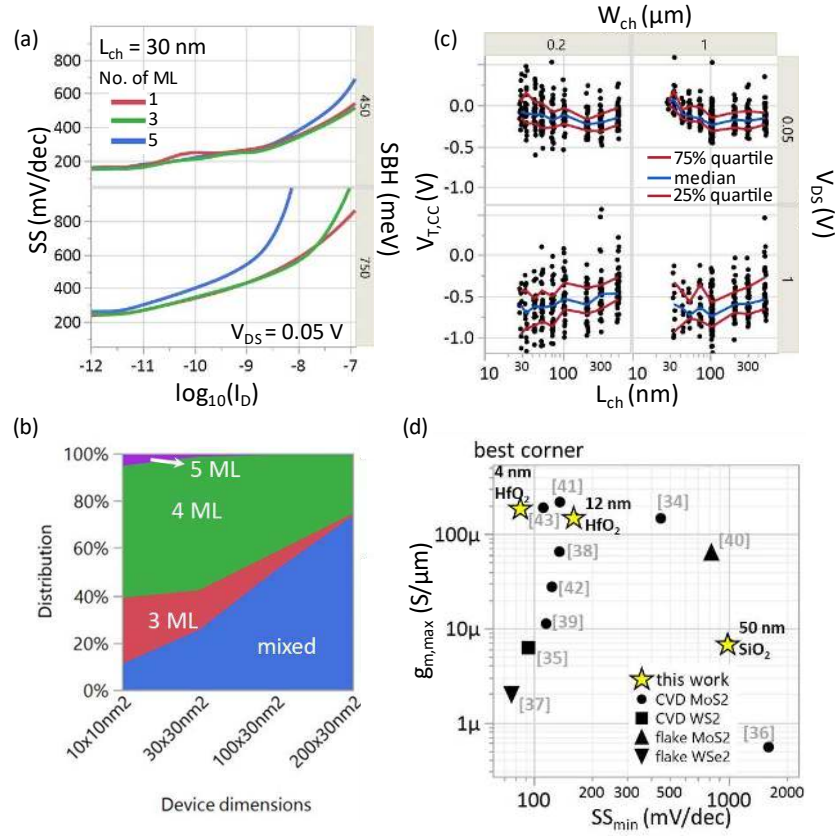


Fig 5. (a) Simulated SS versus  $\log(I_D)$  for a uniform layer of 1, 3 and 5 monolayers of MoS<sub>2</sub> for SBH = 0.45 eV and 0.75 eV. For  $I_D > 1e-9$  A/ $\mu$ m, tunneling through Schottky barrier determines the SS. Subsequently, a thinner channel results in better gate control, shorter tunneling length and therefore better SS. (b) Probability distribution versus device dimensions ( $L_{ch} \times W_{ch}$ ). AFM from Fig 1a was used to compute the probability distribution for fabricating devices on only 3, 4, 5 or a combination of those (mixed). Our experimental devices have a 60-70% probability of being mixed, leading to non-uniform gate control across the channel and contact regions. (c)  $V_{T,CC}$  versus  $L_{ch}$  for two different  $W_{ch}$  (200 nm, 1000 nm) and  $V_{DS}$  (0.05 V, 1 V). No  $V_T$  roll-off at  $V_{DS} = 0.05$  V due to excellent gate control over the channel for 12 nm HfO<sub>2</sub>.  $V_T$  roll-off of about 200 mV for  $V_{DS} = 1$  V due to higher lateral-field at the source contact allowing for more carrier injection. No systematic  $V_T$  deviation between  $W_{ch} = 1$   $\mu$ m and 200 nm. (d) Benchmark plot showing  $g_{m,max}$  versus  $SS_{min}$ . All values are at  $V_{DS} = 1$  V except [39] -  $V_{DS} = 0.1$  V, [34] -  $V_{DS} = 0.5$  V, [40] -  $V_{DS} = 1.2$  V, [37] -  $V_{DS} = 1.5$  V. In this work, 4 nm HfO<sub>2</sub> provides best SS = 86 mV/dec and  $g_{m,max} = 185$   $\mu$ S/ $\mu$ m.

in Fig. S5 where the conduction band energy is flat over most of the device for  $L_{ch} = 100$  nm, while it is lowered for  $L_{ch} = 30$  nm with the region of maximum barrier reducing to a small portion near the center of the device. Note that this effect is similar to conventional MOSFETs.

The second regime ( $I_D > 1e-9$  A/ $\mu$ m) is reached when the conduction band in the channel is lowered further, and carriers can efficiently tunnel through the SB (Fig. S6). Here, the thermionic component over the barrier saturates and the tunneling path length determines the current. Because it continuously changes with higher  $V_g$ , the SS is worse than the first regime. Correspondingly, in the experimental devices, the  $SS_{CC}$  extracted at  $I_D > 1e-8$  A/ $\mu$ m (for  $L_{ch} < 100$  nm) shows a higher value than  $SS_{min}$  and stronger degradation with  $L_{ch}$ . The SS for a given  $I_D$  also becomes nearly independent of  $L_{ch}$ , because the tunneling path length depends

only on the gate voltage and the thicknesses and dielectric permittivities of the TMD<sup>32</sup> and oxide, for the low lateral electric field ( $V_{DS} = 0.05$  V). This is illustrated in Fig. S6 where the conduction band energy and tunneling rate are plotted along the edge carrier injection path for  $L_{ch} = 30$  nm and 100 nm, showing no significant difference. With further reduction in SBH, the SS value in the second regime improves, reaching closer to the thermionic limit of the first regime.

We study the increased SS scatter for short  $L_{ch}$  seen experimentally, using simulations of devices with different uniform MoS<sub>2</sub> channel thickness and SBH. Fig 5a shows the simulated SS value for two different SBH (0.45 eV, 0.75 eV) and three different uniform thicknesses (1,3 and 5 layers) of MoS<sub>2</sub> for  $L_{ch} = 30$  nm. Similar to the above case, we note two different regimes for SS irrespective of the barrier height. For the first regime of low  $I_D$  ( $< 1e-8$  A/ $\mu$ m for SBH = 0.45 eV and  $< 1e-11$  A/ $\mu$ m for SBH = 0.75 eV), the SS is determined only by thermionic emission over the channel barrier. Therefore, the SS is independent on the channel thickness. However, the SS degrades for SBH=0.75 eV compared to 0.45 eV, because the higher Schottky barrier field penetrates deeper into the channel. For the second regime of high  $I_D$  ( $> 1e-8$  A/ $\mu$ m for SBH = 0.45 eV and  $> 1e-10$  A/ $\mu$ m for SBH = 0.75 eV), the SS is dependent on the tunneling length which is sensitive to the thickness of the semiconductor among other parameters<sup>33</sup>. Subsequently, the gate control over the Schottky barrier, and hence the tunneling length, reduces with thicker MoS<sub>2</sub>, resulting in poor SS for the 5 ML MoS<sub>2</sub> (Fig. S7). In agreement with this observation, we also note that the difference in SS between the layers is more pronounced for the higher SBH of 0.75 eV.

In our experiments, we have even more variability due to non-uniform thickness within a single device. Even for the smallest functional device footprint ( $L_{ch} \sim 30$  nm \*  $W_{ch} \sim 200$  nm), we always have a high probability ( $\sim 70\%$ ) of having a mixed device i.e. regions of 3, 4 and 5 layers of MoS<sub>2</sub> within the same device. This is illustrated in Fig 5b where the representative AFM (Fig 1a) image of the material was used to compute the probability of fabricating devices with different dimensions on only 3 (or) 4 (or) 5 or a mix of those layers. These mixed-thickness devices, together with the associated SBH variations, would result in non-uniform gate control and large scatter in the SS values of experimental devices.



