Antiferroelectric Multilayers: a Multifunctional Platform for Energy and Information Storage

Hung-Wei Li, Sadegh Kamaei, Cyrille Masserey, Pascal Morel, Niccolò Martinolli, Tom Schlatter, Océane Mauroux, Igor Stolichnov, and Adrian M. Ionescu.
EPFL, Nanoelectronic Devices Laboratory (NanoLab), 1015, Lausanne, Switzerland

***Abstract*— This work demonstrates the integration of energy storage and multi-bit memory functionalities in a single platform device using multilayer dielectric (DE) with ferroelectric (FE) or antiferroelectric (AFE) materials. Employing a hafnium-based oxide, we achieve an energy storage density of approximately 50 J/cm^3 and advanced multi-bit storage capabilities. We investigate the effects of adjusting silicon doping concentration in multilayer Si:HfO_2 during the atomic layer deposition (ALD) process to form and compare both FE and AFE layers. Inserting an Al_2O_3 dielectric (DE) layer between the FE or AFE layers, we create a built-in electric field, enhancing both energy storage efficiency and multi-bit memory storage beyond the current state of the art. Our innovative approach harnesses the intrinsic properties of these materials to boost energy efficiency and data storage capacity, significantly improving device functionality and advancing electronic component design. Our results reveal that strategically layering DE with AFE materials enhances energy and multi-bit storage, opening new avenues for future multifunctional electronic applications.**

I. Introduction

Hafnium oxide-based ferroelectric (FE) and antiferroelectric (AFE) thin films are emerging as key materials in the evolution of advanced memory and energy storage technologies. Integrating these materials into commercially viable devices demands not only comprehensive material characterization but also innovative engineering to enhance both endurance and performance. This study introduces a groundbreaking device that seamlessly integrates high-density energy storage with robust multi-bit memory functionality within a techmological framework. Utilizing multilayer configurations of dielectric (DE), FE, and AFE materials, we have achieved an unprecedented energy density of approximately 50 J/cm^3, coupled with enhanced multi-bit storage capabilities. This advancement is made possible through meticulous control of silicon doping in Si:HfO_2 during atomic layer deposition (ALD), allowing for the precise formation of distinct FE and AFE phases. A key innovation in our device is the incorporation of an Al_2O_3 dielectric layer between the FE and AFE layers, which generates a built-in electric field. This field significantly boosts energy storage efficiency while simultaneously stabilizing data retention in multi-bit memory configurations (Fig. 1). As the demand for ultracompact electronic devices increases, particularly in wearable and implantable technologies, our research addresses the *multi-functional need for devices that not only store energy efficiently but also retain data reliably*. The miniaturized energy autonomous systems enabled by this technology will integrate energy harvesting and storage with memory, offering a comprehensive solution for long-term, reliable device operation. Our findings not only demonstrate substantial improvements in energy and memory performance but also pave the way for the development of next-generation electronic components.

II. Device Fabrication Process

A. Tailoring silicon doped hafnium oxide properties

In the ALD process for Si:HfO_2 thin films, adjusting the SiO_2 and HfO_2 cycle ratios (Fig. 2a) yields either ferroelectric (FE) or antiferroelectric (AFE) behavior [1]. The metal-ferroelectric-metal (MFM) capacitor exhibits ferroelectric (Fig. 2(b)) or antiferroelectric (Fig. 2(c)) behavior depending on the SiO_2 doping concentration.

B. FeCap fabrication

The fabrication process for the multilayer capacitors illustrated in Fig. 3(a). We begin by preparing a silicon substrate with a 200-nm thermally grown SiO_2 layer on both sides. After the standard cleaning process, the bottom electrodes composed of titanium (Ti) and platinum (Pt) were deposited by sputtering. Without breaking the vacuum, we sputter a titanium nitride (TiN) layer. The multi-layer structure is then created through ALD. Table 1 provides details on the various sample configurations. For consistency in comparison, we maintained the overall thickness of FE/AFE layers at 40 nm across all samples. After the multi-layer structure is complete, we deposit a second TiN and the entire stack then undergoes a rapid thermal annealing (RTA) in a nitrogen atmosphere. This step is crucial in achieving the desired orthorhombic crystalline phase, which is essential for tuning the FE or AFE properties of our device. The final stage of fabrication involves defining individual capacitors, by sputtering Ti and Pt layers, followed by a ion beam etching (IBE) process. This results in capacitors with a well-defined area of in the range of 100 μm^2 for characterization and performance evaluation.

III. Results and Discussion

A. Multi-layer Capacitor Characterization

Fig 5 illustrates the experimental hysteresis curves of dielectric and ferroelectric behaviors in fabricated multilayer

capacitors, highlighting the influence of interfacial engineering and material composition on electrical properties. Fig. 5(a) and (b) show the polarization-electric field (P-E) curves for DE1FE10×4 and DE1AFE10×4, respectively. The P-E curve for DE1FE10×4 displays characteristic ferroelectric behavior with suppressed remanent polarization (P_r), attributed to the dielectric layer's effect on ferroelectric switching. While the MFM capacitor exhibits a $2P_r$ of 40 μC/cm^2, DE1FE10×4 shows a reduced $2P_r$ of 20 μC/cm^2. The P-E curve of DE1AFE10×4 demonstrates antiferroelectric characteristics and the versatility of layering strategies in tailoring material properties. The observed crossover of capacitance-voltage (C-V) branches (Figs 5(c) and (d)) at a non-zero electric field indicates the presence of a built-in electric field, likely resulting from asymmetric charge distributions or interfacial dipole moments within the multilayer structure. Fig. 6(a) and (b), representing DE1FE5×8 and DE1AFE5×8 respectively, demonstrate how a thicker dielectric layer alters the P-E response. This is evident from the linear capacitor behavior observed at lower electric fields, which also contributes to an increase in the electric breakdown voltage at high electric fields.

