

**Hardware Trojans based on Two-dimensional Memtransistor**

Journal:	<i>Nanoscale Horizons</i>
Manuscript ID	NH-COM-12-2022-000568.R1
Article Type:	Communication
Date Submitted by the Author:	24-Jan-2023
Complete List of Authors:	Wali, Akshay; Pennsylvania State University, Engineering Science and Mechanics Ravichandran, Harikrishnan; Pennsylvania State University, Engineering Science and Mechanics Das, Saptarshi; Pennsylvania State University, Engineering Science and Mechanics

# *Hardware Trojans based on Two-dimensional Memtransistor*

*Akshay Wali<sup>1</sup>, Harikrishnan Ravichandran<sup>2</sup>, and Saptarshi Das<sup>1,2,3,4,\*</sup>*

<sup>1</sup>*Electrical Engineering and Computer Science, Penn State University, University Park, PA 16802, USA*

<sup>2</sup>*Engineering Science and Mechanics, Penn State University, University Park, PA 16802, USA*

<sup>3</sup>*Materials Science and Engineering, Penn State University, University Park, PA 16802, USA*

Corresponding author email address: sud70@psu.edu

**Abstract:** Hardware trojans (HTs) have emerged as a major security threat for integrated circuits (ICs) owing to the involvement of untrustworthy actors in the globally distributed semiconductor supply chain. HTs are intentional malicious modifications, which remain undetectable through simple electrical measurements but can cause catastrophic failure in the functioning of the ICs in mission critical applications. In this article, we show how two-dimensional (2D) material based in-memory compute elements such as the memtransistors can be used as hardware trojans. We found that logic gates based on 2D memtransistors can be made to malfunction by exploiting the programming capability of 2D memtransistors. While, we use 2D memtransistors-based ICs as the testbed for our demonstration, the results are equally applicable to any state-of-the-art and emerging in-memory compute technologies.

**KEYWORDS:** *hardware trojans, in-memory compute, 2D materials, hardware security, memtransistors, logic gates*

**Introduction:**

The monumental success of nearly all present day computing and information technology platforms brought about by the von Neumann architecture used in integrated circuits (ICs) has been truly remarkable<sup>1</sup>. However, with the ever-increasing demands of data-intensive applications such as artificial intelligence, machine learning, signal processing, optimization, etc., the conventional von Neumann computing platform suffers from performance bottleneck due to the separation between processing and memory units, a phenomenon widely known as the “Memory Wall”<sup>2</sup>. Despite the introduction of improved architectures such as high bandwidth memory (HBM)<sup>3</sup>, hybrid memory cube (HMC)<sup>4</sup>, and high-parallel graphics processing units (GPUs)<sup>5</sup>, moving massive amounts of information between logic and memory results in significant latency and energy consumption<sup>6</sup>. The emergence of non-von-Neumann architecture, particularly in-memory computing mitigates this problem by allowing computational tasks to be performed within the confines of the computational memory<sup>7</sup> eliminating the aforementioned energy and latency limitations<sup>8,9</sup>. In this context, successful demonstrations of in-memory computing architectures implemented on programmable heterogeneous microprocessor<sup>10</sup>, static random access memory (SRAM)<sup>11</sup>, integrated optics<sup>12</sup>, and resistive switching devices<sup>13</sup> offer a new way forward.

However, as expected, newer technologies are subjected to newer challenges and given that the semiconductor supply chain is a globally distributed ecosystem with likely involvement of untrustworthy actors, it is worthwhile to study the susceptibility of in-memory compute technology to emerging security threats. In this context, hardware trojans (HTs) have been identified as one of the most pervasive security threats to ICs used in military, governments, and industrial

operations in recent years <sup>14, 15</sup>. HTs are undesirable malicious modifications at the device, circuit, or architecture level, which alter the intended IC functionalities. HTs can be inserted at various stages of design and manufacturing flow such as fabrication, testing, assembly, and packaging. Such modifications remain unknown to the original designer and can result in catastrophic consequences in critical applications such as defense, transportation and communications <sup>16, 17</sup>. In fact, as per a report, malicious HTs present in electronic systems recently resulted in severe military mishaps <sup>14, 18</sup>.

Unlike software based trojans (STs) which steal, erase or corrupt information and cause harm to the host by gaining access to the operating system <sup>19</sup>, HTs cannot be removed once inserted in the intellectual property (IP). While trojans exist in many forms, they are most commonly activated by a specific mechanism, called a trigger, and deliver a specific function, called a payload <sup>20</sup>. HTs can be small or large in size with respect to the overall circuit <sup>19</sup>. Although, the intention is to compromise the integrity, authenticity, and functionality of the underlying IC, HTs can also be designed with the sole intention of shortening the operational lifetime or even cause failure of the entire hardware system upon its activation. Trojans must be stealthy, undetectable, and inserted into rare circuit nodes so as to reduce the likelihood of its activation during normal operations. It must be emphasized that a HT is an intentional insertion by an adversary and not a flaw arising from design or manufacturing variability.

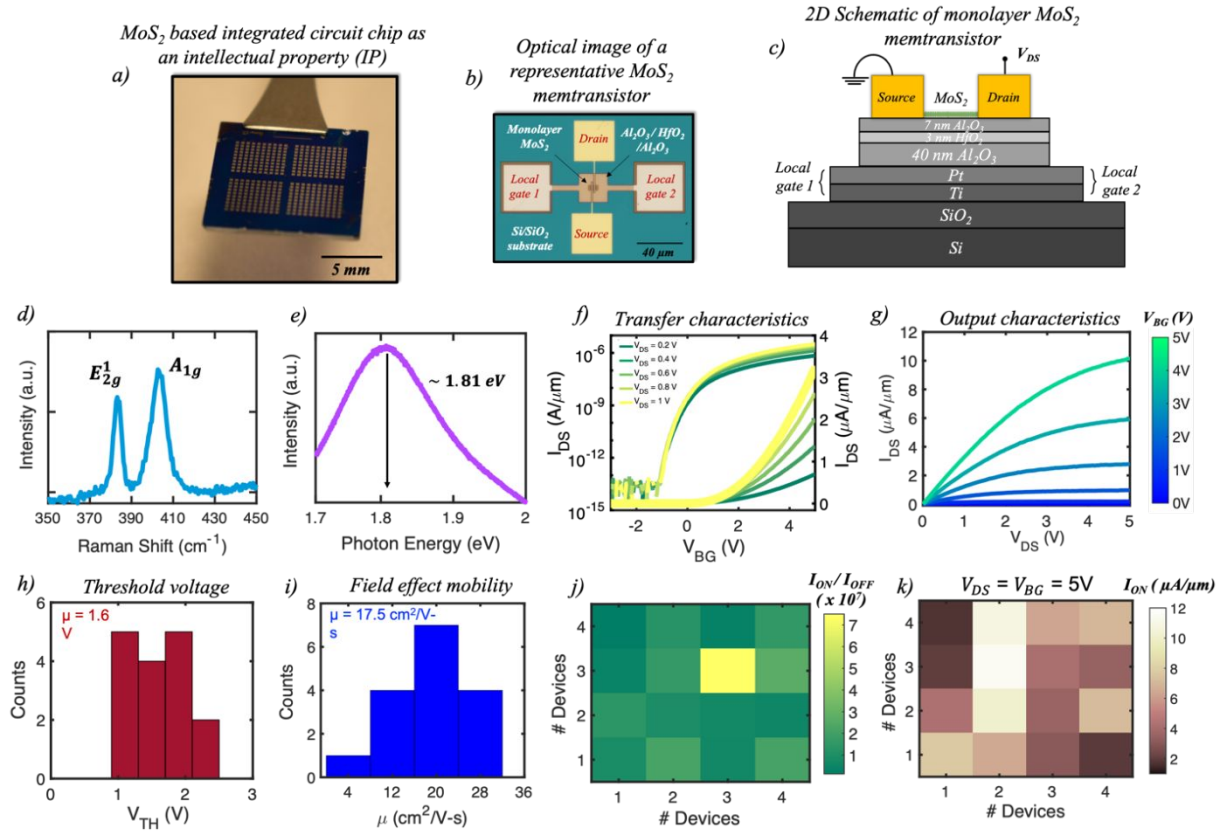
While majority of the research effort collimated towards studying HTs have focused on their detection, few have highlighted their deliberate inclusion in an IP <sup>21-24</sup>. In this article, we adopt a radical approach by placing ourselves in the role of an adversary who is knowledgeable about the

circuit designs and address a scenario where the HT is locally inserted within the IP core. We utilize the in-memory computing technique to engineer our hardware designs to deliberately produce faulty or corrupt circuit outputs, thereby compromising and degrading their overall circuit functionality. Note that contrary to earlier studies <sup>25-30</sup> which have focused on securing the IPs against variety of security threats such as reverse engineering (RE), piracy, counterfeiting, and illegal cloning, our demonstration aims to expose the hardware security vulnerabilities of the emerging in-memory compute paradigm. We have used logic circuits such as *NOT*, *AND*, *NAND*, *NOR* and a three-stage inverter (TSI) based on memtransistors built using atomically thin two-dimensional (2D) materials such as monolayer molybdenum disulfide (MoS<sub>2</sub>) as a testbed for HT insertion. However, the results and HT insertion strategies adopted in this work are equally applicable to any other state-of-the-art in-memory computing technologies such as CMOS-based static random access memories (SRAMs) <sup>31, 32</sup>, resistive RAM (RRAM) <sup>33, 34</sup> phase-change material (PCM) <sup>35, 36</sup>, spintronics <sup>37, 38</sup>, floating gate memories <sup>39, 40</sup>, ferroelectric field effect transistors (FETs) <sup>41, 42</sup>, and memristive arrays <sup>43, 44</sup>.