## F. Threshold voltage control

We analyse  $V_T$  control for decreasing channel length, and Fig 5c shows that there is no significant median  $V_T$  roll-off at  $V_{DS} = 0.05$  V. With a higher  $V_{DS} = 1$  V, we notice a  $V_T$  roll-off of about 200 mV from  $L_{ch} = 500$  nm to 30 nm. We attribute this roll-off to the higher lateral electric field across the reverse-biased Schottky contact, because  $V_{DS}$  is fixed at 1 V for all  $L_{ch}$ . This higher electric field allows for increased carrier injection in short channel devices, which lowers  $V_T$ . This roll-off could be mitigated by improving the gate control through gate-oxide scaling, or by reducing the amount of defects at the MoS<sub>2</sub>/oxide interface.

$V_T$  control for decreasing channel width is also shown in Fig 5c, and no systematic impact is seen as  $W_{ch}$  is scaled from 1  $\mu$ m down to 200 nm. However, we note that there is an increased  $V_T$  variability for all  $L_{ch}$  devices with  $W_{ch} = 200$  nm compared to  $W_{ch} = 1$   $\mu$ m, especially at  $V_{DS} = 0.05$  V. This increased  $V_T$  variability could be attributed to the higher probability of finding devices on discrete layers (Fig 5b) for narrower channel compared to a wider channel where the devices are always mixed. Other sources of variability such as bias-temperature instability, non-uniformity of the MoS<sub>2</sub> grains etc. could also impact the  $V_T$  variability and more dedicated experiments are required.

## IV. BENCHMARK, PROJECTION AND CONCLUSION

We present a benchmark chart (Fig 5d) to compare the performance of our devices against flake and CVD 2D material FETs in literature<sup>34,35,36,37,38,39,40,41,42,43</sup>. We choose the peak of transconductance ( $g_{m,max}$ ) measured at  $V_{DS} = 1$  V and  $SS_{min}$  as the two metrics for comparison, similar to conventional Si transistors. The best corner is on the top-left since low  $SS_{min}$  and high  $g_{m,max}$  are desired. Our SiO<sub>2</sub> devices, owing to the thick EOT, provide low transconductance even for the shortest  $L_{ch}$  devices. Scaling the EOT (12 nm HfO<sub>2</sub> and 4 nm HfO<sub>2</sub>) and using an optimized process flow (see Methods), we gain both in transconductance and SS, achieving a  $R_c < 2$  k $\Omega$   $\mu$ m for Ni contact metal and  $D_{it} < 5 \times 10^{12}$  cm<sup>-2</sup> for a CET of 1.9 nm. We demonstrate the highest



$g_{m,max} = 185 \mu S/\mu m$  at  $V_{DS} = 1 V$  and a minimum SS of 86 mV/dec for 4 nm  $HfO_2$ . We also achieve  $I_{max} = 400 \mu A/\mu m$  at  $V_{DS} = 1 V$  and  $V_{GS} = 4 V$  for our 12 nm  $HfO_2$  samples (Fig. S8).

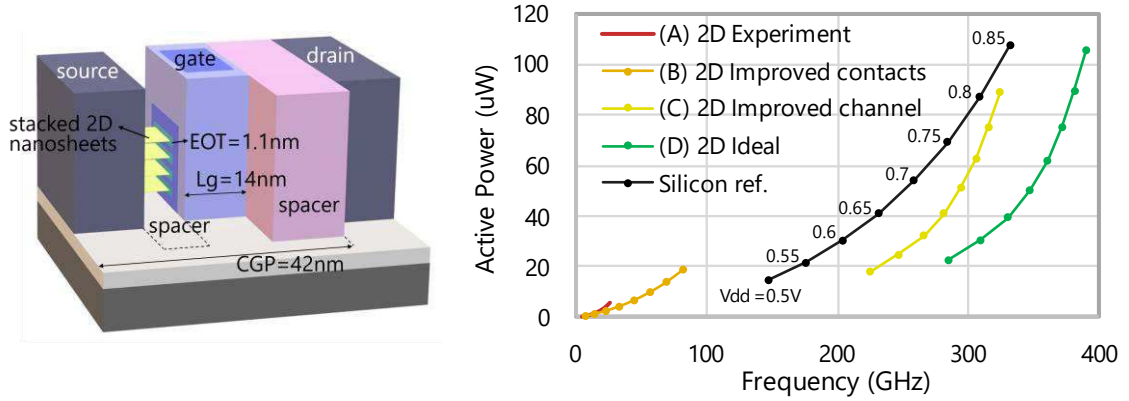


Figure 6: (a) Power performance area (PPA) analysis comparing Silicon and 2D in the same configuration of 4 stacked nanosheets with gate-all-around. (b) The baseline (A) is set with experimental values,  $R_c = 1.5 k\Omega \cdot \mu m$  (corresponding to  $\Phi_{SB} = 0.45 eV$ ),  $\mu = 15 cm^2/Vs$ ,  $D_{it} = 3 \times 10^{12} cm^{-2} eV^{-1}$ ,  $t_{channel} = 3$  layers. For (B) the contacts are improved with  $R_c \leq 50 \Omega \cdot \mu m$  (corresponding to  $\Phi_{SB} = 0.2 eV$ ). For (C) the channel is further improved with  $\mu = 200 cm^2/Vs$ ,  $D_{it} = 1 \times 10^{12} cm^{-2} eV^{-1}$ ,  $t_{channel} = 1$  layer. For (D), more aggressive improvements are done with  $\mu = 450 cm^2/Vs$  and no  $R_c$ . For all curves, the area is the same and the bias conditions are such that at  $V_{dd} = 0.7V$ ,  $I_{off}$  is fixed at 2nA. Methodology from [45].

Despite the fact our 2D performance is among the best in literature, significant improvements are still needed to make 2D materials competitive with Silicon channel devices for high-performance logic applications. Therefore we propose a roadmap using the Power Performance Area (PPA) metric for technology comparison in figure 6. 2D-FET and Silicon nanosheet technology are compared using an inverter-based ring oscillator circuit, where each device consists of 4 vertically stacked sheets with scaled  $L_g = 14 nm$  and gate-all-around structure, corresponding to the imec 2nm node [45]. All devices are retargeted to an  $I_{off} = 2 nA$  at 0.7V  $V_{dd}$  and the inverter-circuit area is kept the same for fairer comparison between technologies. Starting from the baseline case (A) where experimental channel and contact parameters are assumed, the performance strongly improves in (B) when the Schottky barrier height is reduced. In (C), improvements to the 2D channel mobility results in higher ring-oscillator operating frequency compared to Silicon, owing to superior electrostatic control of the 2D devices at shorter gate lengths. In (D), the ideal performance is simulated with more aggressively optimized material parameters.

In conclusion, we have scaled down the different device dimensions of CVD-grown MoS<sub>2</sub> FETs and demonstrated  $g_{m,max} = 185 \mu S/\mu m$  and  $SS_{min} = 86 \text{ mV/dec}$  which are among the best in literature. Using our large dataset, we systematically identified the key obstacles to be tackled to outperform Silicon. First, we showed that scaling  $L_{cont}$  for thin MoS<sub>2</sub> does not impact the short channel performance, which allows for an overall reduction in the device footprint. Second, we identified that for  $L_{ch} < 100 \text{ nm}$ , the on-current is currently limited by high Schottky contact resistance ( $R_c = 1\text{-}2 \text{ k}\Omega \cdot \mu m$ ) at low  $V_{ds}$ , and by a combination of velocity saturation and the Schottky barriers at high  $V_{ds}$ . Third, we identified that our devices suffer from short channel effects (SS degradation), caused by the Schottky barrier at intermediate current level and the thick CET at low current level. Reducing the CET is therefore crucial to keep optimal electrostatic control of the thin channel. We established that a 0.4 nm layer of water or adsorbed hydrocarbons (or combination thereof) at the HfO<sub>2</sub>/MoS<sub>2</sub> interface is the root cause of a lower-than-expected CET. This value is consistent across different thicknesses of HfO<sub>2</sub>. Therefore, an optimized transfer process free of water and carbon is needed to enable gate stack scaling below 1 nm, and additionally allow upscaling to 300 mm-wafer processing. Finally, we have demonstrated using a PPA analysis that if the obstacles of Schottky contacts, gate stack scaling and mobility improvement can be tackled, MoS<sub>2</sub> FETs will significantly outperform Silicon GAA FETs at the imec 2 nm node and beyond, and are therefore excellent candidates to continue logic scaling.

