Physics and Technology of Electronic Insulator-to-Metal Transition (E-IMT) for Record High On/Off Ratio and Low Voltage in Device Applications

Jianqiang Lin[1,2], Khan Alam[1,2], Leonidas Ocola[1], Zhen Zhang[3], Suman Datta[4], Shriram Ramanathan[3], Supratik Guha[1,2]

[1]Center for Nanoscale Materials, Argonne National Laboratory, Lemont, IL 60439 USA. [2]Institute for Molecular Engineering, University of Chicago, Chicago, IL 60615 USA. [3]School of Materials Engineering, Purdue University, West Lafayette, IN 47907 USA. [4]Department of Electrical Engineering, University of Notre Dame, IN 46556 USA.

E-mail: linj@anl.gov. Phone: 1-630-252-1559.

Abstract

New device concepts related to both computing and biological function emulation are emerging rapidly based upon the electronic insulator-to-metal transition (E-IMT) effect that some oxides, such as VO_2, exhibit. However, the experimental E-IMT devices to-date are limited to an ON/OFF ratio of ~10^2, resulting in a small and inadequate dynamic range in device operation. In addition, the voltage that drives the E-IMT is high, typically above 1 V. In this paper, we investigate the physics and technology toward realizing both high ON/OFF and low-voltage E-IMT devices. We show that, the ON/OFF ratio, critical E-IMT voltage, and device reliability are closely coupled. A predictive model is developed and shows that, for reliable operation, the maximum ON/OFF ratio of an E-IMT device should follow a square-root relation with the strength of the thermally driven insulator-to-metal transition (T-IMT). This new design rule is verified by systematic experiments using prototypical VO_2 E-IMT devices. Through this study, we achieve a record value of reliable E-IMT with an ON/OFF ratio of 3.5×10^3 at 1.2 V – greater than 10x improvement over the previous state-of-the-art. A record low voltage of IMT switching at 0.3 V (ON/OFF ratio =20) is also demonstrated. The proposed universal design rule is widely applicable for a range of emerging applications based on E-IMT devices. As an experimental example, the E-IMT based transistors show an ultra-steep subthreshold swing (<1mV/dec) and ON/OFF ratio $>10^3$.

Introduction

Insulator-to-metal-transition (IMT) materials exhibit the unique property of an abrupt, reversible resistivity change at a critical temperature [1], known as thermal IMT (T-IMT). The electronic IMT (E-IMT) is a special case of T-IMT where the temperature change occurs by Joule heating at an applied voltage or current [2]. Recently, many new device concepts have proposed that exploit the E-IMT phenomena, including steep-slope transistors [3], memory or cross-point elements [4,5], emulation of biological function [2,6,7], coupled oscillatory networks [8], among others. The operation of these device prototypes rely upon an electro-thermal positive feedback [2,9]. However, the low ON/OFF ratio (~10^2) and high operating voltage (typically >1 V) [10] in the experimentally demonstrated E-IMT devices to-date remain as outstanding challenges that limit their potential in low-power applications. For the first time, we present a theory that explains the internal coupling of the electrical ON/OFF ratio, reliability, and voltage scaling in the IMT devices. The model result has excellent agreement with our experiments on nano-scale E-IMT devices. Using this theoretical model we have then designed E-IMT devices where we experimentally demonstrate $>10^3$ in the ON/OFF ratio (new record with >10x improvement) and reduction in transition voltage (<0.3 V) at different design points. Combining the nano-scale IMT devices with MOSFETs, we present an example case in which the new E-IMT devices can be used to study and improve an emerging steep-swing transistor [3].