### **Monolayer MoS<sub>2</sub> characterization and device fabrication:**

Our choice of MoS<sub>2</sub> stems from the growing interest in 2D materials as a post-Si solution as indicated by IC manufacturing giants such as Intel and TSMC. It is also noteworthy that there are several recent reports on high-performance, high-density, and low power 2D device technologies <sup>45-52</sup> as well as 2D-based in-memory and in-sensor compute technologies <sup>53-59</sup>. In order to ensure material high quality, the monolayer MoS<sub>2</sub> used in our study was grown using a metal-organic chemical vapor deposition (MOCVD) technique on an epitaxial sapphire substrate at a temperature

of 950 °C<sup>51</sup>. The material growth was followed by fabrication of memtransistor based IC as the IP. **Fig. 1a** shows the optical image of as-fabricated MoS<sub>2</sub>-based IP which comprises of several



**Figure 1. Monolayer MoS<sub>2</sub> characterization and memtransistor fabrication:** a) An optical image of a MoS<sub>2</sub>-based integrated IP comprising of several logic circuits such as NAND, NOR, AND, and a three-stage inverter (TSI) module. b) an optical image along with a corresponding c) 2D schematic of a representative individual MoS<sub>2</sub>-memtransistor. The device structure comprises of Platinum/Titanium (Pt/Ti) local gates, a 7 nm Alumina (Al<sub>2</sub>O<sub>3</sub>) tunnelling layer, a 3 nm Hafnia (HfO<sub>2</sub>) charge-trapping layer, and a 40 nm Al<sub>2</sub>O<sub>3</sub> blocking layer as the gate dielectric stack with Ni/Au (40/30 nm) as the source and drain contact pads. All the memtransistors have a channel length (L) of 1 μm, a channel width (W) of 5 μm and two local gates from either side for better accessibility. d) Raman and e) photoluminescence (PL) spectra, respectively, obtained from a representative MoS<sub>2</sub> film where the characteristic in-plane E<sub>2g</sub><sup>1</sup> mode and out-of-plane A<sub>1g</sub> mode was observed at 385 cm<sup>-1</sup> and 403 cm<sup>-1</sup>, respectively with the PL peak at 1.83 eV. f) Transfer characteristics of a representative MoS<sub>2</sub> memtransistor in logarithmic and linear scale where the source to drain current (I<sub>DS</sub>) is plotted with respect to the local back-gate voltage (V<sub>BG</sub>) measured at different source to drain (V<sub>DS</sub>) biases. A clear n-type dominated carrier transport is observed. g) A plot of the output characteristics where I<sub>DS</sub> is plotted against V<sub>DS</sub> as a function of different V<sub>BG</sub> biases with excellent I<sub>ON</sub> values of 10 μA/μm for V<sub>DS</sub> = 5 V at a V<sub>BG</sub> = 5 V. Histogram plots for the extracted h) threshold voltage (V<sub>TH</sub>) at iso-current of 100 nA/μm with a mean of 1.6 V, i) peak electric field mobility (μ<sub>n</sub>) extracted using transconductance method with a mean of 17.5 cm<sup>2</sup> V<sup>-1</sup> s<sup>-1</sup> along with the colormap plots of j) I<sub>ON</sub>/I<sub>OFF</sub> ratio and k) I<sub>ON</sub> values at a V<sub>DS</sub> = V<sub>BG</sub> = 5V for a population of 16 MoS<sub>2</sub> memtransistors which further confirm the high quality of our channel material.

logic circuits such as *NAND*, *NOR*, *AND*, and a three-stage inverter (TSI) module. **Fig. 1b-c**, respectively show the optical image and the corresponding 2D schematic of a representative individual 2D-memtransistor with monolayer MoS<sub>2</sub> as the semiconducting channel, Pt/Ti as local back-gate electrodes (CG), 40 nm alumina (Al<sub>2</sub>O<sub>3</sub>) as the gate dielectric, a 3 nm hafnia (HfO<sub>2</sub>) as a charge trapping layer in a floating gate (FG) configuration, a 7 nm Al<sub>2</sub>O<sub>3</sub> as the tunnelling layer along with Ni/Au (40/30 nm) as the source and drain electrodes. All our memtransistors have a fixed channel length ( $L$ ) of 1  $\mu\text{m}$  and a channel width ( $W$ ) of 5  $\mu\text{m}$ . Details on the MoS<sub>2</sub> synthesis, film transfer, and memtransistor-based circuit fabrication can be found in the **Methods** section.

**Fig. 1d-e**, respectively show the Raman and the photoluminescence (PL) spectra obtained from a representative MoS<sub>2</sub> film where the characteristic in-plane  $E_{2g}^1$  mode and out-of-plane  $A_{1g}$  mode was observed at 385  $\text{cm}^{-1}$  and 403  $\text{cm}^{-1}$ , respectively, and the PL peak was found to be at 1.81 eV, consistent with monolayer material. **Fig. 1f** shows the transfer characteristics of a representative MoS<sub>2</sub> memtransistor in logarithmic and linear scale where the source to drain current ( $I_{DS}$ ) is plotted with respect to the local back-gate voltage ( $V_{BG}$ ) measured at different source to drain biases ( $V_{DS}$ ). The clear n-type dominated carrier transport observed due to the pinning of the metal Fermi level near the conduction band facilitating enhanced electron injection<sup>60, 61</sup>. In addition, an excellent ON/OFF current ratio of  $\sim 10^7$  and subthreshold slope (SS) of  $\sim 350$  mV/decade extracted for three orders of magnitude change in  $I_{DS}$  is also observed. The presence of interface traps can be accounted for the deviation of the SS from the ideal value of 60 mV/dec. Next, the output characteristics (**Fig. 1g**) where  $I_{DS}$  is plotted against  $V_{DS}$  as a function of different  $V_{BG}$  also shows excellent  $I_{ON}$  values of 10  $\mu\text{A}/\mu\text{m}$  for  $V_{DS} = 5$  V at a  $V_{BG} = 5$  V. Note that further enhancement in

the device characteristics can be accomplished through the optimization of metal contacts, scaling the oxide thickness and channel length<sup>62</sup>

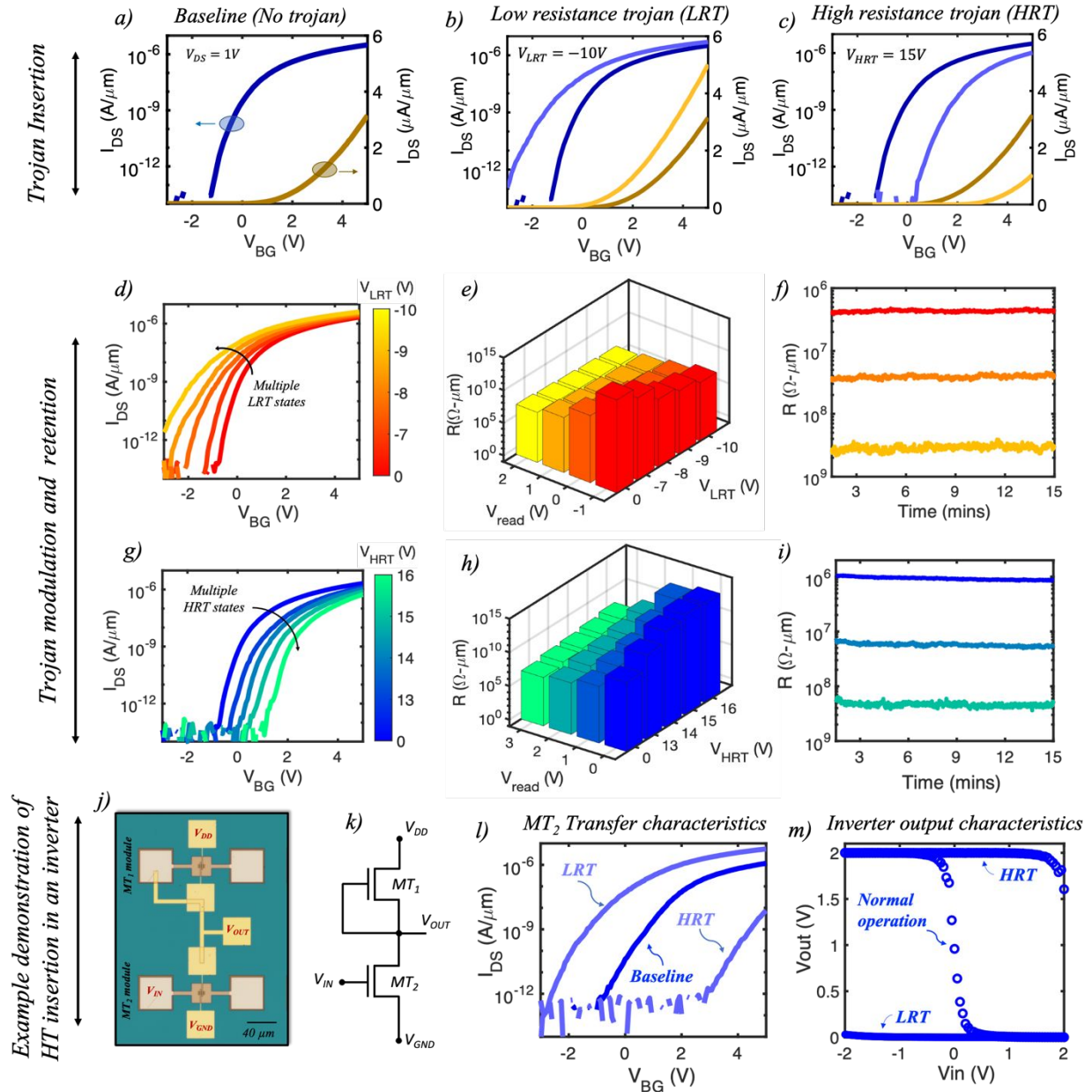
The device-to-device (D2D) variation was also examined as we intend to fabricate ICs involving multiple memtransistors. **Fig 1h-i** show the histogram plots for extracted threshold voltage ( $V_{TH}$ ) at an iso-current of  $100 \text{ nA } \mu\text{m}^{-1}$  and peak electric field mobility ( $\mu_n$ ) extracted from the transconductance ( $g_m = \frac{dI_{DS}}{dV_{BG}}$ ) method, respectively, with corresponding mean values of 1.5 V and  $17.5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for 16 different MoS<sub>2</sub> memtransistors. **Fig 1j-k** show the colormaps of on/off ratio, and  $I_{ON}$  measured at  $V_{DS} = 5 \text{ V}$  and  $V_{BG} = 5 \text{ V}$  for the same device. These results confirm the high quality and uniform growth of MoS<sub>2</sub>, which is critical for successful realization of memtransistor-based ICs. **Supporting Information S1** shows D2D variation with transfer characteristics and linear  $V_{TH}$  extraction for all the 16 MoS<sub>2</sub> memtransistors. Although, a deterrent, these variations are often unavoidable due to factors such as growth and transfer-related defects, adsorbed water molecules and resist residues during the fabrication process. However, they do not hinder or limit our ability as an adversary to insert HTs into the chip designs.