## V. METHODS

### A. Device fabrication

For the device design, we use the back-gate configuration with top-contacts (Fig 1b). The fabrication flow is summarized in Fig 1c. The MoS<sub>2</sub> is delaminated from the sapphire growth substrate using water intercalation and transferred to three different target substrates; (1) Si/50 nm SiO<sub>2</sub> (2) Si/50 nm SiO<sub>2</sub>/5 nm TiN/12 nm HfO<sub>2</sub>, or (3) Si/50 nm SiO<sub>2</sub>/5 nm TiN/4 nm HfO<sub>2</sub>. Before transfer, the target substrates are pre-cleaned using a solvent rinse, followed by an optimized forming gas anneal (FGA) or soft O<sub>2</sub> plasma, for SiO<sub>2</sub> and HfO<sub>2</sub> back-gate oxides, respectively. The active channels are patterned using PMMA mask and e-beam lithography, followed by reactive ion etching (Cl<sub>2</sub> + O<sub>2</sub>) of MoS<sub>2</sub>. Source and drain contacts of different lengths ( $L_{cont}$ ) with different channel lengths ( $L_{ch}$ ) are subsequently defined on the active channel by another e-beam

lithography exposure of ZEP520A-2 resist (ZEON Corp.), e-beam evaporation of 10nm Ni, and metal lift-off in anisole. We ensure a low vacuum pressure  $< 10^{-6}$  Torr while depositing the Ni contact metal. Finally, in a third e-beam lithography step, thicker Ni/Pd contact pads are lifted off.

## B. TCAD calibration

All simulations<sup>44</sup> are performed in Sentaurus Synopsys Device. The low-field mobility ( $\mu_{\text{eff}}$ ) is calibrated from an experimental TLM fit shown in Fig 3c and implemented under a constant mobility model. An estimate for  $D_{\text{it}}$  is obtained from multi-frequency CV measurements as discussed in section D. An acceptor trap distribution uniform over the entire bandgap is assumed with  $D_{\text{it}} = 3 \times 10^{12} \text{ cm}^{-2} \text{ eV}^{-1}$ . With  $\mu_{\text{eff}}$  and  $D_{\text{it}}$  fixed by experiments, the Schottky barrier height is fitted to median transfer characteristics of  $L_{\text{ch}} = 30 \text{ nm}$  devices which are predominantly contact-limited. For the Schottky injection, the non-local tunneling model based on the Wentzel-Kramers-Brillouin approach is used.

**Data availability:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Additional information:** The authors declare that there are no competing interests.

**Author contributions:** Goutham Arutchelvan and Quentin Smets conceived the experiments. Goutham Arutchelvan, with support from Quentin Smets and Surajit Sutar, performed the device fabrication, characterization and analysis. Goutham Arutchelvan and Devin Verreck performed the TCAD simulations. Zubair Ahmad performed the PPA analysis. C-V measurements and analysis were performed by Abhinav Gaur, with support from Dennis Lin. The CVD MoS<sub>2</sub> was grown by Benjamin Groven and e-beam litho support was provided by Julien Jussot. Prof. Marc Heyns, Inge Asselberghs and Iuliana Radu managed the overall activities and allocation of resources under the beyond CMOS program at imec, Belgium. The manuscript was written by Goutham Arutchelvan with contributions from Quentin Smets and Devin Verreck. The manuscript was revised by all co-authors.

**Corresponding author mail address:** goutham.arutchelvan@imec.be

## VI. References

1. Dennard, R. H. *et al.* Design of ion-implanted MOSFET's with very small physical dimensions. *IEEE J. Solid-State Circuits* **9**, 256–268 (1974).
2. Skotnicki, T., Hutchby, J. A., Tsu-Jae King, Wong, H.-. P. & Boeuf, F. The end of CMOS scaling: toward the introduction of new materials and structural changes to improve MOSFET performance. *IEEE Circuits Devices Mag.* **21**, 16–26 (2005).
3. Jin, S., Fischetti, M. V & Tang, T. Modeling of Surface-Roughness Scattering in Ultrathin-Body SOI MOSFETs. *IEEE Trans. Electron Devices* **54**, 2191–2203 (2007).
4. Chhowalla, M., Jena, D. & Zhang, H. Two-dimensional semiconductors for transistors. *Nat. Rev. Mater.* **1**, 16052 (2016).
5. Novoselov, K. S. *et al.* Two-dimensional atomic crystals. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10451–10453 (2005).
6. Szabo, A., Rhyner, R. & Luisier, M. Ab-initio simulations of MoS<sub>2</sub> transistors: From mobility calculation to device performance evaluation. in *2014 IEEE International Electron Devices Meeting* 30.4.1–30.4.4 (2014). doi:10.1109/IEDM.2014.7047142.
7. Majumdar, K., Hobbs, C. & Kirsch, P. D. Benchmarking Transition Metal Dichalcogenide MOSFET in the Ultimate Physical Scaling Limit. *IEEE Electron Device Lett.* **35**, 402–404 (2014).
8. Smithe, K. K. H., Suryavanshi, S. V, Muñoz Rojo, M., Tedjarati, A. D. & Pop, E. Low Variability in Synthetic Monolayer MoS<sub>2</sub> Devices. *ACS Nano* **11**, 8456–8463 (2017).
9. Kang, K. *et al.* High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity. *Nature* **520**, 656–660 (2015).
10. Smets, Q. *et al.* Ultra-scaled MOCVD MoS<sub>2</sub> MOSFETs with 42nm contact pitch and 250μA/μm drain current. in *2019 IEEE International Electron Devices Meeting (IEDM)* 23.2.1–23.2.4 (2019). doi:10.1109/IEDM19573.2019.8993650.
11. Liu, H., Neal, A. T. & Ye, P. D. Channel Length Scaling of MoS<sub>2</sub> MOSFETs. *ACS Nano* **6**, 8563–8569 (2012).
12. Verreck, D., Arutchelvan, G., Heyns, M. M. & Radu, I. P. Device and Circuit Level Gate Configuration Optimization for 2D Material Field-Effect Transistors. in *2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)* 1–4 (2019). doi:10.1109/SISPAD.2019.8870506.
13. Agarwal, T. *et al.* Benchmarking of monolithic 3D integrated MX<sub>2</sub> FETs with Si FinFETs. in *2017 IEEE International Electron Devices Meeting (IEDM)* 5.7.1–5.7.4 (2017). doi:10.1109/IEDM.2017.8268336.
14. Lee, Y., Fiore, S. & Luisier, M. Ab initio mobility of single-layer MoS<sub>2</sub> and WS<sub>2</sub>: comparison to experiments and impact on the device characteristics. in *2019 IEEE International Electron Devices*

Meeting (IEDM) 24.4.1-24.4.4 (2019). doi:10.1109/IEDM19573.2019.8993477.

15. Ma, N. & Jena, D. Charge Scattering and Mobility in Atomically Thin Semiconductors. *Phys. Rev. X* **4**, 11043 (2014).