B. Energy storage enhancement with multilayered AFE/FE

Fig. 7(b) and (c) compares energy storage capabilities of ferroelectric (FE) and antiferroelectric (AFE) multilayers. Our work demonstrated that AFE materials exhibit superior performance due to their unique phase transition properties. Energy storage is calculated from the area under the P-E curve, revealing total stored energy and hysteresis losses (Fig 7(a)). This method highlights fundamental differences in dipole behavior between FE and AFE materials under electric fields. FE multilayers like DE10FE12×2 achieve energy densities up to 30 – 40 J/cm^3 but are limited by breakdown electric field. In contrast, AFE multilayers, particularly DE1AFE10×4, exceed 60 J/cm^3 with efficiencies up to 80% at 3 MV/cm. This performance stems from AFE materials' ability to *switch between non-polar and polar states*, allowing for controlled energy discharge. To further increase the breakdown voltage, we compared more stacks of the multilayer. However, while increasing the number of layers achieves a higher breakdown voltage, it also results in lower energy density. Both phenomena can be observed in FE and AFE devices, which is due to the more pronounced DE contribution suppressing the dipole from FE and AFE layers. Unlike the rapid discharge in ferroelectrics, AFE materials demonstrate *controlled energy release*. This driven by reversible phase transitions, enhancing energy storage capacity and efficiency, and making AFE multilayers suitable for applications requiring efficient, high-density energy storage. Charge-discharge characteristics of FE and AFE multilayer capacitors, providing insights into their practical performance. The setup uses a DC power supply and controlled discharge through a resistor to simulate real-world conditions (Fig. 8(a)). The voltage-time curves in Fig. 8(b) and 8(c) compare different multilayer configurations during charge-discharge cycles, using a 2.2 nF capacitor as a reference which is close to the maximum capacitance (Fig.5 (d)) we can obtain from AFE device. FE structures display rapid voltage decay, typical of quick energy release due to polarization switching, with DE1FE10×4 showing a slightly slower but still significant drop from 15V to near 0V within about 100 μs. In contrast, AFE structures exhibit more controlled voltage decay, with DE1AFE10×4 maintaining higher voltage for longer. This confirms the superior performance of AFE materials in controlled energy release, making them ideal for applications requiring energy efficiency and longevity. Fig. 9 demonstrates the long-term stability of AFE materials under extensive field cycling, crucial for durable energy storage applications.

C. Multibit memory storage enabled by multilayered AFE/FE

FORC measurements (Fig. 9(a-f)) offer insights into the switching behavior of these materials. By inserting the DE layer, the current peaks can be finely tuned due to the built-in electric field (Fig. 9(b)). The DE1FE10×4 structure shows typical ferroelectric hysteresis with distinct current peaks at coercive fields. The presence of the DE layer subtly shifts these coercive fields closer to ±2 MV/cm. Integrating DE layers making DE1FE10×4 suitable for memory and energy storage technologies with fast response times and improved endurance. For DE1AFE10×4 (Fig. 9(e)), the FORC diagram reveals a shift in current peaks towards the same field polarity, a phenomenon not observed without the DE layer. This shift, caused by the DE layer's influence on the internal electric field, indicates multibit nonvolatile properties. By stabilizing specific states, this shift ensures reliable quaternary multibit memory operations. The read-out transient current data further supports these findings (Fig. 11). A ±12V pulse switches one peak, while a ±20V pulse switches both, with the current in the mA range. This demonstrates precise control, enabling accurate multibit memory storage and retrieval. The C-V characteristics of AFE materials (Fig. 12) display a double-peak butterfly shape with significant frequency dispersion, indicating rapid switching speeds in DE+AFE structures.

IV. Conclusion

This study comprehensively analyzes for the first time the energy storage capabilities, switching behavior, and memory performance of ferroelectric (FE) and antiferroelectric (AFE) multilayer capacitors. Through detailed charge-discharge measurements, C-V characteristics, and FORC analysis, it is demonstrated that *AFE multilayers exhibit superior energy storage capacity, controlled energy release, and enhanced nonvolatile properties compared to FE structures*. These advantages position AFE materials as a promising platform candidate for next-generation advanced electronic devices with multifunctional energy storage and memory capability.

References

[1] T. Boscke, et al, *Applied Physics Letters*, vol. 99, 2011. [2] F. Ali, et al, . *J Appl Phys* , vol. 122, 2017. [3] J. P. B. Silva, et al, *J Mater Chem A*, vol. 8, 2020. [4] Cheema, S. S., et al, *Nature*, 2024.