Experimental

The strength of T-IMT (defined as the ratio of the resistivities at 25°C and at 125°C, denoted by M_T) in VO_2 can be controlled by deposition conditions [11]. We grew six VO_2 samples with M_T ranging from 77 to 1.8×10^5 as shown in **Table I** and **Fig. 1**. The characterization of thermal transition (T-IMT) and electronic transition (E-IMT) are respectively shown in **Fig. 2a** and **b**. The device structure and fabrication process follow [2]. Two types of E-IMT devices are fabricated. The regular device (**Fig. 3a**) has constant width. Device in this work will be constant width device unless mentioned otherwise. **Fig. 3b** is device layout with variable width. The current flows in a lateral direction through two Ti/Au contacts. We fabricated devices with length L from 40 nm to 1 μm and width w from 100 nm to 10 μm.

Model for Electronic Insulator-to-Metal Transition

The typical DC sweep on an E-IMT device in shown in **Fig. 4**. The forward sweep is important as it determines the transition threshold (V_{CF}) and the ON/OFF ratio. The ON/OFF ratio is equivalent to M_E ($M_E=I_{ON}/I_{TH}$, defined in **Fig. 4**) under DC sweep. Among all reported devices [2,12-17], the highest M_E to-date is only ~200 [10], comparing to the highest M_T in VO_2 of $>10^5$ [2].

After transition, the voltage across the device remains at V_{CF} but the resistivity changes abruptly, resulting in excess Joule heating. Under this framework, an analytical model is used to explain the relation between M_T and M_E. The key equations are in **Fig. 5**. P_{TH} and P_{PST} are respectively the Joule heating powers at pre-transition and post-transition. The derivation establishes a relationship of P_{TH}, P_{PST}, M_E and M_T. When the device reaches steady state–the case in DC measurement–the Joule heating power is balanced with the dissipated power through, primarily, convective heat transfer to the ambience. In addition, the dissipated power is, to the first order, proportional to T_{IMT}-T_a. A square-root criteria is then derived for M_E and M_T,

$$M_E = c\cdot\sqrt{M_T} \quad , \quad \text{where} \quad c=\sqrt{\frac{T_{PST}-T_a}{T_{CRT}-T_a}}\,.$$

The post-transition temperature (T_{PST}), when the device just turns on, is the key in understanding the relation of M_T and M_E. T_a is the ambient temperature, 25°C for all measurements in this work. T_{CRT} is the critical temperature for T-IMT, which is 60~65°C for VO_2 (**Fig. 1**). The highest possible T_{PST} determine the maximum value of *c*, and therefore M_E. For reliable device operation, T_{PST} should be below 400~600 °C, corresponding to value of *c* from 3 to 4. This defines the critical E-IMT strength, $M_{E,CRT}$, which is the maximum allowable M_E for reliable IMT operation. This work uses *c*=4 to calculate $M_{E,CRT}$ for reliable E-IMT operation (**Table I**). For VO_2 the highest M_T achieved so far is 2×10^5 [2], translating to a $M_{E,CRT}$ of 2×10^3.

We examine the above relationship using a coupled electro-thermal model [8]. A step voltage of V_{CF} is applied to turn the E-IMT device from OFF-state to ON-state in the simulation. T_{PST} is extracted for different combination of M_E and M_T as shown in

U.S. Government work not protected by U.S. copyright

Fig. 6. The contour is triangular because M_E is always lower than M_T. For some combinations of M_T and M_E, T_{PST} remains low (<300ºC)–blue indicates a safe operating region (SOR). Some area show T_{PST} exceeding 600ºC–red indicates a forbidden region (FR). The contour shows a sharp transition (300-600ºC) between the two regions. This boundary sets $M_{E,CRT}$, which follows the square-root relation: $M_{E,CRT}=4.M_T^{1/2}$. This contour, which represent a relation of T_{PST}, M_E and M_T, is observed in simulations even when other parameters are changed, such as device width and length, critical transition voltage (V_{CF}), threshold current (I_{TH}), among others. Therefore the square-root relation is universal in designing any E-IMT devices driven by self-generated Joule heating.