16. Smithe, K. K. H., English, C. D., Suryavanshi, S. V & Pop, E. High-Field Transport and Velocity Saturation in Synthetic Monolayer MoS<sub>2</sub>. *Nano Lett.* **18**, 4516–4522 (2018).

17. Arutchelvan, G. *et al.* Transistors on two-dimensional semiconductors: contact resistance limited by the contact edges. in *2017 IEEE International Interconnect Technology Conference (IITC)* 1–3 (2017). doi:10.1109/IITC-AMC.2017.7968951.

18. Szabó, Á., Jain, A., Parzefall, M., Novotny, L. & Luisier, M. Electron Transport through Metal/MoS<sub>2</sub> Interfaces: Edge- or Area-Dependent Process? *Nano Lett.* **19**, 3641–3647 (2019).

19. Arutchelvan, G. *et al.* From the metal to the channel: a study of carrier injection through the metal/2D MoS<sub>2</sub> interface. *Nanoscale* **9**, 10869–10879 (2017).

20. English, C. D., Shine, G., Dorgan, V. E., Saraswat, K. C. & Pop, E. Improved Contacts to MoS<sub>2</sub> Transistors by Ultra-High Vacuum Metal Deposition. *Nano Lett.* **16**, 3824–3830 (2016).

21. Guo, Y. *et al.* Study on the Resistance Distribution at the Contact between Molybdenum Disulfide and Metals. *ACS Nano* **8**, 7771–7779 (2014).

22. de la Rosa, C. J. L. *et al.* Insight on the Characterization of MoS<sub>2</sub> Based Devices and Requirements for Logic Device Integration. *ECS J. Solid State Sci. Technol.* **5**, Q3072--Q3081 (2016).

23. Smithe, K. K. H., English, C. D., Suryavanshi, S. V & Pop, E. Intrinsic electrical transport and performance projections of synthetic monolayer MoS<sub>2</sub> devices. *2D Mater.* **4**, 11009 (2016).

24. Yuan, H. *et al.* Field effects of current crowding in metal-MoS<sub>2</sub> contacts. *Appl. Phys. Lett.* **108**, 103505 (2016).

25. Liu, H. *et al.* Statistical Study of Deep Submicron Dual-Gated Field-Effect Transistors on Monolayer Chemical Vapor Deposition Molybdenum Disulfide Films. *Nano Lett.* **13**, 2640–2646 (2013).

26. Cheng, Z. *et al.* New Observations in Contact Scaling for 2D FETs. in *2019 Device Research Conference (DRC)* 227–228 (2019). doi:10.1109/DRC46940.2019.9046480.

27. Wang, Y. *et al.* Van der Waals contacts between three-dimensional metals and two-dimensional semiconductors. *Nature* **568**, 70–74 (2019).

28. Nalin Mehta, A. *et al.* Grain-Boundary-Induced Strain and Distortion in Epitaxial Bilayer MoS<sub>2</sub> Lattice. *J. Phys. Chem. C* **124**, 6472–6478 (2020).

29. Zhou, W. *et al.* Intrinsic Structural Defects in Monolayer Molybdenum Disulfide. *Nano Lett.* **13**, 2615–2622 (2013).

30. Gaur, A. *et al.* Analysis of admittance measurements of MOS capacitors on CVD grown bilayer MoS<sub>2</sub>. *2D Mater.* **6**, 35035 (2019).

31. Kwon, J. *et al.* Thickness-dependent Schottky barrier height of MoS<sub>2</sub> field-effect transistors. *Nanoscale* **9**, 6151–6157 (2017).

32. Laturia, A., Van de Put, M. L. & Vandenberghe, W. G. Dielectric properties of hexagonal boron nitride

and transition metal dichalcogenides: from monolayer to bulk. *npj 2D Mater. Appl.* **2**, 6 (2018).

33. Penumatcha, A. V., Salazar, R. B. & Appenzeller, J. Analysing black phosphorus transistors using an analytic Schottky barrier MOSFET model. *Nat. Commun.* **6**, 8948 (2015).

34. Nourbakhsh, A. *et al.* Serially connected monolayer MoS<sub>2</sub> FETs with channel patterned by a 7.5 nm resolution directed self-assembly lithography. in *2016 IEEE Symposium on VLSI Technology* 1–2 (2016). doi:10.1109/VLSIT.2016.7573376.

35. Cheng, C. *et al.* First demonstration of 40-nm channel length top-gate WS<sub>2</sub> pFET using channel area-selective CVD growth directly on SiO<sub>x</sub>/Si substrate. in *2019 Symposium on VLSI Technology* T244–T245 (2019). doi:10.23919/VLSIT.2019.8776498.

36. Huyghebaert, C. *et al.* 2D materials: roadmap to CMOS integration. in *2018 IEEE International Electron Devices Meeting (IEDM)* 22.1.1–22.1.4 (2018). doi:10.1109/IEDM.2018.8614679.

37. Pang, C., Thakuria, N., Gupta, S. K. & Chen, Z. First Demonstration of WSe<sub>2</sub> Based CMOS-SRAM. in *2018 IEEE International Electron Devices Meeting (IEDM)* 22.2.1–22.2.4 (2018). doi:10.1109/IEDM.2018.8614572.

38. Kai-Shin Li *et al.* MoS<sub>2</sub> U-shape MOSFET with 10 nm channel length and poly-Si source/drain serving as seed for full wafer CVD MoS<sub>2</sub> availability. in *2016 IEEE Symposium on VLSI Technology* 1–2 (2016). doi:10.1109/VLSIT.2016.7573375.

39. Patel, K. A., Grady, R. W., Smithe, K. K. H., Pop, E. & Sordan, R. Ultra-scaled MoS<sub>2</sub> transistors and circuits fabricated without nanolithography. *2D Mater.* **7**, 15018 (2019).

40. Lingming Yang *et al.* High-performance MoS<sub>2</sub> field-effect transistors enabled by chloride doping: Record low contact resistance (0.5 kΩ·μm) and record high drain current (460 μA/μm). in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers* 1–2 (2014). doi:10.1109/VLSIT.2014.6894432.

41. Chen, M. *et al.* TMD FinFET with 4 nm thin body and back gate control for future low power technology. in *2015 IEEE International Electron Devices Meeting (IEDM)* 32.2.1–32.2.4 (2015). doi:10.1109/IEDM.2015.7409813.

42. Park, S. & Akinwande, D. First demonstration of high performance 2D monolayer transistors on paper substrates. in *2017 IEEE International Electron Devices Meeting (IEDM)* 5.2.1–5.2.4 (2017). doi:10.1109/IEDM.2017.8268331.

43. Chou, A. *et al.* High On-Current 2D nFET of 390 uA/um at V<sub>DS</sub> = 1V using Monolayer CVD MoS<sub>2</sub> without Intentional Doping. *Vlsi 2020* 10–11 (2020).

44. Verreck, D. *et al.* The Role of Nonidealities in the Scaling of MoS<sub>2</sub> FETs. *IEEE Trans. Electron Devices* **65**, 4635–4640 (2018).