### **Basis of in-memory computing inspired hardware trojan in MoS<sub>2</sub> IP:**

In this section, we demonstrate the non-volatile electrical programmability in our 2D memtransistors as the basis of inserting the HT in the IP core. **Fig. 2a** shows the pre-trojan baseline transfer characteristics of a representative memtransistor in logarithmic and linear scale with expected carrier transport behavior. The trojan is inserted by applying a voltage pulse to the local gate of the MoS<sub>2</sub> memtransistor which results in an altered device characteristic. **Fig.2b-c**, respectively, show the transfer characteristics (logarithmic and linear scale) for the same



representative memtransistor post trojan insertion. It can be clearly seen that an application of a negative voltage pulse shifts the device threshold voltage ( $V_{TH}$ ) towards a more negative value thereby making it less resistive (more conductive) whereas a positive voltage pulse shifts the device  $V_{TH}$  towards a more positive value, making it more resistive (less conductive). The shift is attributed to the tunnelling of charge carriers across the 7nm  $Al_2O_3$  into the  $HfO_2$  charge-trapping layer. The ability to draw the carriers into the FG is enabled by the Fowler-Nordheim tunnelling<sup>63</sup> where a large positive and a negative voltage pulse applied to CG facilitates the accumulation of electrons and holes, respectively through the thin oxide layer (7nm  $Al_2O_3$ ) into the FG thereby screening the electric field which accounts for the observed  $V_{TH}$  shifts. **Supporting Information S2** shows the band diagram associated with the described tunnelling phenomenon. Note that similar charge trapping phenomenon with a different gate stack has also been extensively reported in our earlier reports<sup>30, 53, 54</sup>. While similar  $V_{TH}$  shifts are also observed due to the long-term drifts of the device characteristics as a result of bias-temperature instability (BTI)<sup>64</sup>, the slow and often prolonged nature of the process renders it infeasible for timely insertion of HT from an adversary's perspective. Due to their distinct effect on memtransistor characteristics, we have defined the negative voltage pulse as low resistance trojan voltage ( $V_{LRT}$ ) and the positive voltage pulse as high resistance trojan voltage ( $V_{HRT}$ ). We applied a -10V and a 15V for  $V_{LRT}$  and  $V_{HRT}$ , respectively with a pulse width ( $\tau_w$ ) of 100  $\mu s$  each, which remains constant for all subsequent discussions. Note that for our demonstration,  $V_{LRT}$  and  $V_{HRT}$  are defined as trigger mechanisms whereas the affected logic designs are defined as the payload. Note that similar resistive memories have also been demonstrated earlier for memristive<sup>65, 66</sup> and oxide nanosheets materials<sup>67</sup>.



**Figure 2. Basis of Non-von-Neumann based Hardware Trojan in monolayer  $\text{MoS}_2$  memtransistor:** a) Transfer characteristics of a representative memtransistor for three different scenarios: a) no trojan, b) post LRT and c) post HRT insertion. d) transfer characteristics of a representative  $\text{MoS}_2$  memtransistor for different  $V_{LRT}$  magnitudes ranging from -7V to -10V and their e) corresponding resistance values extracted at different  $V_{read} = V_{BG}$  values. f) Retention plot for three different LRT states which were found to be stable for 15 minutes. g) transfer characteristics of the same  $\text{MoS}_2$  memtransistor for different  $V_{HRT}$  magnitudes ranging from 13V to 16V and their h) corresponding resistance values extracted at different  $V_{read} = V_{BG}$ . i) Retention plot for three different HRT states which were also stable for 15 minutes. j) An optical image and a k) circuit diagram of an inverter, fabricated using a depletion-mode (D-mode) memtransistor ( $MT_1$ ) and an enhancement-mode (E-mode) memtransistor ( $MT_2$ ). l) Transfer characteristics of  $MT_2$ , measured pre-trojan (baseline), post  $V_{LRT} = -10V$  (LRT) and  $V_{HRT}$  of 15V (HRT) insertion. m) Inverter output characteristics showing normal operation along with post LRT and HRT operation with clear performance degradation.

It is also interesting to note that our MoS<sub>2</sub> memtransistors offer a unique capability of modulating both the low resistance trojan (LRT) and high resistance trojan (HRT) states as a function of the applied magnitude of  $V_{LRT}$  and  $V_{HRT}$ , respectively. Although, the overall aim is to corrupt or provide faulty logic at the output, the modulation becomes important from D2D variation point of view which is generally observed across a large population of 2D-memtransistors and can necessitate application of higher or even lower magnitudes of  $V_{LRT}$  and  $V_{HRT}$ . **Fig. 2d** shows the transfer characteristics of a representative MoS<sub>2</sub> memtransistor for different  $V_{LRT}$  magnitudes ranging from -7V to -10V and their corresponding resistance values (**Fig. 2e**) extracted at different  $V_{read} = V_{BG}$  values. Two observations can be made from these plots. First, for any given  $V_{LRT}$ , the values of LRT states decrease as  $V_{read}$  increases. This is expected since the  $V_{read}$  values gradually shift from the sub-threshold region towards the ON state. Second, the LRT values also decrease as a function of the magnitude of  $V_{LRT}$  at any given  $V_{read}$  value. This phenomenon can be attributed to the higher number of electrons tunnelling into the FG with increasing  $V_{LRT}$  which results in a higher negative  $V_{TH}$  shift. The LRT states were found to be very retentive (**Fig. 2f**) when measured for over 15 minutes. Similarly, **Fig. 2g-h**, respectively, show the transfer characteristics for different  $V_{HRT}$  magnitudes ranging from 13V to 16V and their corresponding resistance values (**Fig. 2h**) extracted at different  $V_{read} = V_{BG}$ . While following the similar trend as a function of  $V_{read}$  values, the HRT states show the converse effect with increasing  $V_{HRT}$  magnitudes. Once again, this is expected due to the higher positive  $V_{TH}$  shift with increasing  $V_{HRT}$  magnitudes which would suggest larger amount of electron tunnelling out of the FG. Note that this ability to electrostatically alter the channel conductance through application of the  $V_{BG}$ , which acts as the third terminal is what differentiates our memtransistor architecture from other similar prototypes such as metal-insulator-metal (MIM) or even memristors which are simple two terminal

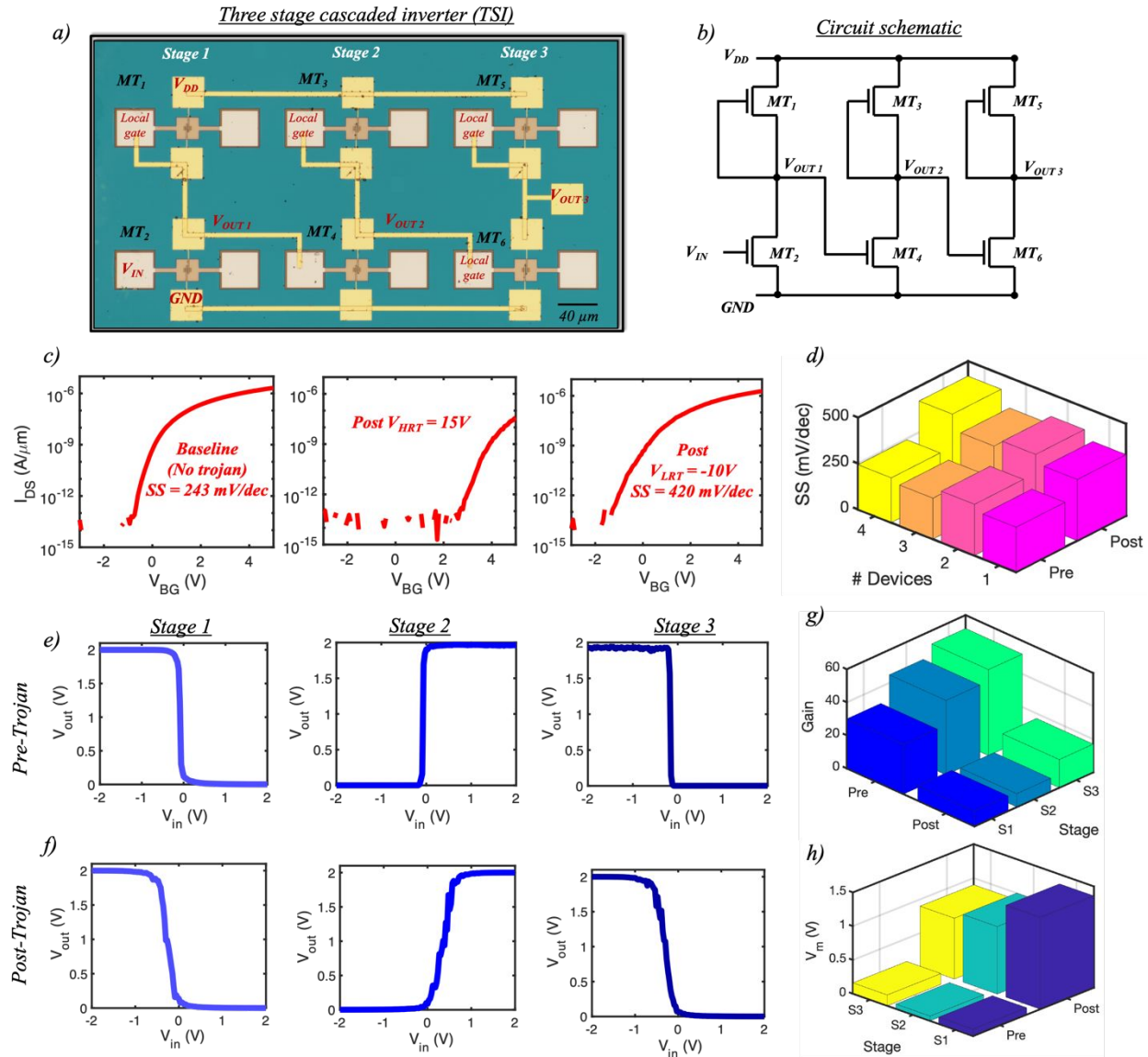
devices offering limited tunability and electrostatic control. Like the LRT states, the HRT states were also found to be retentive (**Fig. 2i**).

Next, as an example demonstration, we successfully insert a trojan in an inverter module and corrupt its functionality. An inverter is one of the most fundamental and well known logic circuits whose output reflects the opposite logic-level of its input. **Fig. 2j-k**, respectively, show an optical image and a circuit diagram of the inverter, fabricated using a depletion-mode (*D*-mode) memtransistor (*MT1*) and an enhancement-mode (*E*-mode) memtransistor (*MT2*). Note that the gate and source terminals of the *MT1* in our design are shorted in order for it to act as a n-type depletion load in the absence of a p-type memtransistor. **Fig. 2l** shows the transfer characteristics of *MT2*, measured as fabricated (baseline), post  $V_{LRT} = -10\text{V}$  (LRT) and  $V_{HRT}$  of 15V (HRT) insertion whereas **Fig. 2m** shows the inverter output characteristics considering all the three scenarios for a representative applied voltage of ( $V_{DD}$ ) of 2V. For the baseline state, a  $V_{in} = 0\text{V}$  (logic state ‘0’) applied to the gate of *MT2*, results in a  $V_{out} = V_{DD} = 2\text{V}$  since at a  $V_{BG} = 0\text{V}$ , *MT2* is highly resistive compared to *MT1* (logic state ‘1’). Conversely, a  $V_{in} = 2\text{V}$  (logic state ‘1’), clamps the  $V_{out}$  to 0V (logic state ‘0’) since *MT2* is now less resistive than *MT1* at a  $V_{BG} = 2\text{V}$ . However, post HT insertion, a faulty inverter behavior is observed. For both the LRT and the HRT states, respectively, the  $V_{out}$  always remains clamped to 0V and 2V with no apparent transition or switching point. This is because *MT2* now remains either always “ON” (LRT state) or always ‘OFF’ (HRT state) compared to *MT1*. Interestingly, despite the HRT state of *MT2*, the inverter exhibits normal operation for a  $V_{DD} = 5\text{V}$  (**Supporting Information S3**). Therefore, assuming its lifetime functioning at this higher supply voltage, the presence of HT remains hidden at a  $V_{DD} = 2\text{V}$  (rare node). It must be noted that for the LRT state,  $V_{out}$  always remains clamped to 0V regardless of

the  $V_{DD}$  value. We would also like to point out here that for this particular case and the subsequent discussions, the adversary attack model relies on accelerated aging of the IC. Since HT insertion techniques continue to evolve, there is no guarantee that every attack mechanism is addressed. Therefore, run-time monitors are often deployed in the ICs as the last line of defense with an aim to switch the IC operating conditions to a safe mode if a HT is detected<sup>68, 69</sup>. For our MoS<sub>2</sub> IC, this safe mode can only be enabled through higher  $V_{DD}$  operations which inevitably leads to higher power consumption and energy dissipation, both of which are detrimental for any circuit/chip operation in terms of its operational lifetime span.