Results and Discussion

The model predicts that when M_E>$M_{E,CRT}$, the operation falls in the FR (**Fig. 6**), resulting in excess heating and irreversible breakdown of the IMT material. **Therefore $M_{E,CRT}$ is the highest ON/OFF ratio that can be reliably sustained without runaway breakdown of the device.**

Optimizing the external series resistance, R_S (**Fig. 2b**), is the key of maximizing M_E while protecting the device from excessive heating. **Fig. 7a** shows the results of a series of measurements where we have reduced R_S (starting from a high value) while measuring M_E in a VO_2 IMT device (fabricated on Sample D in **Table I**) configured as in the circuit in **Fig 2b**. As R_S reduces, M_E increases, and the maximum value of M_E repeatedly achievable in the same device is 125. For the M_E=225 sweep, the device failed after the first sweep. The experimental results are in good agreement with the theoretical prediction of $M_{E,CRT}$=150. Based on the experimental M_T and M_E in **Fig. 7a**, we extract T_{PST} from the simulation (**Fig. 6**). It shows the T_{PST} in the devices increases rapidly as M_E increases –both increase in the exponential way (**Fig. 7b**). Upon reducing R_S, the E-IMT finally falls in the FR, leading to a permeant breakdown. The SEM image in **Fig. 8** shows a device after breakdown due to runaway overheat. The damages in VO_2 and contacts are evident.

From the above experiments in devices in different samples, we extract I_{ON} and I_{TH} (**Fig. 9**) for the maximum reliable E-IMT. Here, the separation of I_{ON} and I_{TH} is the ON/OFF ratio (M_E), which is plotted against M_T in **Fig. 10**. The highest reliable M_E in experiment follows the square-root criteria very well. Results from literature are also plotted in **Fig. 10** and they fall into the operating region predicted by the model. The experimental M_E of 1300 (Sample A in **Table I**) is a new record value, showing 10x improvement over the state-of-the-art E-IMT switches.

The above measurements were carried out using a conventional slab geometry for the VO_2 IMT element, with a constant width as shown in **Fig. 3a**. We now show that introducing device layouts with a variable width can further improve the device ON/OFF ratios. Consider the staggered layout shown in **Fig. 3b**. The narrowest spacing of the two contacts is L=40 nm, whose width is 100 nm. This "necked section" is then connected to sections of larger widths and lengths as shown (schematic in **Fig. 3b**, SEM image in **Fig 11)**. Scaling L reduces V_{CF} [8]. When the voltage ramps up, E-IMT initiates between the two closely spaced contacts in the necked section (inner circle in **Fig. 3b**). This ensures a low value of V_{CF} and I_{TH}. Once the positive feedback of E-IMT starts, the transition then propagates to the adjacent regions of larger width and reduces its resistivity. E-IMT will eventually occurs throughout all sections as shown in the outer circles in **Fig. 3b**. In this fashion—with the low value of I_{TH} thus enabled— reliable M_E can be further improved from 1300 to 3500 at reduced V_{CF} (**Fig. 12**). For M_E of ~3500, three consecutive I-V sweeps are performed to illustrate reliable E-IMT operation in the device.

Devices reported in this work demonstrated record low voltage operation (M_E=7 at 0.2 V), and >10x improvement for the ON/OFF ratio. Benchmarking with E-IMT devices reported in literature is shown in **Fig. 13**.

Application to Steep-Swing Transistor

With the new design criteria developed here, the E-IMT devices with reliable operation, high ON/OFF ratio, and low voltage can be used in a range of applications. An experimental example of ultra-steep swing transistors is demonstrated.

Two configurations for constructing IMT steep-swing transistor are shown in **Fig. 14 a** and **b**. The FETs used here are InGaAs quantum-well (QW) MOSFETs [18] with appropriate gate length and width to match the current required to switch the IMT device. The approach will work when integrated with Si and other MOSFETs as well—the example with the III-V MOSFET is a representative one. The IMT device can be connected to either the drain or the source, denoted as IMT-D and IMT-S respectively. The source-connected "hybrid FET" was first proposed and demonstrated by Shukla et. al. [3].