45. Weckx, P. *et al.*, Novel forksheet device architecture as ultimate logic scaling device towards 2nm. *IEEE International Electron Devices Meeting (IEDM)* 36.5.1 (2019) doi: 10.1109/IEDM19573.2019.8993635

# Figures

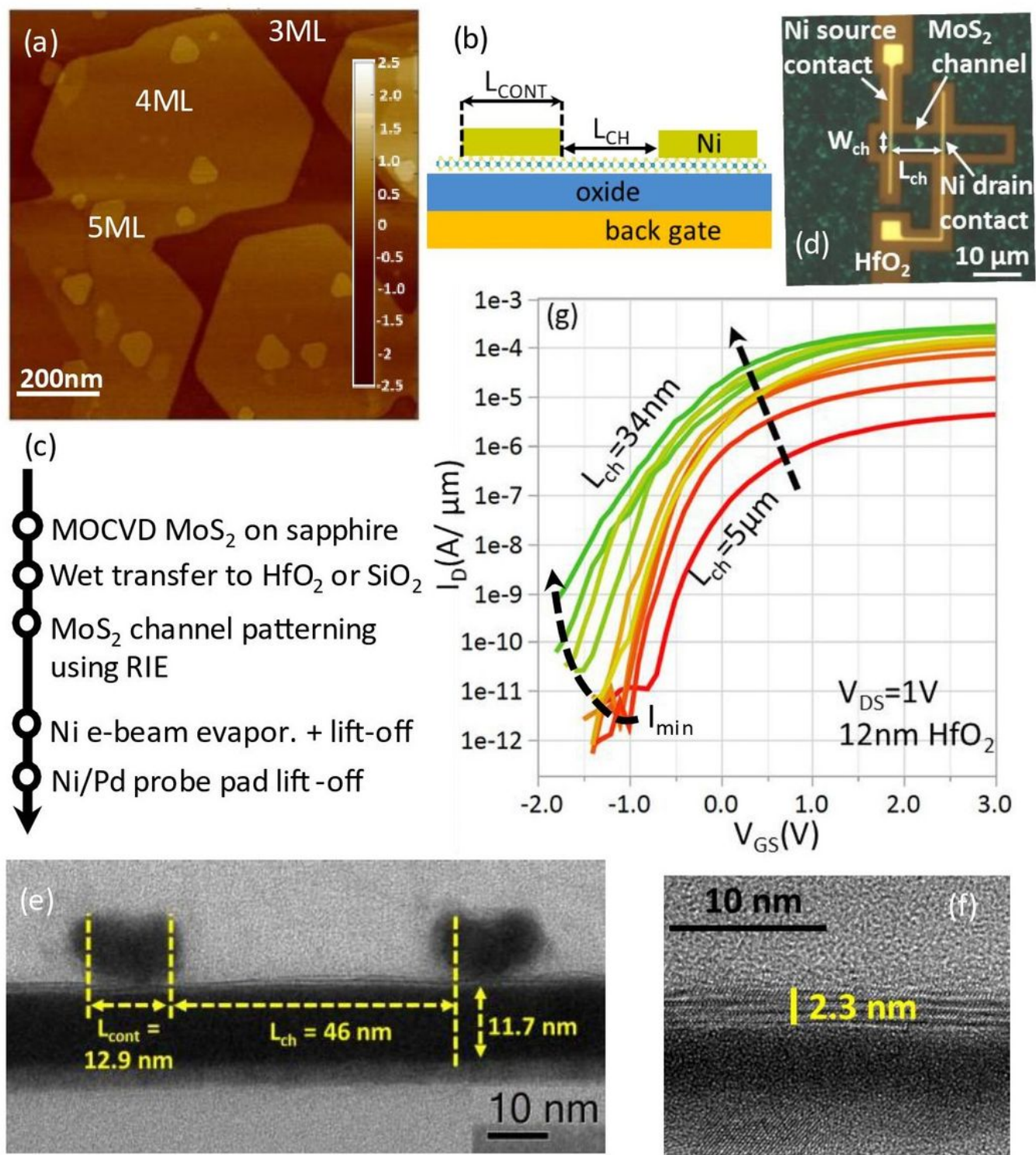
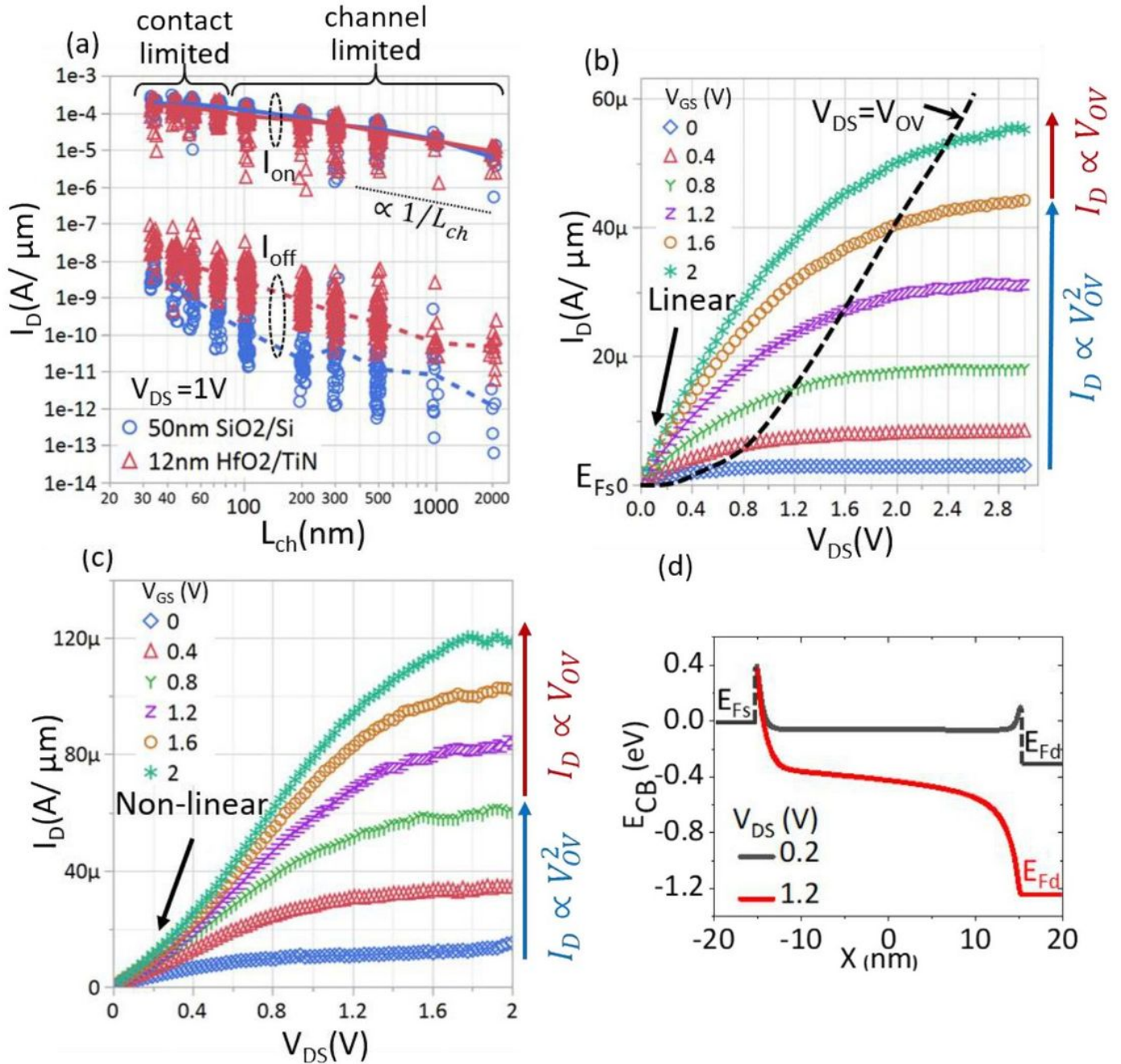


Figure 1

(a) Atomic force micrograph of CVD MoS<sub>2</sub> on sapphire template shows a closed 3ML layer with islands of 4 and 5 monolayers distributed randomly. (b) Device schematic with global back-gate and top source/drain contacts. (c) Fabrication flow for the back-gated devices. (d) Optical micrograph showing



the patterned MoS2 channel with 10 nm Ni contacts. (e) Cross-TEM shows a fabricated device with  $L_{\text{cont}} = 13$  nm and  $L_{\text{ch}} = 46$  nm on 12 nm HfO2. (f) Zoomin of the channel region for another device showing 3 monolayer MoS2 on nominal 4 nm HfO2 (g) Transfer characteristics at a fixed  $V_{\text{DS}} = 1$  V. Maximum drive current at  $V_{\text{GS}} = 3$  V scales with  $L_{\text{ch}}$  saturating for short-channel devices. The plot shows  $L_{\text{ch}} = 34$  nm, 44 nm, 50 nm, 70 nm, 100 nm, 200 nm, 300 nm, 500 nm, 1000 nm, 5000 nm.