Multilayer Structure Fabrication Process

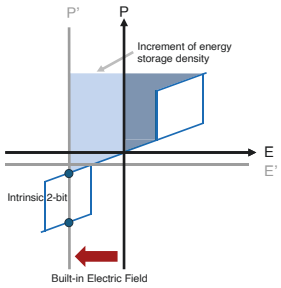


Fig 1. Schematic illustration of the qualitative P-E loop for the multilayer metal-insulator-ferro-electric-metal (MIFM) capacitor structure. The integration of a dielectric (DE) layer with antiferroelectric (AFE) or ferroelectric (FE) materials introduces a built-in electric field. This field significantly enhances the energy storage density and enables multi-bit memory functionality within the same MIFM capacitor structure, optimizing both memory capacity

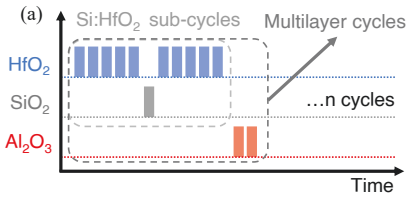
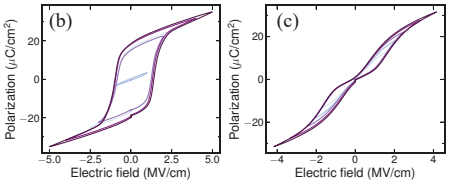


Fig 2. (a) ALD process schematic for Si:HfO_2 thin films. Si:HfO_2 sub-cycle ratio adjustments yield FE or AFE behavior.



P-E hysteresis loops of MFM capacitors: (b) 3.12% Si concentration (ferroelectric), (c) 5% Si concentration (antiferroelectric).

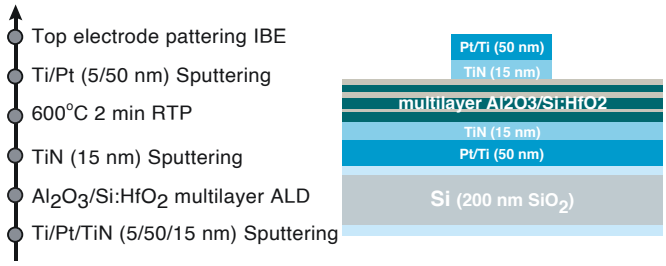


Fig 3. Key fabrication steps for the multilayer capacitor structure. The ALD multilayer cycles are adjusted to achieve different material combinations in the multilayer structure.

Sample ID	T_{AlO} (nm)	T_{SiHfO} (nm)	Layer numbers	Si:HfO_2 properties
DE10FE10x2	10	10	2	Ferroelectric
DE1FE10x4	1	10	4	Ferroelectric
DE1FE8x5	1	8	5	Ferroelectric
DE1FE5x8	1	5	8	Ferroelectric
DE1AFE10x4	1	10	4	Antiferroelectric
DE1AFE8x5	1	8	5	Antiferroelectric
DE1AFE5x8	1	5	8	Antiferroelectric

Fig 4. The tables summarize key samples, providing its thickness of each layer and the ferroelectric/antiferroelectric properties.

Experimental Device Characterisitics

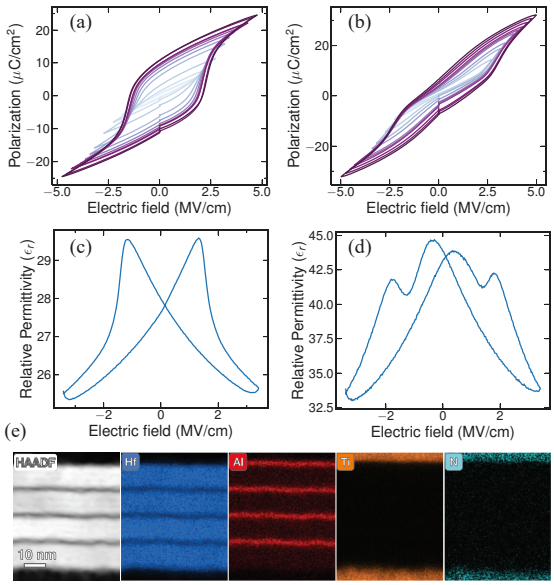


Fig 5. (a) P-E curve of DE1FE10x4 showing typical ferroelectric behavior, with suppressed remanent polarization due to the dielectric layer's contribution. (b) P-V curve of DE1AFE10x4 demonstrating antiferroelectric characteristics. (c) Butterfly-like dielectric response of the FE capacitor after interfacial engineering of interlayers, with a crossover of C-V branches at a non-zero electric field indicating a built-in field. (d) Butterfly-like dielectric response of the AFE capacitor under similar conditions, also showing a crossover due to a built-in field. (e) TEM image and EDS results for the multilayer structure, providing a detailed view of the material composition and interfaces.