In this work, IMT-D with devices in Sample A demonstrates an ON/OFF ratio of 2000 at ΔV_G of 1 mV (**Fig. 15**). Its subthreshold swing is S=0.3 mV/dec. **Fig. 16** also show the transistor output characteristics by sweeping V_{DD}.

In the IMT-S configuration, we construct the hybrid-FET using the one IMT device and several QW-MOSFETs with different current drive. The applied V_{DD} is 1.1 V. **Fig. 17a** show the current increases by 1200 for ΔV_G of 1 mV (S=0.32 mV/dec). Two observations are made in **Fig. 17**. First, while the off currents at the lowest V_G are the same for both the standalone MOSFET and the hybrid FET, the ON currents at the highest V_G are significantly lower in the hybrid FET for all four cases. To quantify it, the parameter I^*_{ON}/I^O_{ON} (ratio between hybrid FET ON current, I^*_{ON}, and MOSFET ON current, I^O_{ON}) is extracted whose value is always less than unity. Second, the ON/OFF ratio in the steep region (I^*_{ON}/I_{TH}) decreases in the hybrid FET as the current drive reduces in the standalone MOSFET. This results in the loss of steepness (I^*_{ON}/I_{TH} for a fixed V_G step). The extraction of I^*_{ON}, I^O_{ON} and I_{TH} are illustrated in **Fig. 17a**.

The two ratios, respectively representing the loss of ON current and loss of steepness, are plotted in **Fig. 18**. Three regions of operation emerge in **Fig. 18**. This can be explained by the ON-state conductance in the IMT (σ_{IMT}) and MOSFET (σ_{FET}). In region 1, σ_{FET} is comparable to (or greater than) σ_{IMT}, therefore there is a significantly voltage drop in the IMT in ON-state. σ_{FET} in region 2 is significantly lower than σ_{IMT}. In region 3, σ_{FET} is too low to reach the required I_{TH} to trigger transition.

Engineering the IMT element for low M_E values enables the design of low voltage devices since the starting currents are high [2]. As a representative example, **Fig. 19** shows the results of a hybrid FET processed using sample F (M_E~77) where we are able to achieve switching at 0.33 V, the lowest demonstrated to date. Results from the hybrid-FETs from this work are summarized in **Table II** along with previous reports.

Conclusions

In this work, we have derived a guideline in designing electronic insulator-to-metal transition based phase change devices that are electrically driven. The proposed square-root relation between the strengths of E-IMT and T-IMT is compared against experiment, showing excellent agreement. Our device shows record E-IMT switching of $M_E=3.5x10^3$ (at 1.2 V) and low voltage at 0.3 V (M_E=20). The new devices are used in an emerging application of steep-swing transistor and demonstrate significant improvement over the state-of-the-art.

Table 1. Six VO_2 samples with different IMT characteristics are grown for this study. The thermal IMT strengths vary from 10^2 to 10^5.

Sample ID	ρ (T=25 °C) [Ω-cm]	ρ (T=125 °C) [Ω-cm]	M_T	Predicted critical strength of E-IMT, $M_{E,crt}$	Substrate
A	30	0.00017	$1.8X10^5$	1700	Sapphire
B	9.1	0.00035	$2.6X10^4$	650	Sapphire
C	4.6	0.0012	3700	240	SiO_2/Si
D	1.8	0.0012	1500	150	SiO_2/Si
E	1.4	0.0029	480	88	SiO_2/Si
F	0.69	0.009	77	35	SiO_2/Si

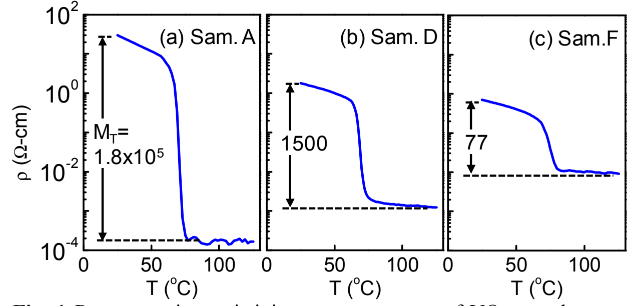

Fig. 1 Representative resistivity vs. temperature of VO_2 samples grown through different conditions. M_T for the samples shown here are (a) $1.8x10^5$, (b) $1.5x10^3$ and (c) 77 respectively.