#### **Demonstration of in-memory hardware trojan in a three-stage inverter (TSI) circuit:**

Continuing from the discussion above, here we exploit the degradation observed in the subthreshold slope (SS) of a 2D memtransistor in a cascaded three-stage inverter (TSI) circuit module as a form of HT. **Fig. 3a-b**, respectively show the optical image and the corresponding circuit schematic of our TSI design which consists of three inverters connected in series wherein the  $V_{out}$  from a single inverter is fed as the  $V_{in}$  to the next stage inverter. **Fig. 3c** shows the process flow of HT insertion which begins by first measuring the baseline transfer characteristics of the *MT2* from stage 1. Thereafter, the HT is inserted in two steps. First, a  $V_{HRT}$  of 15V is applied to one of the local gates of *MT2*, the effect of which can be seen through a large positive  $V_{TH}$  shift in the device characteristics. This is followed by an application of a large  $V_{LRT}$  of -12V in an attempt to bring the device back to its baseline state. Interestingly, the device never fully recovers, the evidence of which can be clearly seen through the SS degradation from 243 mV/dec to 420 mV/dec. In order to be more substantiative of this finding, we repeated the experiment with three



**Figure 3. Demonstration of in-memory hardware trojan in a cascaded three-stage inverter (TSI) circuit:** a) An optical image and the b) corresponding circuit schematic of our TSI design consisting of three inverters connected in series where the  $V_{out}$  from a single inverter is fed as the  $V_{in}$  to the next stage inverter. c) Evolution of degradation of subthreshold slope (SS) of MT2 through application of first, a  $V_{HRT}$  of 15V followed by a  $V_{LRT}$  of -12V to bring the device back to its baseline state. Both the voltage pulses are applied to one of the local gates of MT2 from stage 1. d) A 3D bar plot showing similar SS degradation for four separate 2D memtransistors which is attributed to a more permanent damage caused to the underlying  $Al_2O_3$  oxide due to repeated high magnitude voltage stressing of the devices. Inverter output characteristics plots for e) pre-HT and f) post-HT clearly showing the deterioration of the overall gain as a result of HT insertion which is further confirmed by g) analyzing the maximum gain as shown using the 3D bar plot. h) A 3D bar plot of the extracted  $V_m$  which clearly shows a significant positive  $V_m$  shift as compared to the original pre-HT state of the inverters.

different individual 2D memtransistors and analyzed their SS both before and after HT insertion with similar results as shown in the 3D bar plot in **Fig. 3d**. This degradation in SS can be attributed to a more permanent damage caused to the underlying  $\text{Al}_2\text{O}_3$  oxide due to repeated high magnitude voltage stressing of the devices. **Supporting Information S4** shows the transfer characteristics for the three memtransistors both pre- and post-HT insertion.

Next, **Fig. 3e-f**, respectively, show the inverter output characteristics plots for all the three stages of the cascaded TSI both before and after the HT insertion. The degradation in the gain is clearly visible from the plots. Note that the inverter gain which is defined as the slope of the inverter characteristics, i.e.,  $\frac{dV_{out}}{dV_{in}}$  prior to HT insertion increases with the number of stages. The overall reduction in the maximum gain for all the three stages post-trojan as shown using the 3D bar plot in **Fig. 3g** is attributed to the degraded SS of *MT2*. This is because a steeper SS would signify a more rapid change in corresponding current values as a function of the applied gate-bias which would in turn imply a more abrupt transition between the two logic levels, i.e.,  $V_{DD} = 2\text{V}$  and  $0\text{V}$ . In addition to the SS and the corresponding gain degradation, we also evaluated the change in the switching voltage ( $V_m$ ), which is defined as the magnitude of  $V_{in}$  for which  $V_{out}$  reaches  $V_{DD}/2$  for all the inverters by applying a single  $V_{HRT}$  of  $15\text{V}$  of  $\tau_w = 100\mu\text{s}$  to *MT2*. **Fig. 3h** shows the 3D bar plot of the extracted  $V_m$  which clearly shows a significant positive  $V_m$  shift as compared to the original pre-HT state of the inverters. **Supporting Information S5** shows the corresponding output inverter plots post  $V_{HRT}$  application on *MT2* for all the stages.

### Demonstration of in-memory hardware trojan in a NOR logic design:

**Fig. 4a-b**, respectively, show the optical image and the circuit schematic of a *NOR* gate fabricated using our 2D memtransistors. Note that in the optical image, our design is contained within the demarcated white dashed line. A *NOR* is the negation of the logical *OR* operation and is also a universal logic gate which operates as per the truth table shown in **Fig. 4c**. In our design, we have combined two *E*-mode memtransistors *MT1* and *MT2* with a *D*-mode memtransistor *MT3*. Prior to HT insertion, we first measure the baseline transfer characteristics of both *MT1* and *MT2* as shown in **Fig. 4d-e**, respectively. This was followed by measuring the output characteristics of the *NOR* gate under normal operating conditions for all the possible input combinations. As shown in **Fig. 4f**,  $V_{out} = V_{DD} = 2V$  (logic '1') only when both  $V_{in1}$  and  $V_{in2}$  are at logic '0'. This behavior can be attributed to the highly resistive state of both *MT1* and *MT2*, at a  $V_{BG} = 0V$ , which effectively

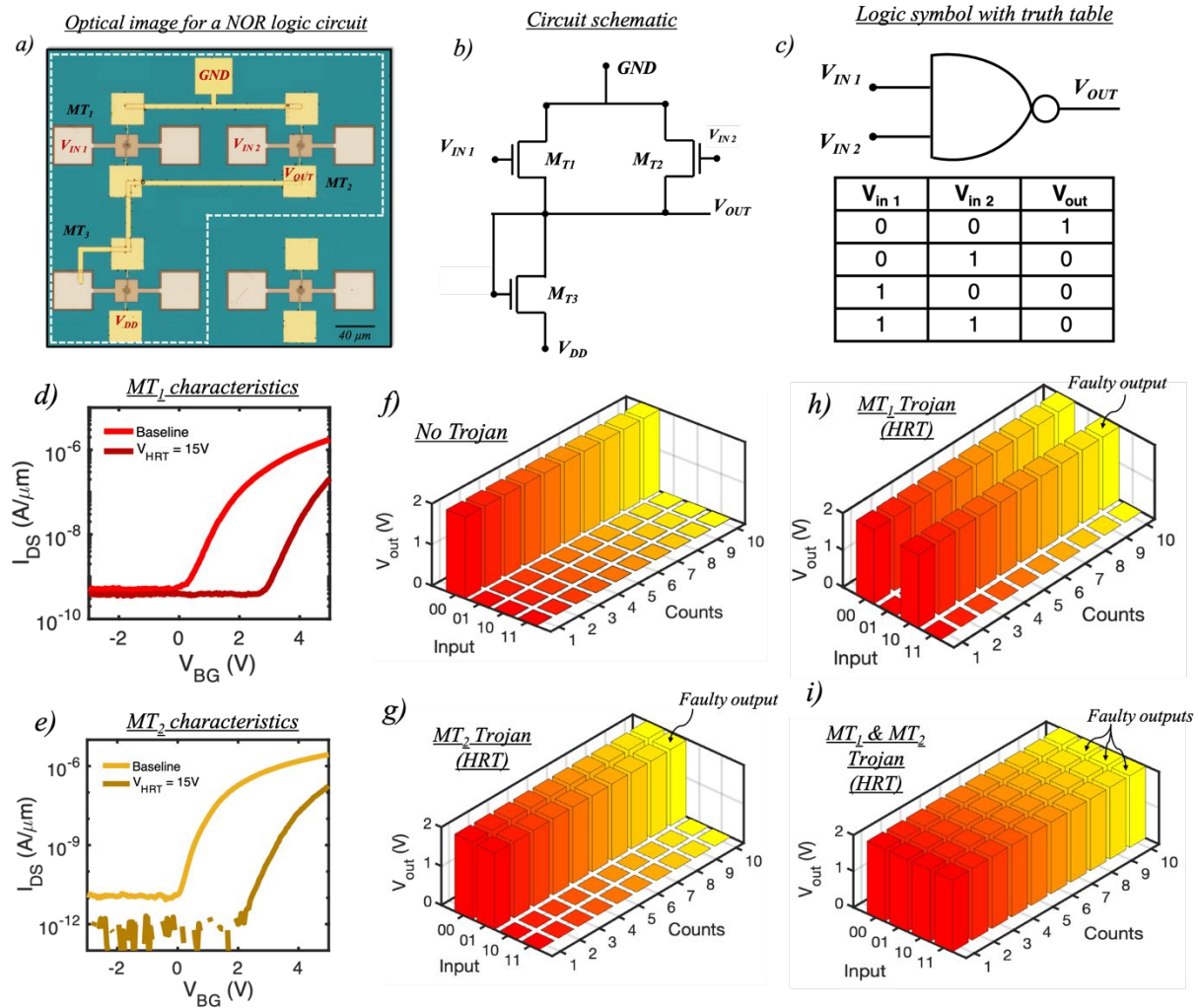
clamps the  $V_{out}$  to  $V_{DD}$  supply. For all the other input combinations,  $V_{out}$  always remains at logic '0', since at  $V_{BG} = 2V$ , either one or both *MT1* and *MT2*, remain at lower resistance state than *MT3*, thereby clamping  $V_{out}$  to  $V_{GND}$ . These results are in agreement with the truth table shown earlier.

Note that to ensure correct logical operations (without HT), the *E*-mode memtransistors must initially remain OFF at a  $V_{BG} = 0V$ . Therefore, any memtransistor which conducts at this  $V_{BG}$  as a result of the D2D variation must be biased properly through application of a programming pulse.