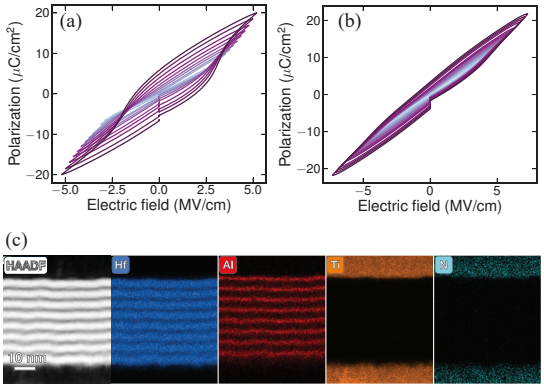
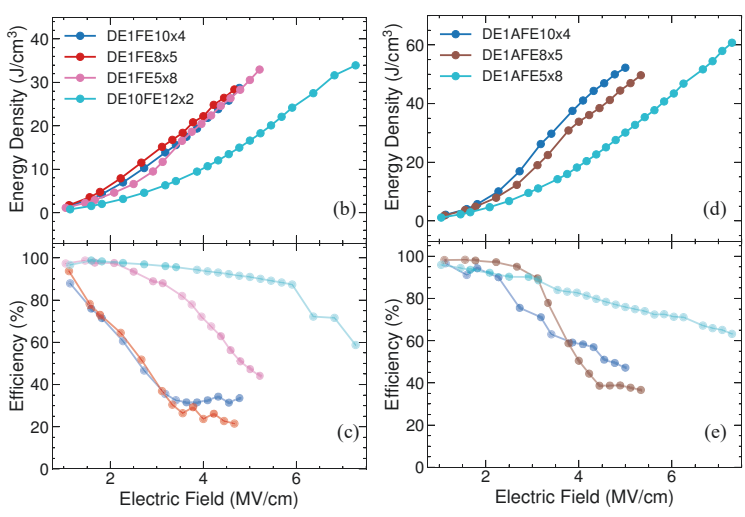


Fig 6. (a) P-E curve of DE1FE5x8, showing ferroelectric properties with a noticeable linear capacitor response at low electric fields due to the increased dielectric layer contribution. (b) P-V curve of DE1AFE5x8, also exhibiting a linear capacitor response at low electric fields, influenced by the dielectric layer. (c) TEM image and EDS results, providing detailed insights into the material structure and composition.

Experimental Device Characterisitics: Energy Storage Density



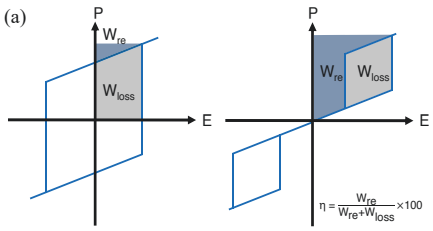


Fig 7. (a) Calculation of energy density and efficiency derived from the P-E loop, involving the integration of the area under the polarization-electric field curve to assess the energy stored and the corresponding efficiency based on energy loss. (b, c) Energy density versus electric field plots for ferroelectric multilayers, showing the material's capacity to store more energy when adjusting the inserting DE layers. (d, e) Similar plots for antiferroelectric multilayers, highlighting the distinct energy storage characteristics compared to their ferroelectric counter-parts, including differences in energy density and efficiency.

Charge - Discharge Measurement

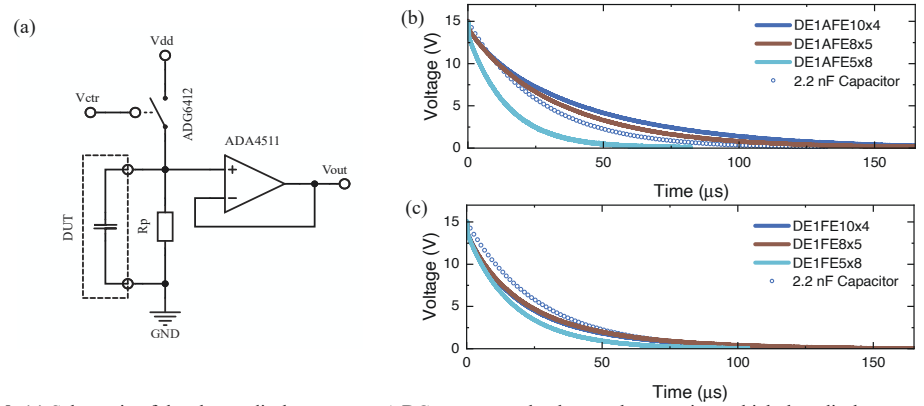


Fig 8. (a) Schematic of the charge-discharge setup: A DC power supply charges the capacitor, which then discharges through a 12 kΩ resistor. The pulse duration (10 µs) is controlled by a switch (ADG6412), and the voltage is monitored using an oscilloscope. A follower (ADA4511) reduces the parastics from the oscilloscope. (b, c) Comparison of discharge curves for ferroelectric and antiferroelectric multilayer structures with those of an ideal 2.2 nF capacitor. These simulated values are used as a reference to assess the performance of the multilayer structures.

Endurance

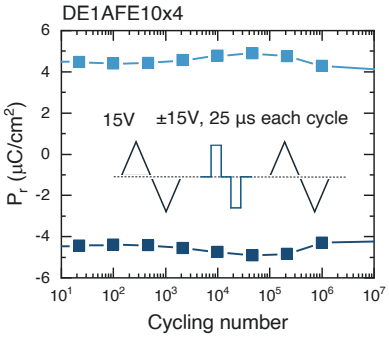


Fig 9. The device demonstrates fatigue-free behavior during field cycling, benefiting from the intrinsic properties of the antiferroelectric (AFE) material.