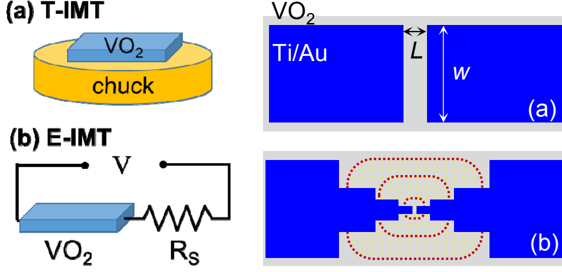

Fig. 2 Schematic of the characterization for VO_2 devices under (a) T-IMT and (b) E-IMT.

Fig. 3 Device layouts with (a) constant width, and (b) variable width. In device with variable width, the transition first occurs in the "necked section" (inner red circle) and trigger IMT in other sections (outer red circles).

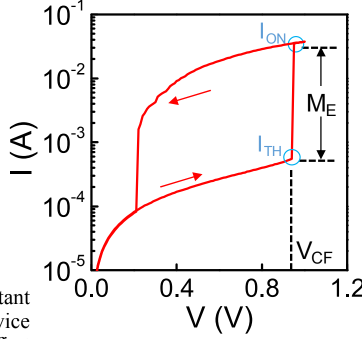

Fig. 4 Feature of a hysteresis I-V sweep for E-IMT. At forward critical voltage (V_{CF}), the current abruptly increases from I_{TH} (pre-transition) to I_{ON} (post-transition). The strength for E-IMT, M_E, is I_{ON}/I_{TH}.

$$M_T = \frac{R_{INS}}{R_{MET}} \quad (1)$$

$$M_E = \frac{R_{INS}+R_S}{R_{MET}+R_S} \sim \frac{R_{INS}}{R_{MET}+R_S} \quad (2)$$

$$P_{TH} = \frac{V_{CF}^2}{R_{INS}} \quad (3)$$

$$P_{PST} = \left(\frac{V_{CF}}{R_S+R_{MET}}\right)^2 R_{MET} \quad (4)$$

Combined 1-5: $$\frac{P_{PST}}{P_{TH}} = \frac{M_E^{\ 2}}{M_T}$$

Fig. 5 Analytical relationship of M_E and M_T through the Joule heating powers in the pre-transition and post-transition state. R_{INS} and R_{MET} are the insulating and metallic resistance of the IMT.

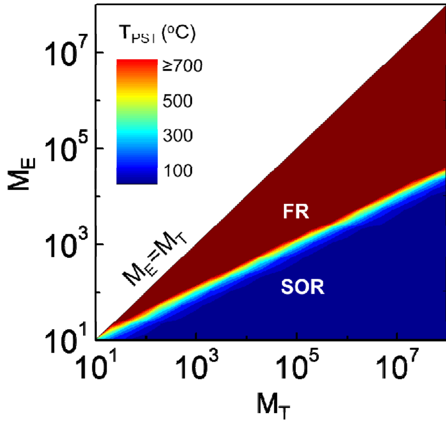

Fig. 6 Simulated post-transition temperature (T_{PST}) contours from the electro-thermal model for different combination of M_E and M_T. The device operation can be in the safe operating region (SOR) or the forbidden region (FR) depending on the post-transition temperatures.