Next, we demonstrate our capability as an adversary to produce corrupt  $V_{out}$  values for specific combinations of  $V_{in1}$  and  $V_{in2}$ . For example, it is possible to engineer a faulty  $V_{out}$  value for the input combination '01' by simply inserting a HT in *MT2* by applying a  $V_{HRT}$  of 15V while keeping *MT1* in its baseline state (refer to **Fig. 4d-e**). Here, *MT2* remains in an HRT state for a  $V_{BG} = 2V$



and therefore, for an input combination ‘01’, both  $MT_1$  and  $MT_2$  now act as an open circuit path which ensures  $V_{out} = V_{DD} = 2V$  (logic ‘1’) as seen in **Fig. 4g**. Following similar logic, applying a

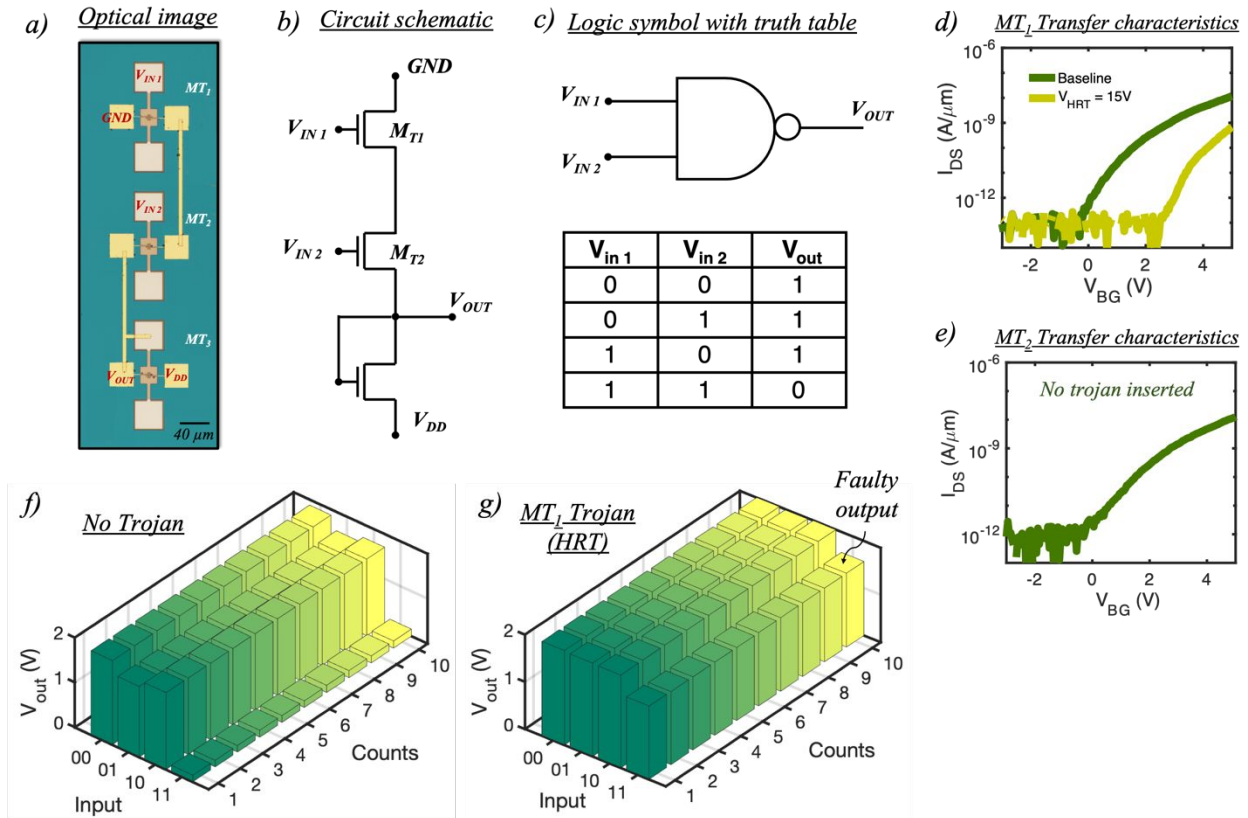


**Figure 4: Demonstration of in-memory hardware trojan in a NOR logic design:** a) An optical image and the b) circuit schematic of a NOR gate fabricated using two E-mode memtransistors  $MT_1$  and  $MT_2$  with a D-mode memtransistor  $MT_3$  along with its c) logic symbol and the truth table. Note that the NOR design is contained within the demarcated white dashed line in the optical image. Transfer characteristics of d)  $MT_1$  and e)  $MT_2$  memtransistors for both the baseline and post  $V_{HRT} = 15V$  HT insertion. Output characteristics of the NOR gate for all the possible input combinations under f) normal operating conditions which is in agreement with the truth table, g) post  $MT_2$  HT insertion resulting in a corrupt  $V_{out} = V_{DD} = 2V$  (logic ‘1’) value for specific input combination of ‘01’, h) post  $MT_1$  HT insertion resulting in a corrupt  $V_{out} = V_{DD} = 2V$  (logic ‘1’) value for specific input combination of ‘10’, and i) post  $MT_1$  and  $MT_2$  HT insertion resulting in a totally corrupt NOR logical operation where regardless of any input combinations,  $V_{out} = V_{DD} = 2V$ .

$V_{HRT}$  of 15V to  $MT1$  while keeping  $MT2$  in its baseline state ensures  $V_{out} = V_{DD} = 2V$  for the input combination '10' as seen in **Fig. 4h**. However, in this case,  $MT1$  remains in an HRT state for a  $V_{BG} = 2V$ . Finally, a  $V_{HRT}$  of 15V to both  $MT1$  and  $MT2$  results in a totally corrupt *NOR* logical operation where regardless of any input combinations,  $V_{out} = V_{DD} = 2V$  as shown in **Fig. 4i** which is due to the HRT state of both  $MT1$  and  $MT2$  at a  $V_{BG} = 2V$ . Once again, when the  $V_{DD}$  is scaled to 5V, *NOR* gate resumes its normal functionality as shown in **Supporting Information S6** while the trojan remains hidden at lower  $V_{DD}$  supply of 2V (rare node).

#### **Demonstration of in-memory hardware trojan in a NAND logic design:**

**Fig. 5a-b**, respectively, show the optical image and the circuit schematic of a *NAND* gate fabricated using two *E*-mode memtransistors  $MT1$  and  $MT2$  that are in series with one *D*-mode memtransistor  $MT3$ . Being a universal gate too, a *NAND* produces low output (logic '0') only when all of its inputs are high (logic '1'). Therefore, its output is complemented to that of an *AND* gate as shown in the truth table in **Fig. 5c**. Moving along similar lines, we first begin with measuring the baseline transfer characteristics of both  $MT1$  and  $MT2$  as shown in **Fig. 5d-e**, respectively, followed by the output characteristics (**Fig. 5f**) prior to HT insertion. Clearly, the expected *NAND* logic functionality is observed where  $V_{out} = V_{DD} = 2V$  (logic '1') for the input combinations of '00', '01', and '10' due to the high resistive state of both  $MT1$  and/or  $MT2$  compared to  $MT3$ . However, for the input combination '11',  $V_{out} = V_{GND} = 0V$  (logic '0') since now both  $MT1$  and  $MT2$  become more conductive than  $MT3$ . Note that is in line with the truth table shown earlier. The HT is then inserted by applying a  $V_{HRT}$  of 15V ( $\tau_w = 100\mu s$ ) to the local gate of  $MT1$  which results in the positive shift in the  $MT1$  device characteristics. On remeasuring



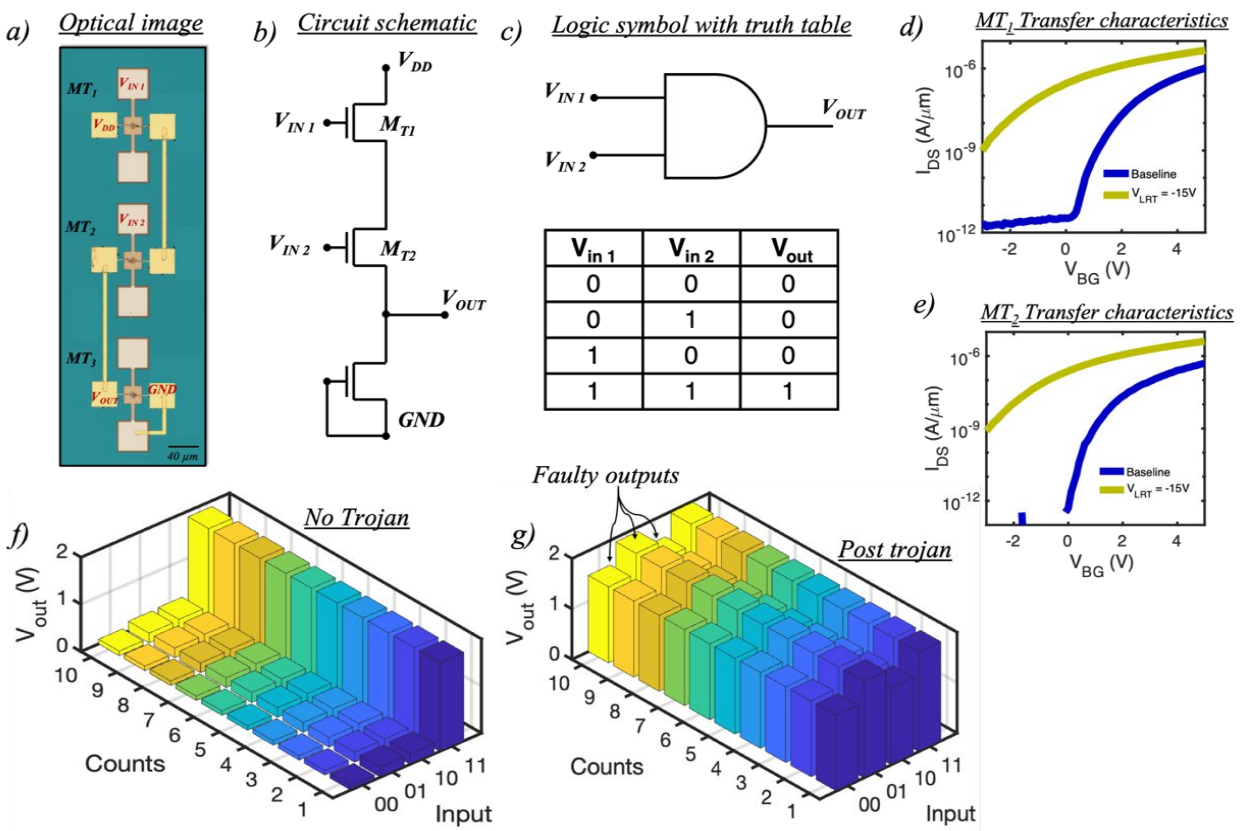
**Figure 5: Demonstration of in-memory hardware trojan in a NAND logic design:** a) An optical image and the b) circuit schematic of a NAND gate fabricated using two E-mode memtransistors  $MT_1$  and  $MT_2$  with a D-mode memtransistor  $MT_3$  along with its c) logic symbol and the truth table. Transfer characteristics of d)  $MT_1$  for both the baseline and post  $V_{HRT} = 15V$  HT insertion along with the baseline transfer characteristics of e)  $MT_2$  memtransistors. Note that for our NAND logic design, only one of the two memtransistors require insertion of HT. Output characteristics of the NAND gate for all the possible input combinations under f) normal operating conditions which is in agreement with the truth table and g) post  $MT_1$  HT insertion resulting in a corrupt  $V_{out} = V_{DD} = 2V$  (logic '1') value for specific input combination of '11'.