First-order Reversal Curve (FORC)

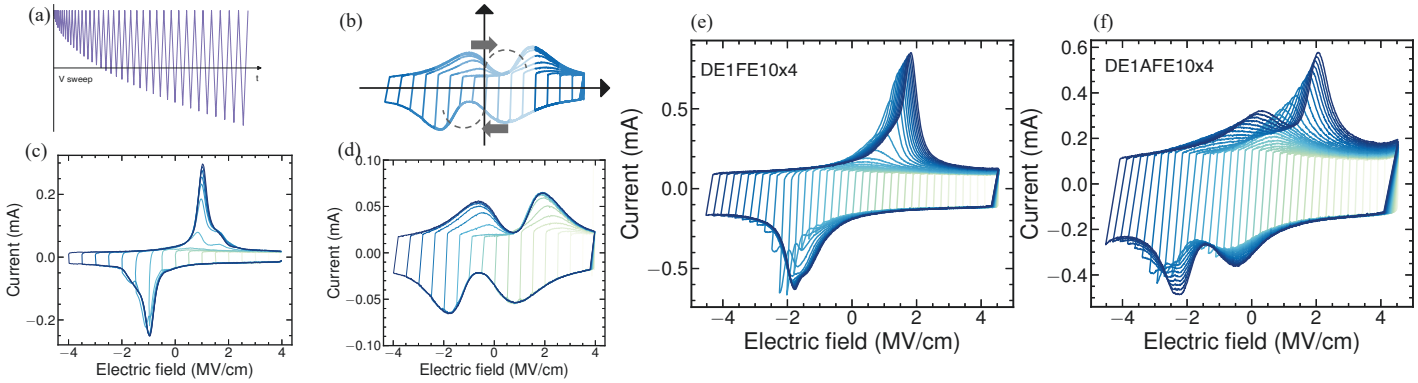


Fig 10 (a) Schematic outlining the First-order Reversal Curve (FORC) measurement approach. The applied field is swept from positive saturation to a reversal point, then back to positive saturation. This process is repeated for multiple reversal points to generate a series of curves that together form a comprehensive map of the material's electric behavior. (b) Schematic of AFE with DE multilayer, the current peaks will shift. FORC result of the MFM capacitors exhibiting (c) FE properties and (d) AFE properties. (e) FORC results showing the effect of the DE layer on the FE multilayer capacitor, illustrating changes in the coercive field and interaction fields due to the DE layer. (f) FORC result of the AFE multilayer capacitor with a DE interlayer, where the addition causes the two peaks at opposite field polarities to shift toward the same polarity, indicating multi-bit nonvolatile properties. This shift suggests a modification in the energy landscape of the AFE material due to the presence of DE layers.

Multi-bit Cell Write and Read

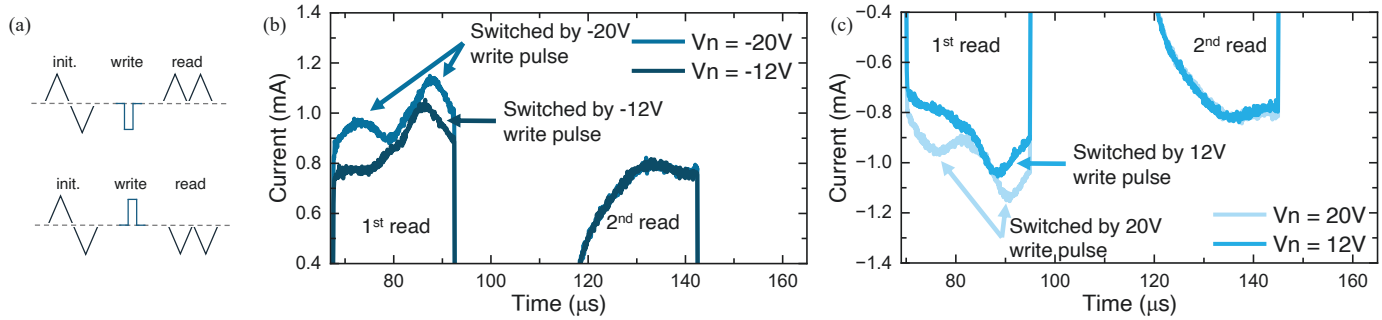


Fig. 11 (a) Schematic of the pulse sequence used to verify the write operation. After the write operation, two pulses (18V) with opposite polarity are applied to read out the switching polarization achieved during the write process. (b, c) Read-out transient current for different write pulse levels. The results show that a ± 12 V pulse can switch only one peak, while a ±20V pulse can switch both peaks, demonstrating the control over the switching behavior with varying pulse levels.

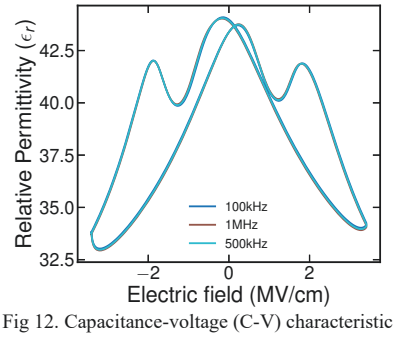


Fig 12. Capacitance-voltage (C-V) characteristic showing a double-peak antiferroelectric (AFE) butterfly shape. The robust frequency dispersion observed indicates a fast switching speed in the multilayer DE AFE structure.

	Material	ESD (J/cm³)	Efficiency	Memory and Energy Storage Integration
F. Ali, et al [2]	Si-HfO_2	61	65%	-
J. P. B. Silva, et al [3]	Al_2O_3: DIL-$Hf_{0.5}Zr_{0.5}O_2$	54	51%	-
Cheema, S. S. et al [4]	HfO_2/ZrO_2	115	90%	-
This work	Al_2O_3/Si:HfO_2	50	80%	Yes

Table 2. Benchmarking of various on-chip CMOS compatible energy storage capacitors.