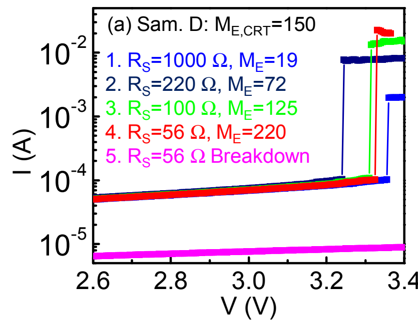

Fig. 7 (a) I-V sweeps of the IMT device in series with R_S, in sequence from high R_S to low R_S, to examine the maximum allowable M_E. (b) Temperature for the three cases of M_E obtained from the I-V sweeps. T_{pst} for sweep #1-3 is in SOR, and that for #4 is in the FR.

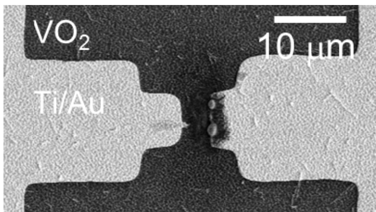

Fig. 8 SEM image of device after breakdown. At the center of the device, the contact region and the VO_2 are damaged.

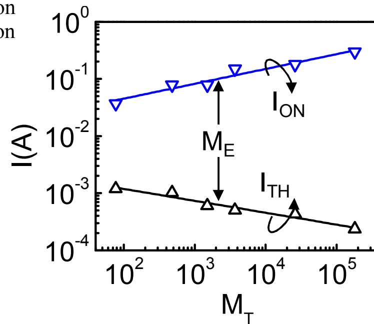

Fig. 9 I_{ON} and I_{TH} of E-IMT in six samples with different M_T. The separation of I_{ON} and I_{TH} reflects the electronic IMT strength, M_E.

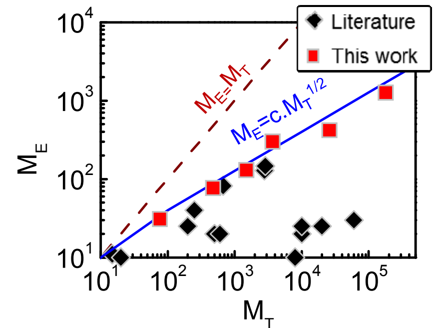

Fig. 10 M_E vs.M_T from data in **Fig. 9**, showing good agreement with the square-root relation (blue line). All literature data falls into the safe operating region predicted by the model.

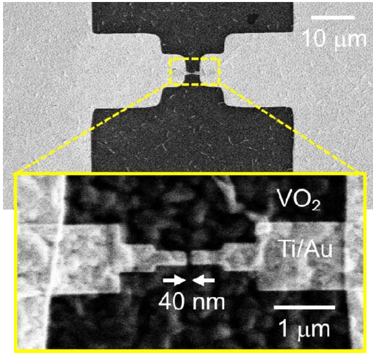

Fig. 11 SEM top view of the devices with variable width. Inset shows the enlarged view at the center of the device. The length of IMT device (L) is 50 nm.

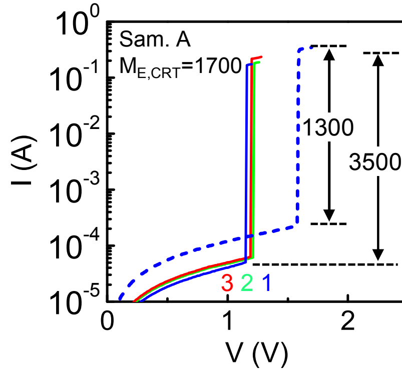

Fig. 12 I-V sweeps on devices fabricated on Sample A with different layout: constant-width (dashed lines) and variable width (solid lines) Three consecutive sweeps are measured for the later cases.

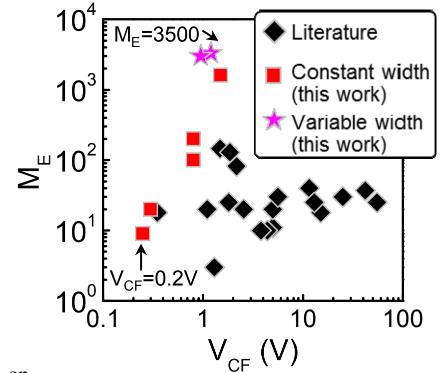

Fig. 13 Benchmarking of M_E and operating voltage for VO_2 E-IMT devices.