the NAND, we now see a faulty  $V_{out}$  for the input combination of '11'. This behavior is attributed to the HRT state of  $MT_1$  which will now always clamp the  $V_{out} = V_{DD} = 2V$ . Note that unlike the NOR logic, our NAND logic design requires HT insertion in only one of the two E-mode memtransistors ( $MT_1$  or  $MT_2$ ). Once again, despite the presence of HT at the rare node ( $V_{DD} = 2V$ ), normal NAND functionality is observed for a  $V_{DD} = 5V$  as shown in **Supporting Information**

S7

### Demonstration of in-memory hardware trojan in a AND logic design:

Finally, **Fig. 6a-b**, respectively, show the optical image and the circuit schematic of a *AND* gate with a similar memtransistor configuration as that of a *NAND* gate except for the switching of the  $V_{DD}$  and  $V_{GND}$  polarities along with the shorting of the gate terminal of  $MT_3$  to  $V_{GND}$ . The *AND* logic can be summarized using the truth table shown in **Fig. 6c** where  $V_{out}$  remains high (logic '1') only when both  $V_{in1}$  and  $V_{in2}$  are high. **Fig. 6d-e**, respectively, show the baseline transfer characteristics of  $MT_1$  and  $MT_2$ . Prior to HT insertion, a normal *AND* logic functionality is

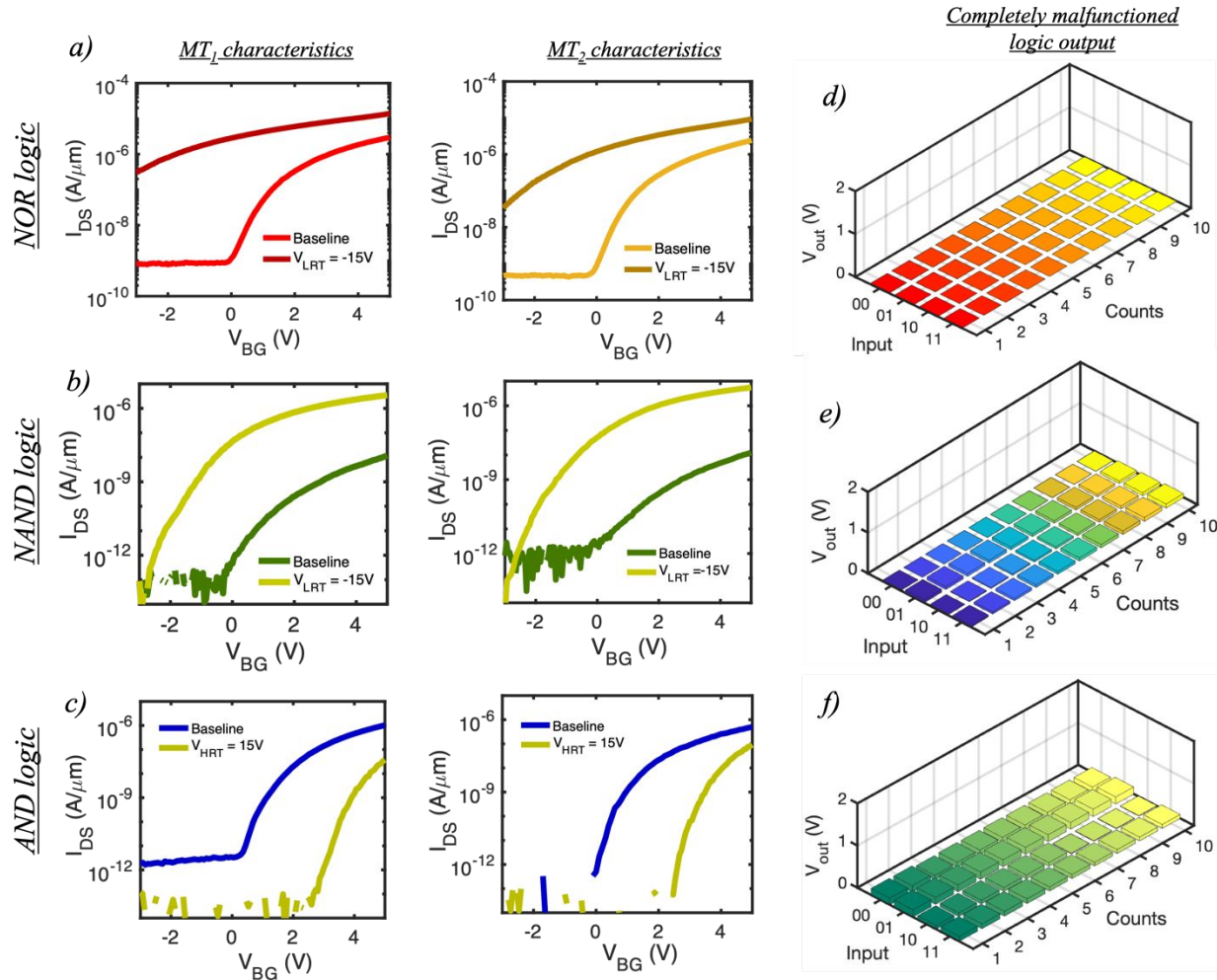


**Figure 6: Demonstration of in-memory hardware trojan in a AND logic design:** a) An optical image and the b) circuit schematic of a *NAND* gate fabricated using two *E*-mode memtransistors  $MT_1$  and  $MT_2$  with a *D*-mode memtransistor  $MT_3$  along with its c) logic symbol and truth table. Transfer characteristics of d)  $MT_1$  and e)  $MT_2$  memtransistors for both the baseline and post  $V_{LRT} = -15V$  HT insertion. Note that for our *AND* logic design, both memtransistors require insertion of HT. Output characteristics of the *AND* gate for all the possible input combinations under f) normal operating conditions where for the input combinations of '00', '01', or '10',  $V_{out} = V_{GND} = 0V$  (logic '0') since either one or both  $MT_1$  and  $MT_2$  are more resistive than  $MT_3$  and g) post  $MT_1$  and  $MT_2$  HT insertion where regardless of the input combination,  $V_{out} = V_{DD} = 2V$  (logic '1') since both  $MT_1$  and  $MT_2$  are now in *LRT* state offer more conductive path compared to  $MT_3$ .

observed as shown in **Fig. 6f** where for the input combinations of ‘00’, ‘01’, or ‘10’,  $V_{out} = V_{GND} = 0V$  (logic ‘0’) since either one or both  $MT1$  and  $MT2$  are more resistive than  $MT3$ . However, for the input ‘11’  $MT1$  and  $MT2$  become more conductive than  $MT3$ , which clamps  $V_{out} = V_{DD} = 2V$  (logic ‘1’). Next, the HT is inserted by applying a  $V_{LRT} = -15V$  ( $\tau_w = 100\mu s$ ) to both  $MT1$  and  $MT2$  resulting in a negative  $V_{TH}$  shift as seen in **Fig. 6d-e**. The effect of inserting this HT can be understood by once again observing the output characteristics as shown in **Fig. 6g** where now regardless of the input combination,  $V_{out} = V_{DD} = 2V$  (logic ‘1’) since both  $MT1$  and  $MT2$  are now in LRT state offer more conductive path compared to  $MT3$ .

#### **Enabling complete failure of logic designs with in-memory HT:**

In addition to corrupting the circuit outputs for a particular set of input combinations, it is also possible for an adversary to engineer a complete failure or malfunction of the logic designs using the described in-memory HT insertion technique. In this context, we define a complete logic failure as a scenario where  $V_{out}$  always remains low (logic ‘0’), regardless of the input combination. In other words, the chip now becomes completely non-functional. **Fig. 7a-c**, respectively show the transfer characteristics of  $MT1$  and  $MT2$  measured for the *NOR*, *NAND*, and *AND* logic designs for both baseline and post-HT configurations. Note that in each case, the type of HT inserted ( $V_{LRT}$  or  $V_{HRT}$ ), would be different. For example, in order to make the *NOR* and *NAND* logic malfunction, a  $V_{LRT}$  of  $-15V$  ( $\tau_w = 100\mu s$ ) is applied to both  $MT1$  and  $MT2$  whereas the *AND* logic malfunctioning requires an application of a  $V_{HRT}$  of  $15V$  ( $\tau_w = 100\mu s$ ). On observing their respective output characteristics as shown in **Fig. 7d-f**, it is evidently clear that each logic design has malfunctioned since  $V_{out} = V_{GND} = 0V$  (logic ‘0’) for all input combinations. In case of the



**Figure 7: Enabling complete failure of logic designs with in-memory HT:** Transfer characteristics of  $MT_1$  and  $MT_2$  measured for the  $MT_1$  and  $MT_2$  measured for the a) NOR, b) NAND, and c) AND logic designs for both baseline and post-HT configurations. d-f) Malfunctioned output characteristics for each respective logic design since  $V_{out} = V_{GND} = 0V$  (logic '0') regardless of the input combinations. In case of the NOR and NAND logic design, this can be attributed to the LRT state of both  $MT_1$  and  $MT_2$  whereas for the AND logic design, both  $MT_1$  and  $MT_2$  are in the HRT state compared to  $MT_3$ . It must be noted the type of HT inserted ( $V_{LRT}$  or  $V_{HRT}$ ) is different for each individual design. For example, in order to make the NOR and NAND logic malfunction, a  $V_{LRT}$  of  $-15V$  ( $\tau_w = 100\mu s$ ) is applied to both  $MT_1$  and  $MT_2$  whereas the AND logic malfunctioning requires an application of a  $V_{HRT}$  of  $-15V$  ( $\tau_w = 100\mu s$ ).

NOR and NAND logic design, this can be attributed to the LRT state of both  $MT_1$  and  $MT_2$  whereas for the AND logic design, both  $MT_1$  and  $MT_2$  are in the HRT state compared to  $MT_3$

### Trojan lifetime and detection countermeasures avoidance:

From the above discussions, it is abundantly clear that from an adversary's perspective, the ultimate goal is to enable a HT attack mechanism which is precise and stealthy with small footprint such that it is capable of evading detection countermeasures<sup>70</sup>. For our in-memory HT demonstration, we have achieved all of the aforementioned requirements. First, despite the  $V_{TH}$  shifts observed as a result of  $V_{LRT}$  or  $V_{HRT}$  application, the presence of HT is obscured since the electrical measurements of the affected individual 2D memtransistor do not offer any additional information. Although, it can be argued that device performance parameters such as SS and  $I_{ON}$  values degrade slightly, they are more likely to be attributed to fabrication constraints and D2D variation related factors. Furthermore, our HT injection scheme also fulfills the criteria of having a small footprint since HT can be triggered ( $V_{LRT}$  and  $V_{HRT}$  application) without requiring any additional peripheral components within the design.