Wafer-Scale 2D MoS_2 Transistors Using Transfer-Free Location-on-Demand Selective Synthesis

Anthony Cabanillas,[1] Chu Te Chen,[1] Asma Ahmed,[1] Anthony Butler,[1] Yu Fu,[1] Ajay Yadav,[2] Gabriel Lee,[2] Keith Wong,[2]* Fei Yao,[1]* and Huamin Li[1]*
[1]University at Buffalo, The State University of New York, Buffalo, NY, USA
[2]Applied Materials, Inc., Sunnyvale, CA, USA
*Email: keith_wong@amat.com, feiyao@buffalo.edu, huaminli@buffalo.edu

Abstract — Compatible integration of emerging two-dimensional (2D) materials with mature Si-CMOS technology is promising to enable high-performance energy-efficient electron devices. In this work, we exploited wafer-scale, location-on-demand, selective growth of 2D semiconducting transition metal dichalcogenide (TMD), MoS_2, on a SiO_2/Si substrate for transfer-free electron device applications. We investigated the impact of native oxide MoO_3 dielectrics on the performance of MoS_2 field-effect transistor (FET) arrays through a comparative study with SiO_2 dielectrics, and demonstrated great potential of the selective growth of 2D semiconductors to lower down the technological requirement for practical integration with Si-CMOS technology.

I. Introduction

To continue the Moore's law and maintain device miniaturization, 2D semiconducting TMDs have been extensively explored as one of the most promising channel materials. Despite the early success in fundamental science explorations and proof-of-concept device demonstrations [1-4], TMD films with good scalability, uniformity, crystallinity, and compatibility on suitable substrates remain a significant roadblock to the realization of commercially viable TMD-based electron devices [5, 6]. To mitigate this problem, we investigate a location-on-demand selective growth methodology to realize high-quality TMD layers with consistent layer characteristics. Without any wet or dry transfer process which inevitably leads to material degradation, the TMD growth is catalyzed directly at the desired channel area with improved time and cost effectiveness. Specifically, we take 2D MoS_2 as an example, and exploit patterned MoO_3 seedings to enable wafer-scale location-on-demand selective growth of MoS_2 arrays for transfer-free FET applications. We directly observe the chemical vapor deposition (CVD) growth of MoS_2 on micrometer scale within a few seconds (a high growth rate of 0.5 μm/s), and investigate the impacts of the dielectric interfaces (MoO_3, SiO_2, and Al_2O_3) on the synthetic MoS_2 FET performance in terms of on-current density ($J_{D,on}$), field-effect mobility (μ_{FE}), on/off ratio, contact resistance (R_C), transfer length (L_T), and Schottky barrier height (SBH) etc. We find that the polycrystal MoS_2 with MoO_3 dielectric interface has comparable and even superior performance compared to other MoS_2 FETs using selective and non-selective growth, despite the presence of rich grain boundaries (GBs). Our work demonstrates the wafer-scale, transfer-free, location-on-demand selective growth of 2D TMD, which can significantly ease the integration with Si-CMOS processing and paves a feasible way for realizing 2D semiconductor implementation.

II. Device Fabrication and Measurement

MoS_2 growth and its dynamic evolution. First, the MoO_3 seedings were patterned using electron-beam lithography (EBL) and sputtered on p-Si substrates (1-10 Ω·cm) with 285 nm SiO_2 dielectric. Next, the sample was loaded in a two-zone CVD furnace for controlled sulfurization, and the MoS_2 thin films were formed at the seeding areas, which spatial morphology and crystallinity can be manipulated by the CVD synthetic parameters. Especially, the dynamic evolution of MoS_2 growth was monitored by a micro-chamber CVD with time-resolved in-situ microscopy, as shown in Fig. 1(a) and (b). A single crystal MoS_2 domain, up to 10 μm in size, can be grown within ~20 s, indicating a high growth rate of 0.5 μm/s.

Controlled sulfurization. MoO_3 is not only a Mo precursor but also a high-k dielectric (k = ~35) [7]. In principle, MoO_3 can serve as a better dielectric interface to screen the charge impurities and boost the charge transport. Compared to a full sulfurization process which leads to a single crystal "triangle" domain formation, a limited sulfurization process only at the MoO_3 surface creates polycrystal MoS_2 around the MoO_3 seedings, as shown in Fig. 1(c). This "circular" MoS_2 pattern suggests isotropic homogeneity of MoS_2 growth on SiO_2, and the remaining MoO_3 layer provides a native high-k dielectric interface, benefiting the charge transport of the synthetic MoS_2 on top of it. Confocal Raman and photoluminescence (PL) spectroscopies also confirm the signature modes of MoS_2 on both MoO_3 and SiO_2 dielectrics, suggesting excellent uniformity of localized homogeneity.

Large-scale location-on-demand selective growth of MoS_2. Scanning electron microscopy (SEM) and Raman spectroscopy were performed to confirm the MoO_3 seedings (e.g., an array of 5×5 μm^2 squares) and the corresponding MoS_2 growth (e.g., an array of circular MoS_2 with a diameter of ~14 um) across a 1×1 cm^2 area, as shown in Fig. 1(d-g). Atomic force microscopy (AFM) confirms the presence of a non-sulfurized MoO_3 layer which is less than 10 nm, as shown in Fig. 1(h). Excellent uniformity of the synthetic MoS_2 size across a 1×1 cm^2 area was evaluated, as shown in Fig. 1(i). With the increased seeding density (i.e., the miniaturized spacing distance down to 1 μm), the selective MoS_2 growth is still consistent, as shown in Fig. 1(j). All these characterizations suggest excellent reproducibility and uniformity of selective growth for the ease of integration, especially at a large scale.