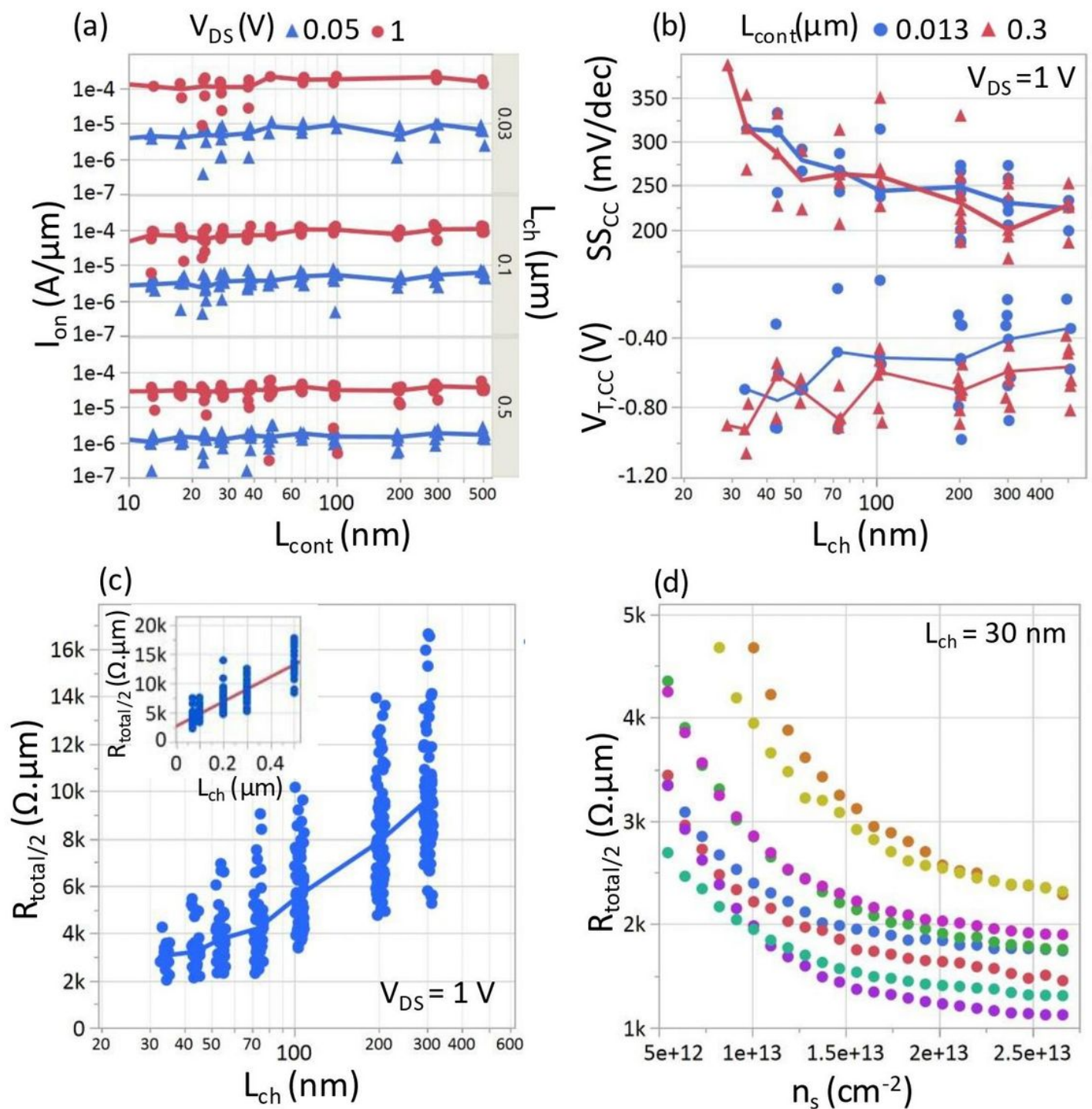


**Figure 2**

(a) Scatter plot (with median line) showing  $I_{\text{on}}$  extracted at  $n_s = 10^{13} \text{ cm}^{-2}$  and  $I_{\text{off}}$  at a fixed displacement field of 0.4 V/nm below  $V_{\text{T,CC}}$  for  $V_{\text{DS}} = 1$  V.  $I_{\text{on}}$  for the 50 nm SiO2 and 12 nm HfO2 devices overlap indicating no impact on low-field mobility and contact barrier.  $I_{\text{on}}$  roughly scales as