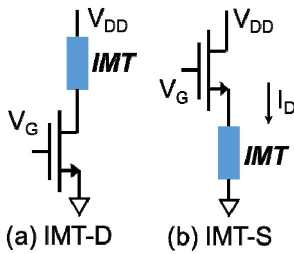

Fig. 14 Schematic of one IMT and one transistor connected in series: (a) IMT connected to the drain, IMT-D, and (b) IMT connected with the source, IMT-S.

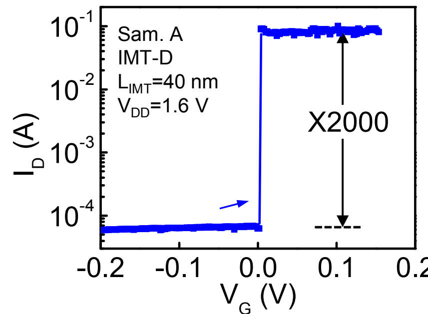

Fig. 15 I_d vs. V_G, the transfer characteristics at different V_{DD} of 1.6 V for the IMT-D steep swing hybrid FET. With an appropriate QW-MOSFET, record ON/OFF ratio of 2000 is achieved with S=0.3 mV/dec.

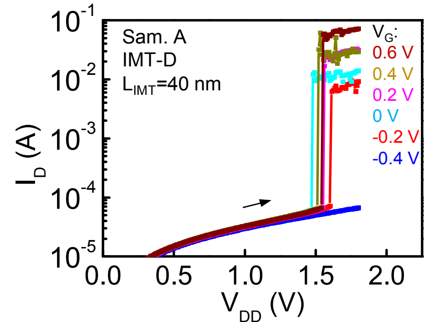

Fig. 16 I_d vs. V_{DD}, the output characteristics at different V_G.

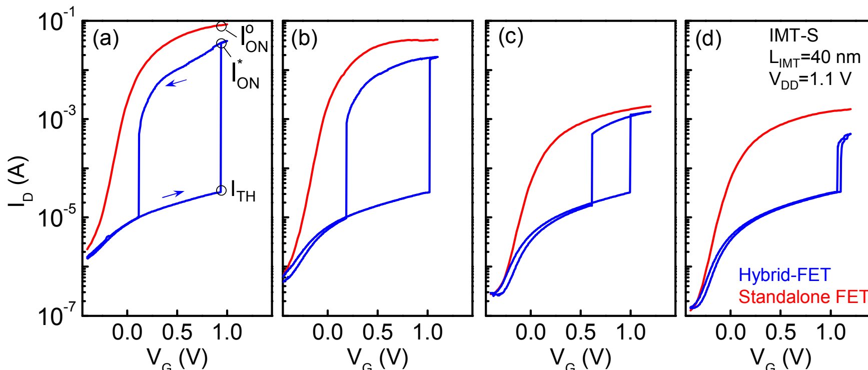

Fig. 17 I_d vs. V_G transfer characteristics at different V_{DD} of 1.1 V for the IMT-S hybrid-FETs (blue) and standalone QW-MOSFETs (red). From **a** to **d**, the QW MOSFETs has decreasing current while the IMT remains the same and. (a) demonstrate an ON/OFF current ratio of 1200 across 1 mV gate voltage step (S=0.32 mV/dec).