Additionally, from our point of view as an adversary, a good retention of both LRT and HRT states is desirable for delivering the payload. Therefore, we examined the extended retention of our memtransistors for two different states for a total of  $10^4$  seconds or  $\sim 3$  hours (**Supporting Information S8**). Although, the two memory states will eventually converge after  $\sim 15$  hours, as evident through the slow collapsing of the memory window (MR) due to an exponential decay, an adversary can employ this attribute to cause catastrophic malfunctions and failures in critical technological infrastructure such as military equipment and aircraft navigation and still evade detection since the presence of HT can only be confirmed if the black box retrieved from the failed equipment is analyzed within the MR retention time frame. As soon as the MR window vanishes and the HT disappears, no evidence of a trojan being ever inserted can be recovered. Note that this decay might seem counterintuitive as far as deliberately aging or causing a total logic malfunction

of the IC is concerned. This is because both of these adversarial intentions would ideally require the HT to be present within an IC at all times. However, in the present day's era of stiff market competition and time-to-market demands requiring increased productivity and shorter design time, any degradation or malfunction observed within this global supply chain will likely result in the affected IP being discarded completely, leading to severe monetary loss in addition to reputational dissolvment for the involved parties.

### **Conclusion:**

In conclusion, we have exploited and exposed the vulnerability of the in-memory computing architecture to insert hardware trojans (HTs) within basic digital logic designs such as *AND*, *NAND*, *NOR* and *NOT* gates fabricated using monolayer MoS<sub>2</sub> memtransistors. In addition, we have also exploited the degradation in subthreshold slope (SS) due to presence of HT to drastically compromise the performance of a cascaded three stage inverter (TSI). The HT is inserted by applying an electrical voltage pulse ( $V_{HRT}$  or  $V_{LRT}$ ) to the local gate of our memtransistor devices, programming them into either HRT or LRT states which significantly corrupts and even causes complete logic malfunction of our circuit. This ability to enable both the LRT and HRT states as a function of the applied magnitude of  $V_{LRT}$  and  $V_{HRT}$  can be attributed to the tunnelling of charge carriers across the 7nm Al<sub>2</sub>O<sub>3</sub> into the HfO<sub>2</sub> charge-trapping layer due to Fowler Nordheim tunneling process. Moreover, the trojan remains hidden within the rare node during the normal operation of the logic designs at a higher  $V_{DD} = 5V$  and only gets activated when the supply voltage is scaled down ( $V_{DD} = 2V$ ). These results are not just limited to 2D-based technologies but are applicable to any state-of-the-art in-memory computing architecture. Our approach in the role of an adversary offers a rather unconventional, uncomfortable and an eye-opening insight into the



vulnerabilities of in-memory computing in today's highly complex and globalized IP supply chain system which can be exploited in unimaginable ways to compromise the security, reliability and even the lifetime of electronic components.

## Methods

Large area monolayer MoS<sub>2</sub> film growth: Metal-organic chemical vapor deposition (MOCVD) technique was used to deposit monolayer MoS<sub>2</sub> on an epi-ready 2” c-sapphire substrate. A cold-wall horizontal reactor with an inductively heated graphite susceptor and a wafer rotation was utilized to assist in achieving a uniform monolayer deposition<sup>71</sup>. Molybdenum hexacarbonyl (Mo (CO)<sub>6</sub>) and hydrogen sulfide (H<sub>2</sub>S) were used as precursors. Prior to the material growth, the substrate was heated to 1000 °C in H<sub>2</sub> and maintained for 10 mins. (Mo (CO)<sub>6</sub>). For the growth process, Mo (CO)<sub>6</sub> was maintained at 10°C and 650 Torr in a stainless-steel bubbler to deliver  $1.1 \times 10^{-3}$  sccm of the metal precursor along with 400 sccm of H<sub>2</sub>S. The monolayer MoS<sub>2</sub> deposition was achieved in 18 min at 950 °C in 50 Torr H<sub>2</sub> ambient. Following the growth of the MoS<sub>2</sub> film, the substrate was cooled down to 300 °C in H<sub>2</sub>S to prevent its decomposition. More details can be found in our earlier work<sup>51, 72, 73</sup>.

Fabrication of local back-gate islands: To define the features with local gate island configuration, the 285 nm SiO<sub>2</sub> on p<sup>++</sup>-Si substrate (commercially purchased) was first spin-coated with a bilayer Poly(methylmethacrylate) or PMMA resist in a two-step process and subsequently baked at 185 °C and 95 °C for 60s and 90s, respectively. Next, an EBPG 5200 Vistec Electron Beam Lithography tool was used to pattern the individual islands. The exposed resist was the developed using a 4-methyl-2-pentanone (MIBK) and 2-propanol (IPA) solution in a 1:1 mixture. Following a sputter cleaning of the substrate with O<sub>2</sub> plasma, the local-gate electrodes comprising of Ti/Pt (20/50 nm) layers were deposited using Temescal e-beam evaporator. This was followed by lifting off the excess metal in a cold acetone bath. Next, atomic layer deposition (ALD) process was performed to grow the Al<sub>2</sub>O<sub>3</sub>/HfO<sub>2</sub>/ Al<sub>2</sub>O<sub>3</sub> layers. Subsequently, Heidelberg Maskless Aligner

(MLA 150) was utilized to pattern and subsequently define the access points to the local back-gate electrodes. The excess oxide in the exposed areas was subsequently removed using  $\text{BCl}_3$  based reactive ion etch (RIE) chemistry.

*MoS<sub>2</sub> film transfer to local back-gate islands:* In order to fabricate the MoS<sub>2</sub> memtransistors, transfer of monolayer MoS<sub>2</sub> film from the growth substrate (sapphire) to the application substrate ( $\text{SiO}_2/\text{p}^{++}\text{-Si}$ ) with local back-gate islands is the first step. This was achieved using a polymethylmethacrylate (PMMA) assisted wet transfer process. First, the sapphire substrate with the monolayer film was spin-coated with PMMA and left idle for 24 hours in order to ensure good adhesion between PMMA and MoS<sub>2</sub>. Next, the corners of the spin-coated film were lightly scratched with a razor blade followed by its immersion in a 2 M NaOH solution at 90 °C. Owing to the hydrophilic and hydrophobic nature of sapphire and MoS<sub>2</sub>/PMMA, respectively, NaOH is be drawn towards the substrate/MoS<sub>2</sub> interface due to the capillary action thereby separating the PMMA/MoS<sub>2</sub> stack from the sapphire substrate. Using a clean glass slide, the separated film is then carefully taken out from the NaOH solution and rinsed in three separate water baths for 15 minutes each before being finally transferred onto the application substrate. Post transfer, the substrate is baked at 50 °C and 70 °C for 10 min each in order to remove the excess moisture and promote good film adhesion. This ensures a pristine interface. Finally, the PMMA is removed using acetone followed by a final cleanse with IPA.

*Fabrication of 2D memtransistors:* For fabricating the MoS<sub>2</sub> memtransistors, the substrate with MoS<sub>2</sub> on-top was spin-coated with PMMA and baked at 180 °C for 90 s. Electron beam (e-beam) lithography was used to define the features by exposing the resist which was subsequently

developed using a 1:1 mixture of 4-methyl-2-pentanone (MIBK) and IPA for 60 seconds and 45 seconds, respectively. The exposed monolayer MoS<sub>2</sub> film was etched using a sulfur hexafluoride (SF<sub>6</sub>) RIE at 5 °C for 30 s followed by an acetone and IPA rinse to remove the e-beam resist. Next, the sample was again spin-coated with methyl methacrylate (MMA) and A3 PMMA resist which was accompanied by E-beam lithography to define the source and drain contacts. The subsequent developing was once again carried out using a 1:1 mixture of MIBK and IPA for the same times as previously described. Finally, Nickel/Gold (40/30 nm) metal contacts were deposited using e-beam evaporation followed by a lift-off process to remove the evaporated Ni/Au except from the source/drain patterns by immersing the sample in acetone for 30 min followed by IPA for another 30 mins.

*Multidevice integration:* The connections between individual memtransistors were achieved using similar resist, e-beam lithography, and development processes as discussed previously. The interconnect metals Ni/Au (60/30 nm) were deposited using the E-beam evaporator. A greater metal thicknesses was required compared to the source/drain contacts due to the requirement for a conformal metal coating over/along the sidewalls. Finally, the e-beam resist was removed using a lift-off process with acetone and IPA.

*Electrical Characterization:* The electrical measurements for devices and circuits were performed in a Lake Shore CRX-VF probe station in air using a Keysight B1500A parameter analyzer.

*Dataset and Code Availability:* The experimental data generated in this study and codes used to analyze them are available from the corresponding author on reasonable request.

## **ASSOCIATED CONTENT**

### **Supporting Information**

Transfer characteristics of 16 MoS<sub>2</sub>-based memtransistors in logarithmic scale and linear scale for examining device to device variation along with a histogram plot of the threshold voltage ( $V_{TH}$ ) extracted from the linear scale. Band diagram profile of our 2D memtransistor architecture demonstrating mechanism of charge trapping and detrapping. Normal inverter output characteristics for a supply voltage  $V_{DD} = 5V$ . Degradation of the subthreshold slope (SS) for three different 2D memtransistor in a cascaded three-stage inverter (TSI) circuit module. Output inverter plots post  $V_{HRT}$  application showing a clear shift in the switching voltage  $V_m$  for all the three stages of a TSI circuit module. Normal NOR functionality for a  $V_{DD} = 5V$ . Normal NAND functionality for a  $V_{DD} = 5V$ . Long term retention of two representative analog resistance states for a total  $10^4$  seconds or 3 hours along with a memory ratio (MR) plot showcasing the eventual merging of the two resistance states.

## **ACKNOWLEDGEMENT**

The work was supported by National Science Foundation (NSF) through CAREER Award under Grant Number ECCS-2042154. The authors also acknowledge the support from The Pennsylvania State University 2D Crystal Consortium-Materials Innovation Platform (2DCCMIP) for the material growth under NSF cooperative agreement DMR-2039351.