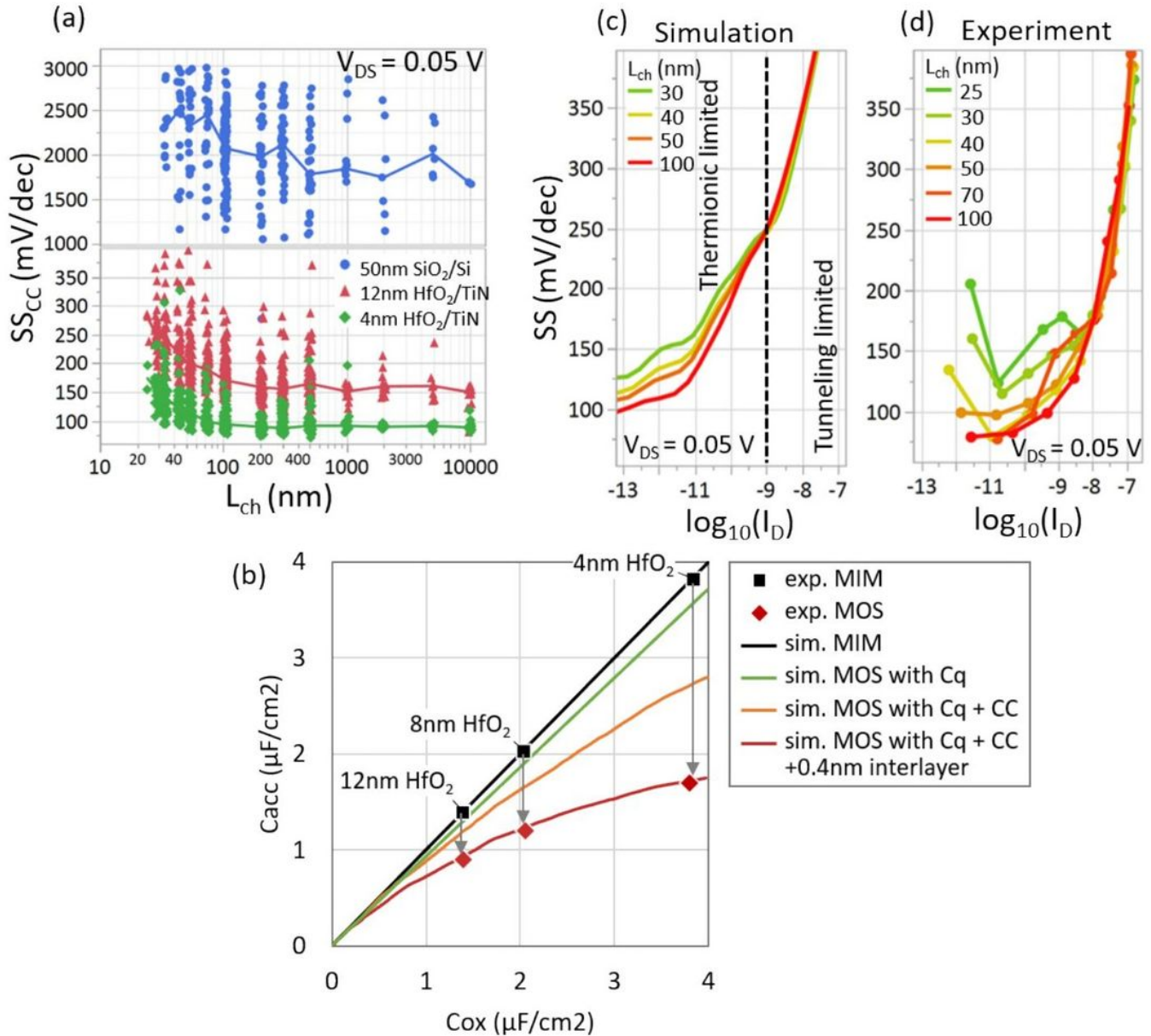
$1/L_{ch}$  for  $L_{ch} > 500$  nm and saturates for  $L_{ch} < 50$  nm.  $I_{off}$  is higher for  $HfO_2$  compared to  $SiO_2$  due to high interface trap density. (b)  $I_D$ - $V_{DS}$  for  $L_{ch} = 500$  nm shows linear triode regime and saturation at high  $V_{DS}$ . The dashed line follows the current at  $V_{DS} = V_{OV}$ . While the onset of saturation follows the  $V_{OV}$  at low  $V_{GS}$ , it saturates at  $V_{DS} = 2.4$  V for high  $V_{GS}$ . The saturation current roughly scales  $V_{OV}^2$  and  $V_{OV}^1$ , at low and high  $V_{GS}$ , respectively. (c)  $I_D$ - $V_{DS}$  for  $L_{ch} = 30$  nm shows non-linear triode regime due to Schottky contacts and saturation at high  $V_{DS}$ . Saturation current follows a similar trend as  $L_{ch} = 500$  nm but  $V_{DS}$  at onset of velocity saturation is reduced to 1.4 V. (d) Conduction band profile for  $L_{ch} = 30$  nm device with Schottky contacts shown for low and high  $V_{DS}$ . The Fermi-level at the source and drain are indicated by  $E_{FS}$  and  $E_{FD}$ , respectively. The Schottky barrier is shown as the abrupt potential change at the contact-channel interface. At low  $V_{DS}$ ,  $I_{DS}$  is determined by Schottky contacts. At high  $V_{DS}$ , though the potential drops significantly at the source contact, velocity saturation near the drain determines the  $I_D$  characteristics.



**Figure 3**

Scatter plot (with median line) of (a)  $I_{on}$  (at  $n_s = 1e13$   $cm^{-2}$ ) versus  $L_{cont}$  for  $L_{ch} = 30$  nm (contact-limited), 100 nm (intermediate regime), and 500 nm (mobility limited). No dependence on  $L_{cont}$  down to 13 nm indicates carrier injection from the edge of the metal directly into the MoS2 channel with  $LT < 13$  nm. (b)  $SS_{CC}$  and  $V_{T,CC}$  versus  $L_{ch}$  for  $L_{cont} = 13$  nm and 300 nm. No systematic deviation with  $L_{cont}$  indicates identical electrostatics in both cases. (c)  $R_{total/2}$  (at  $n_s = 1e13$   $cm^{-2}$ ) versus  $L_{ch}$  show saturation below  $L_{ch} = 50$  nm due to contact resistance. Upper limit for RC is obtained as median

$R_{\text{total}}/2$  for  $L_{\text{ch}} = 30$  nm. Median RC values of  $3 \text{ k}\Omega \cdot \mu\text{m}$  with best performers at  $2 \text{ k}\Omega \cdot \mu\text{m}$  are obtained. (inset) TLM fit of  $R_{\text{total}}/2$  (at  $n_s = 1 \times 10^{13} \text{ cm}^{-2}$ ) versus  $L_{\text{ch}}$  gives  $R_c = 2.7 \text{ k}\Omega \cdot \mu\text{m}$  and field-effect mobility =  $15 \text{ cm}^2/\text{V}\cdot\text{s}$  (d)  $R_{\text{total}}/2$  versus  $n_s$  for  $L_{\text{ch}} = 30$  nm at  $V_{\text{DS}} = 1$  V of 8 devices. RC significantly reduces at  $n_s = 2 \times 10^{13} \text{ cm}^{-2}$  due to better carrier injection into the accumulated channel.



**Figure 4**

(a) Scatter plot (with median line) of  $SS_{\text{cc}}$  versus  $L_{\text{ch}}$  for the three different oxides. While  $SS$  improves with lower EOT, the degradation and scatter for short channel devices are attributed to electrostatic potential fluctuations caused by non-uniform thickness of  $\text{MoS}_2$  in the contact and channel regions. (b) Experimentally measured maximum accumulation capacitance from MOScap ( $C_{\text{max}}$ ) versus MIMcap capacitance ( $C_{\text{ox}}$ ). Systematically, the  $C_{\text{max}}$  is lower than  $C_{\text{ox}}$  corresponding to an additional 1 nm CET over the measured EOT. Simulations show this is caused by the quantum capacitance  $C_q$  ( $\text{MoS}_2$  having

lower DOS than metal), the impact of the charge centroid (CC) further away in MOS than MIM, and additionally due to 0.4 nm of water or carbon residues stuck at the HfO<sub>2</sub>/MoS<sub>2</sub> interface during transfer. Qualitative comparison between (c) simulated and (d) experimental SS versus log (I<sub>D</sub>) for different L<sub>ch</sub>. The simulated SS is for a uniform 3 monolayers MoS<sub>2</sub> with SBH = 0.45 eV. Two transport regimes at the contacts– thermionic emission and tunneling through the SB are identified. In the thermionic regime, the relative increase of field in the channel from the source/drain Schottky contacts degrades gate control for short L<sub>ch</sub> devices. In the tunneling regime, the nearly equal tunneling lengths for the different L<sub>ch</sub> results in a similar but degraded SS compared to the thermionic regime.