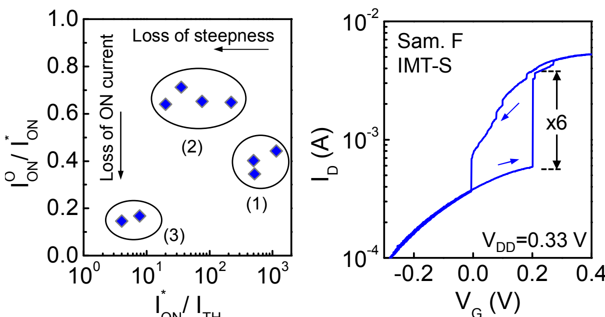

Fig. 18 Ratio of current components (denoted in **Fig. 16a** indicating the loss of ON-state current versus loss of steepness in IMT-S hybrid FETs. Regions 1-3 are resulted from different conductance matching. (see text)

Fig. 19 I_D-V_G characteristics for a hybrid-FET with low supply voltage of 0.33 V.

Table II. Summary of steep-swing transistors from a hybrid structure with E-IMT. (*I_{ON}/I_{OFF} is extracted in the steep switching region, see **Figs. 16-17, 19**)

Ref.	Connetion	L_{IMT}	V_{DD}	I_{ON}/I_{OFF}*
This work	IMT-D	40 nm	1.6 V	2000
This work	IMT-S	40 nm	1.1 V	1200
This work	IMT-S	40 nm	0.33 V	6
[3]	IMT-S	100 μm	5.2 V	35
[3]	IMT-S	200 nm	1.6 V	5

Acknowledgements

The authors acknowledge Prof. J. A. del Alamo in the Massachusetts Institute of Technology for his technical support. This work was performed, in part, at the Center for Nanoscale Materials, a U.S. Department of Energy Office of Science User Facility under Contract No. DE-AC02-06CH11357. Aspects of the device work was supported jointly by the National Science Foundation under grant 1640081, and the Nanoelectronics Research Corporation (NERC), a wholly owned subsidiary of the Semiconductor Research Corporation (SRC), through Extremely Energy Efficient Collective Electronics (EXCEL), an SRC-NRI Nanoelectronics Research Initiative under Research Task ID 2698.001., and research grants ARO W911NF-16-1-0289, ONR N00014-16-1-2398.

Reference

[1] Y. Zhou and S. Ramanathan, Proc. IEEE, vol. 103, no. 8, pp. 1289–1310, 2015.
[2] J. Lin, et al., in IEDM Tech. Dig., 2016, pp. 862–865.
[3] N. Shukla, et al., Nat. Commun., vol. 6, p. 7812, 2015.
[4] M. Son et al., IEEE Electron Device Lett., vol. 32, no. 11, pp. 1579–1581, 2011.
[5] E. Cha, et al., in ISCAS, 2014, pp. 428–431.
[6] M. D. Pickett, et. al., Nat. Mater., vol. 12, no. 2, pp. 114–117, 2013.
[7] M. Jerry, et al., in Proc. of DRC, 2016.
[8] N. Shukla, et al., Sci. Rep., vol. 4, p. 4964, 2014.
[9] D. Li, et al., ACS Appl. Mater. Interfaces, vol. 8, no. 20, pp. 12908–12914, 2016.
[10] Y. Zhou, et al., IEEE Electron Device Lett., vol. 34, no. 2, pp. 220–222, 2013.
[11] H.-T. Zhang, et al., Nat. Commun., vol. 6, p. 8475, 2015.
[12] H.-T. Kim et al., New J. Phys., vol. 6, p. 52, 2004.
[13] A. Crunteanu, et al., SPIE Proceedings vol. 9364, p. 93640J, 2015.
[14] A. Zimmers, et al., Phys. Rev. Lett., vol. 110, no. 5, p. 56601, 2013.
[15] Z. Yang, et al, J. Appl. Phys., vol. 110, no. 3, p. 33725, 2011.
[16] J. Leroy, et al, Appl. Phys. Lett., vol. 100, no. 21, p. 213507, 2012.
[17] I. P. Radu, et al., Nanotechnology, vol. 26, no. 16, p. 165202, 2015.
[18] J. Lin, et al, IEEE Electron Device Lett., vol. 37, no. 2, pp. 381–384, 2016.