Device fabrication and measurement. On top of the as-grown MoS_2, Bi/Au electrodes (20 nm/50 nm) were patterned by EBL and sputtered to form a back-gate FET configuration, and Al_2O_3 was deposited by atomic layer deposition (ALD) to form a top dielectric in a top-gate FET configuration. Electrical and photoresponse characterizations were performed using a semiconductor parameter analyzer, a temperature-variable vacuum probe station, and an integrated laser system. With the grounded source, drain current (I_D) was measured as a function of the applied drain, back-gate, and top-gate voltages (V_D, V_{BG}, and V_{TG}), and was normalized in current density ($J_D = I_D/W$ where W is the channel width) for comparison.

III. Device Results and Discussion

MoO_3 dielectric interface. A comparative investigation is performed for the synthetic MoS_2 FETs with the MoO_3 and SiO_2 dielectric interfaces, in terms of output and transfer characteristics (J_D-V_D and J_D-V_{BG}), and a statistical analysis of $J_{D,on}$, subthreshold swing (SS), μ_{FE}, threshold voltage (V_{th}), hysteresis window (ΔV_{th}), and on/off ratios, as shown in Fig. 2(a-j). The linear J_D-V_D characteristics suggest Ohmic contact for both FET types. With the identical synthetic process and device geometry, the MoO_3 interface provides comparable FET performance metrics, and more importantly, in much narrower distributions. On the wafer scale, the MoS_2 arrays formed by the controlled sulfurization at on-demand locations are still reproduceable, which gives the average $J_{D,on}$ of 2 μA/μm, μ_{FE} of 5 cm^2/Vs, and on/off ratio exceeding 10^5, as shown in Fig. 2(k-m). By optimizing the synthetic parameters, our best MoS_2 FET device possesses $J_{D,on}$ of 3 μA/μm, μ_{FE} of 20 cm^2/Vs, and on/off ratio up to 10^6, as shown in Fig. 2(n).

Compared to conventional SiO_2, MoO_3 as both the precursor and native oxide provides a much intimate contact to MoS_2 with less defects, traps, mismatch, strain, or interfacial states, as shown in Fig. 3(a) and (b). These facts are evidenced by the extracted interfacial trap density (D_{it}), as shown in Fig. 3(c). Here D_{it} was calculated from SS [8], and the MoO_3 dielectric interface provides much higher homogeneity compared to the SiO_2 dielectric interface through a wafer-scale statistical analysis. We also fabricated the top-gate MoS_2 FETs, and the synthetic MoS_2 has good interfacial states to enable ALD-produced Al_2O_3 dielectrics, as shown in Fig. 3(d).

Impacts of crystallinity and GBs. Compared to the single crystal triangle MoS_2, the polycrystal circular MoS_2 possesses abundant GBs. To understand the impact of crystallinity and GBs on synthetic MoS_2, we design one MoS_2 FET with the channel being parallel with a GB, and another device with the channel being perpendicular to the same GB, as shown in Fig. 3(e). Our results show comparable FET performance including $J_{D,on}$, on/off ratio, and V_{th}, suggesting negligible impact of GMs on the synthetic MoS_2 in this work.

On-demand geometric manipulation of MoS_2 growth and metal contact improvement. Owing to the well-controlled growth, we can define the geometry of as-grown MoS_2 in arbitrary shapes without any lithography and etching process. For example, we create a long MoO_3 ribbon (5 μm × 200 μm) for transmission line measurement (TLM). The MoS_2 growth is remarkably uniform along the seeding pattern, as shown in Fig. 4(a). It is intriguing that the more involvement of MoO_3 interface, in contrast to SiO_2 interface, can lead to greater improvement of the metal contact condition such as the lowering of R_C and T_L, but keep the MoS_2 channel resistance (R_{CH}) intact, as shown in Fig. 4(b-d). The SHB, evaluated from a temperature-variable measurement, suggests a flat-band barrier height of 12 meV with the MoO_3 interface, which is much lower than that with the SiO_2 interface (60 meV) and is consistent with the improvement of metal contact condition. Moreover, the extraction of R_C in this work was the first report of any selectively grown MoS_2, thanks to the ease of the MoS_2 geometric definition. The metrics such as R_C of ~200 kΩ·μm and L_T of ~0.1 μm in this work can be further improved by doping, phase transition, and semimetal contact which are well explored for the non-selectively grown MoS_2.

Photoresponse. The back-gate MoS_2 FET arrays can act as 2D phototransistors, and their photoresponsive performance is evaluated, as shown in Fig. 5. Both the power-dependent static photocurrent (PC) generation and time-resolved high-speed photo-switching dynamics suggest excellence photoresponse of MoS_2 FETs in the visible spectrum.

Performance benchmarking. The MoS_2 FETs using location-on-demand direct CVD synthesis in this work were benchmarked with other MoS_2 FETs using various selectively grown techniques (i.e., patterned seeding, plasma treatment, and laser annealing), as shown in Fig. 6(a-d). Our work shows superior metrics in $J_{D,on}$, μ_{FE}, and on/off ratios, owing to the excellent MoO_3 dielectric interface. Meanwhile, our devices were also compared with other MoS_2 FETs using non-selective grown strategies on special substrates (e.g., HfO_2, sapphire, Au, polymer, and glass), as shown in Fig. 6(e). Our wafer-scale location-on-demand MoS_2 FETs don't require any transfer process, and the performance metrics, such as μ_{FE} and on/off ratios, are comparable with the state-of-the-art MoS_2 FETs.