**References:**

1. L. Yin, R. Cheng, Y. Wen, C. Liu and J. He, *Advanced Materials*, 2021, **33**, 2007081.
2. O. Mutlu, S. Ghose, J. Gómez-Luna and R. Ausavarungnirun, *Microprocessors and Microsystems*, 2019, **67**, 28-41.
3. D. U. Lee, K. W. Kim, K. W. Kim, H. Kim, J. Y. Kim, Y. J. Park, J. H. Kim, D. S. Kim, H. B. Park and J. W. Shin, 2014.
4. J. T. Pawlowski, 2011.
5. S. W. Keckler, W. J. Dally, B. Khailany, M. Garland and D. Glasco, *IEEE micro*, 2011, **31**, 7-17.
6. M. Horowitz, 2014.
7. A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh and E. Eleftheriou, *Nature nanotechnology*, 2020, **15**, 529-544.
8. G. Indiveri and S.-C. Liu, *Proceedings of the IEEE*, 2015, **103**, 1379-1397.
9. M. Di Ventra and Y. V. Pershin, *Nature Physics*, 2013, **9**, 200-202.
10. H. Jia, H. Valavi, Y. Tang, J. Zhang and N. Verma, *IEEE Journal of Solid-State Circuits*, 2020, **55**, 2609-2621.
11. Z. Jiang, S. Yin, J.-S. Seo and M. Seok, *IEEE Journal of Solid-State Circuits*, 2020, **55**, 1888-1897.
12. C. Ríos, N. Youngblood, Z. Cheng, M. Le Gallo, W. H. Pernice, C. D. Wright, A. Sebastian and H. Bhaskaran, *Science advances*, 2019, **5**, eaau5759.
13. D. Ielmini and H.-S. P. Wong, *Nature Electronics*, 2018, **1**, 333-343.
14. S. Adee, *IEEE SpEctrum*, 2008, **45**, 34-39.
15. J. Kumagai, *IEEE Spectrum*, 2000, **37**, 43-48.
16. R. S. Chakraborty, S. Narasimhan and S. Bhunia, 2009.
17. M. Tehranipoor and F. Koushanfar, *IEEE design & test of computers*, 2010, **27**, 10-25.
18. Y. Alkabani and F. Koushanfar, 2009.
19. S. Bhunia, M. S. Hsiao, M. Banga and S. Narasimhan, *Proceedings of the IEEE*, 2014, **102**, 1229-1247.
20. S. Bhunia and M. Tehranipoor, *Cham., Switzerland: Springer*, 2018.
21. R. S. Chakraborty, I. Saha, A. Palchaudhuri and G. K. Naik, *IEEE Design & Test*, 2013, **30**, 45-54.
22. M. Potkonjak, 2010.
23. L. Lin, W. Burleson and C. Paar, 2009.
24. S. S. Ali, R. S. Chakraborty, D. Mukhopadhyay and S. Bhunia, 2011.
25. A. Wali, S. Kundu, A. J. Arnold, G. Zhao, K. Basu and S. Das, *ACS nano*, 2021, **15**, 3453-3467.
26. P. Wu, D. Reis, X. S. Hu and J. Appenzeller, *Nature Electronics*, 2021, **4**, 45-53.
27. A. Dodda, S. S. Radhakrishnan, T. F. Schranghamer, D. Buzzell, P. Sengupta and S. Das, *Nature Electronics*, 2021, **4**, 364-374.
28. C. Wen, X. Li, T. Zanotti, F. M. Puglisi, Y. Shi, F. Saiz, A. Antidormi, S. Roche, W. Zheng and X. Liang, *Advanced Materials*, 2021, 2100185.
29. A. Oberoi, A. Dodda, H. Liu, M. Terrones and S. Das, *ACS nano*, 2021.
30. A. Wali, H. Ravichandran and S. Das, *ACS Nano*, 2021, **15**, 17804-17812.
31. A. Biswas and A. P. Chandrakasan, 2018.
32. A. Jaiswal, I. Chakraborty, A. Agrawal and K. Roy, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2019, **27**, 2556-2567.

33. C.-X. Xue, Y.-C. Chiu, T.-W. Liu, T.-Y. Huang, J.-S. Liu, T.-W. Chang, H.-Y. Kao, J.-H. Wang, S.-Y. Wei and C.-Y. Lee, *Nature Electronics*, 2021, **4**, 81-90.
34. H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen and M.-J. Tsai, *Proceedings of the IEEE*, 2012, **100**, 1951-1970.
35. W. Zhang, R. Mazzarello, M. Wuttig and E. Ma, *Nature Reviews Materials*, 2019, **4**, 150-168.
36. C. D. Wright, P. Hosseini and J. A. V. Diodado, *Advanced Functional Materials*, 2013, **23**, 2248-2254.
37. X. Fong, Y. Kim, K. Yogendra, D. Fan, A. Sengupta, A. Raghunathan and K. Roy, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2015, **35**, 1-22.
38. D. Fan, S. Angizi and Z. He, 2017.
39. D. W. Graham, E. Farquhar, B. Degnan, C. Gordon and P. Hasler, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2007, **54**, 951-963.
40. L. Danial, E. Pikhay, E. Herbelin, N. Wainstein, V. Gupta, N. Wald, Y. Roizin, R. Daniel and S. Kvatinsky, *Nature Electronics*, 2019, **2**, 596-605.
41. J. Hoffman, X. Pan, J. W. Reiner, F. J. Walker, J. Han, C. H. Ahn and T. Ma, *Advanced materials*, 2010, **22**, 2957-2961.
42. A. I. Khan, A. Keshavarzi and S. Datta, *Nature Electronics*, 2020, **3**, 588-597.
43. C. E. Graves, C. Li, X. Sheng, D. Miller, J. Ignowski, L. Kiyama and J. P. Strachan, *Advanced Materials*, 2020, **32**, 2003437.
44. C. Li, D. Belkin, Y. Li, P. Yan, M. Hu, N. Ge, H. Jiang, E. Montgomery, P. Lin and Z. Wang, 2018.
45. D. Akinwande, C. Huyghebaert, C.-H. Wang, M. I. Serna, S. Goossens, L.-J. Li, H.-S. P. Wong and F. H. Koppens, *Nature*, 2019, **573**, 507-518.
46. N. Li, Q. Wang, C. Shen, Z. Wei, H. Yu, J. Zhao, X. Lu, G. Wang, C. He and L. Xie, *Nature Electronics*, 2020, **3**, 711-717.
47. Z. Lin, Y. Liu, U. Halim, M. Ding, Y. Liu, Y. Wang, C. Jia, P. Chen, X. Duan and C. Wang, *Nature*, 2018, **562**, 254-258.
48. C. Liu, H. Chen, S. Wang, Q. Liu, Y.-G. Jiang, D. W. Zhang, M. Liu and P. Zhou, *Nature Nanotechnology*, 2020, **15**, 545-557.
49. K. S. Novoselov, A. K. Geim, S. V. Morozov, D.-e. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva and A. A. Firsov, *science*, 2004, **306**, 666-669.
50. W. Cao, J. Kang, D. Sarkar, W. Liu and K. Banerjee, *IEEE Transactions on Electron Devices*, 2015, **62**, 3459-3469.
51. A. Sebastian, R. Pendurthi, T. H. Choudhury, J. M. Redwing and S. Das, *Nature communications*, 2021, **12**, 1-12.
52. S. Das, A. Sebastian, E. Pop, C. J. McClellan, A. D. Franklin, T. Grasser, T. Knobloch, Y. Illarionov, A. V. Penumatcha and J. Appenzeller, *Nature Electronics*, 2021, **4**, 786-799.
53. A. Sebastian, S. Das and S. Das, *Advanced Materials*, 2022, **34**, 2107076.
54. A. Dodda, N. Trainor, J. Redwing and S. Das, *Nature communications*, 2022, **13**, 1-12.
55. R. Cheng, L. Yin, F. Wang, Z. Wang, J. Wang, Y. Wen, W. Huang, M. G. Sendeku, L. Feng and Y. Liu, *Advanced Materials*, 2019, **31**, 1901144.
56. M. Wang, S. Cai, C. Pan, C. Wang, X. Lian, Y. Zhuo, K. Xu, T. Cao, X. Pan and B. Wang, *Nature Electronics*, 2018, **1**, 130-136.

57. Y. Wang, E. Liu, A. Gao, T. Cao, M. Long, C. Pan, L. Zhang, J. Zeng, C. Wang and W. Hu, *ACS nano*, 2018, **12**, 9513-9520.
58. R. Cheng, F. Wang, L. Yin, Z. Wang, Y. Wen, T. A. Shifa and J. He, *Nature Electronics*, 2018, **1**, 356-361.
59. X. Zhu, D. Li, X. Liang and W. D. Lu, *Nature Materials*, 2019, **18**, 141-148.
60. S. Das, H. Y. Chen, A. V. Penumatcha and J. Appenzeller, *Nano Lett*, 2013, **13**, 100-105.
61. S. Das, H.-Y. Chen, A. V. Penumatcha and J. Appenzeller, *Nano letters*, 2013, **13**, 100-105.
62. P.-C. Shen, C. Su, Y. Lin, A.-S. Chou, C.-C. Cheng, J.-H. Park, M.-H. Chiu, A.-Y. Lu, H.-L. Tang, M. M. Tavakoli, G. Pitner, X. Ji, Z. Cai, N. Mao, J. Wang, V. Tung, J. Li, J. Bokor, A. Zettl, C.-I. Wu, T. Palacios, L.-J. Li and J. Kong, *Nature*, 2021, **593**, 211-217.
63. M. Lenzlinger and E. Snow, *Journal of Applied physics*, 1969, **40**, 278-283.
64. Y. Y. Illarionov, K. K. Smithe, M. Walzl, T. Knobloch, E. Pop and T. Grasser, *IEEE Electron Device Letters*, 2017, **38**, 1763-1766.
65. A. C. Khot, T. D. Dongale, K. A. Nirmal, J. H. Sung, H. J. Lee, R. D. Nikam and T. G. Kim, *ACS Applied Materials & Interfaces*, 2022, **14**, 10546-10557.
66. W. Huh, D. Lee and C. H. Lee, *Advanced Materials*, 2020, **32**, 2002092.
67. A. C. Khot, T. D. Dongale, J. H. Park, A. V. Kesavan and T. G. Kim, *ACS Applied Materials & Interfaces*, 2021, **13**, 5216-5227.
68. A. Valdes-Garcia, R. Venkatasubramanian, J. Silva-Martinez and E. Sánchez-Sinencio, *IEEE Transactions on Instrumentation and Measurement*, 2008, **57**, 1470-1477.
69. H. Le Gall, R. Alhakim, M. Valka, S. Mir, H.-G. Stratigopoulos and E. Simeu, 2015.
70. A. Pavlidis, E. Faehn, M.-M. Louërat and H.-G. Stratigopoulos, 2022.
71. Y. Xuan, A. Jain, S. Zafar, R. Lotfi, N. Nayir, Y. Wang, T. H. Choudhury, S. Wright, J. Feraca and L. Rosenbaum, *Journal of Crystal Growth*, 2019, **527**, 125247.
72. A. Dodda, A. Oberoi, A. Sebastian, T. H. Choudhury, J. M. Redwing and S. Das, *Nature communications*, 2020, **11**, 1-11.
73. D. Jayachandran, A. Oberoi, A. Sebastian, T. H. Choudhury, B. Shankar, J. M. Redwing and S. Das, *Nature Electronics*, 2020, **3**, 646-655.



### New concepts' statement

Here we show, for the first time, that logic circuits constructed using two-dimensional memtransistors can be used as Hardware trojans (HTs). Existing research has not explored this aspects inherent to in-memory compute elements. Therefore, while we use two-dimensional memtransistors-based ICs as the testbed for our demonstration, the results are equally applicable to any state-of-the-art and emerging in-memory compute technologies such as FeFETs, RRAMs, FPGAs, etc.