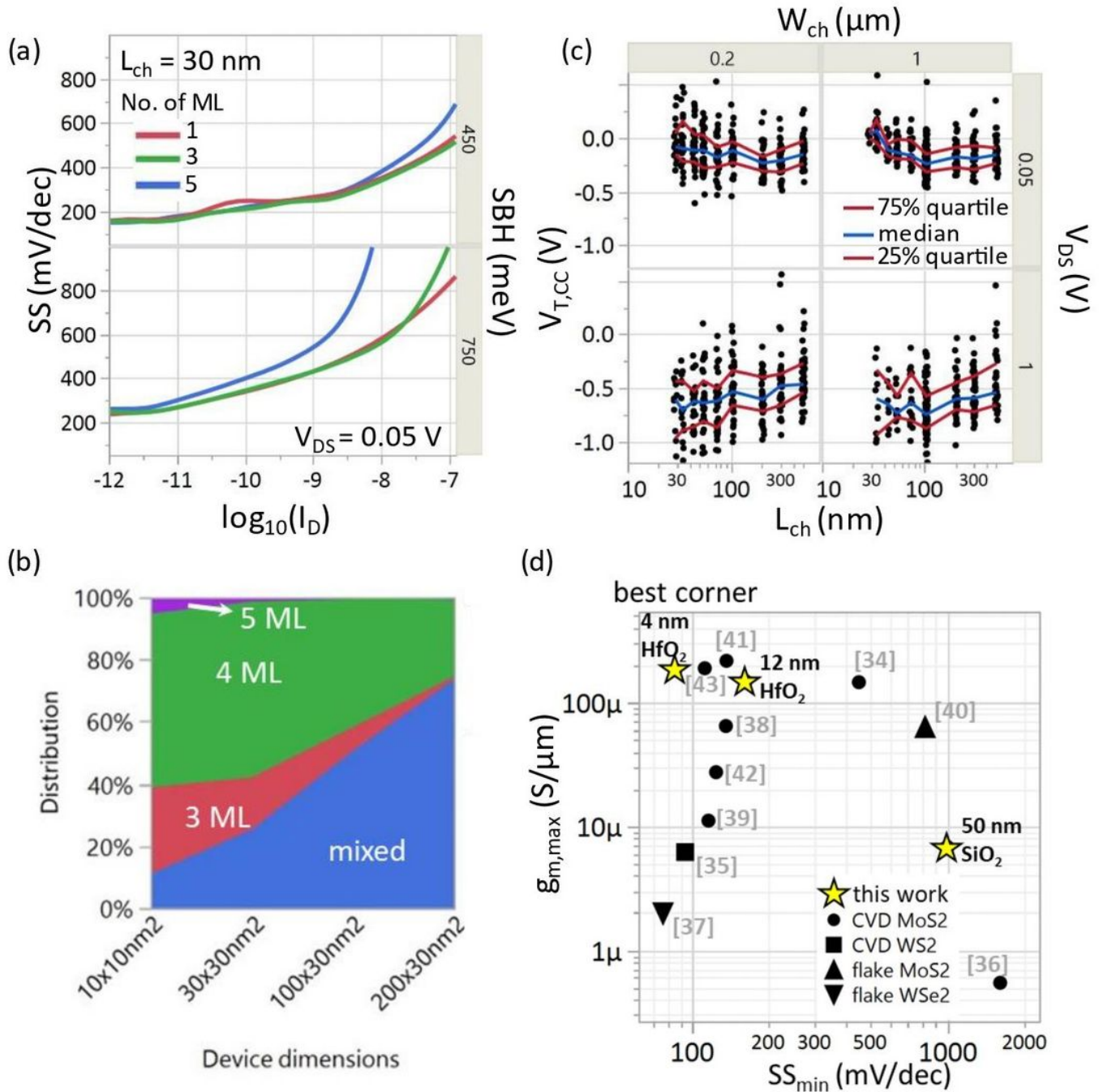
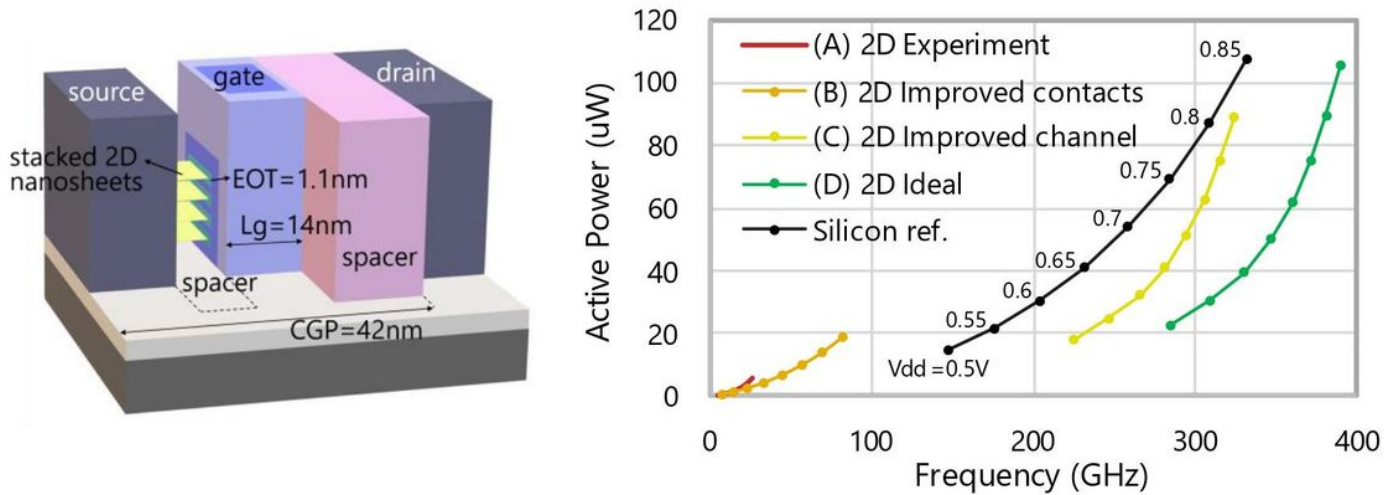


Figure 5



(a) Simulated SS versus log (ID) for a uniform layer of 1, 3 and 5 monolayers of MoS<sub>2</sub> for SBH = 0.45 eV and 0.75 eV. For ID > 1e-9 A/μm, tunneling through Schottky barrier determines the SS. Subsequently, a thinner channel results in better gate control, shorter tunneling length and therefore better SS. (b) Probability distribution versus device dimensions (Lch x Wch). AFM from Fig 1a was used to compute the probability distribution for fabricating devices on only 3, 4, 5 or a combination of those (mixed). Our experimental devices have a 60-70% probability of being mixed, leading to non-uniform gate control across the channel and contact regions. (c) VT,CC versus Lch for two different Wch (200 nm, 1000 nm) and VDS (0.05 V, 1 V). No VT roll-off at VDS = 0.05 V due to excellent gate control over the channel for 12 nm HfO<sub>2</sub>. VT roll-off of about 200 mV for VDS = 1 V due to higher lateral-field at the source contact allowing for more carrier injection. No systematic VT deviation between Wch = 1 μm and 200 nm. (c) Benchmark plot showing gm,max versus SSmin. All values are at VDS = 1 V except [39] - VDS = 0.1 V, [34] - VDS = 0.5 V, [40] - VDS = 1.2 V, [37] - VDS = 1.5 V. In this work, 4 nm HfO<sub>2</sub> provides best SS = 86 mV/dec and gm,max = 185 μS/μm.



**Figure 6**

(a) Power performance area (PPA) analysis comparing Silicon and 2D in the same configuration of 4 stacked nanosheets with gate-all-around. (b) The baseline (A) is set with experimental values,  $R_c = 1.5 \text{ k}\Omega\text{-}\mu\text{m}$  (corresponding to  $\Phi_{SB} = 0.45 \text{ eV}$ ),  $\mu = 15 \text{ cm}^2/\text{Vs}$ ,  $D_{it} = 3 \times 10^{12} \text{ cm}^{-2} \text{ eV}^{-1}$ ,  $t_{\text{channel}} = 3$  layers. For (B) the contacts are improved with  $R_c \leq 50 \Omega\text{-}\mu\text{m}$  (corresponding to  $\Phi_{SB} = 0.2 \text{ eV}$ ). For (C) the channel is further improved with  $\mu = 200 \text{ cm}^2/\text{Vs}$ ,  $D_{it} = 1 \times 10^{12} \text{ cm}^{-2} \text{ eV}^{-1}$ ,  $t_{\text{channel}} = 1$  layer. For (D), more aggressive improvements are done with  $\mu = 450 \text{ cm}^2/\text{Vs}$  and no  $R_c$ . For all curves, the area is the same and the bias conditions are such that at  $V_{dd} = 0.7 \text{ V}$ ,  $I_{off}$  is fixed at 2 nA. Methodology from [45].

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinfo.pdf](#)