IV. Conclusion

In this work, we presented wafer-scale MoS_2 FET arrays using transfer-free location-on-demand selective growth. This technique allows time- and cost-efficient synthesis (0.5 μm/s) of 2D semiconductor channel arrays at designed locations, and the controlled sulfurization creates MoO_3 dielectric interfaces to enhance FET performance. Our devices show comparable and even superior performance, such as $J_{D,on}$, μ_{FE}, and on/off ratios, compared to other MoS_2 FETs using selective and non-selective growth, and demonstrate great potential to ease the integration of 2D semiconductors for various electron devices.

Acknowledgment

The authors acknowledge support from the SUNY Applied Materials Research Institute (SAMRI) and the National Science Foundation (NSF) under Award ECCS-1944095.

References

[1] S. Das et al., *Nat. Electron.*, 4, 786 (2021).
[2] M. Liu et al., *ACS Nano*, 15, 5762 (2021).
[3] H. N. Jaiswal et al., *Adv. Mater.*, 32, 2002716 (2020).
[4] M. Liu et al., *IEEE IEDM*, 251 (2020).
[5] N. Briggs et al., *2D materials*, 6, 022001 (2019).
[6] Y. Zhang et al., *Adv. Mater.*, 31, 1901694 (2019).
[7] B. Holler et al., *Adv. Electron. Mater.*, 6, 2000635 (2020).
[8] C. J. McClellan et al., *ACS Nano*, 15, 1587 (2021).

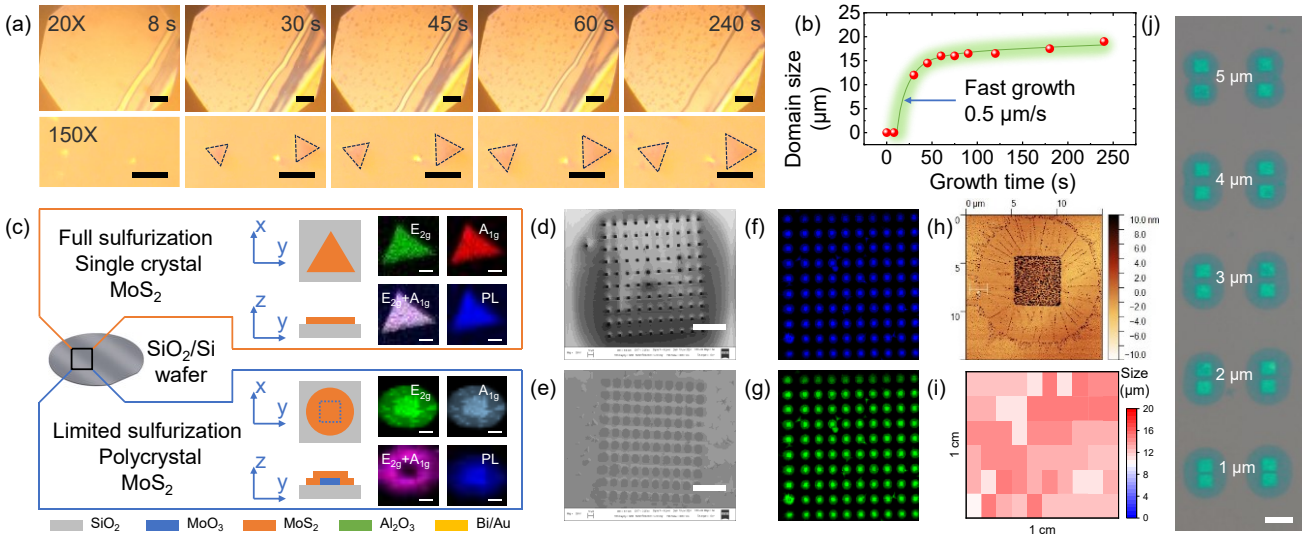


Fig. 1 Synthesis and characterization of wafer-scale location-on-demand MoS_2 growth on SiO_2/Si substrates. (a, b) Time-resolved in-situ microscope images to visualize the MoS_2 growth evolution, and the extracted growth rate up to 0.5 μm/s. Scale bar: 50 μm (top) and 20 μm (bottom). (c) Comparison of single crystal "triangle" MoS_2 and polycrystal "circular" MoS_2 synthesized by full sulfurization and limited sulfurization, respectively, and the corresponding Raman/PL spectroscopy mapping. Scale bar: 5 μm. (d, e) SEM images of a MoO_3 seeding array before growth and a MoS_2 array after growth. Scale bar: 60 μm. (f, g) E_{2g} and A_{1g} Raman spectroscopy mapping for the large-scale synthetic MoS_2 arrays. (h) AFM mapping of a circular MoS_2 with the MoO_3 seeding at the center. (i) Spatial uniformity of the MoS_2 sizes across a 1 × 1 cm^2 area. The average size is about 14 μm. (j) Microscopy image of the circular MoS_2 thin films which merges as the spacing distance of the MoO_3 seedings scales down from 5 to 1 μm. Scale bar: 10 μm.

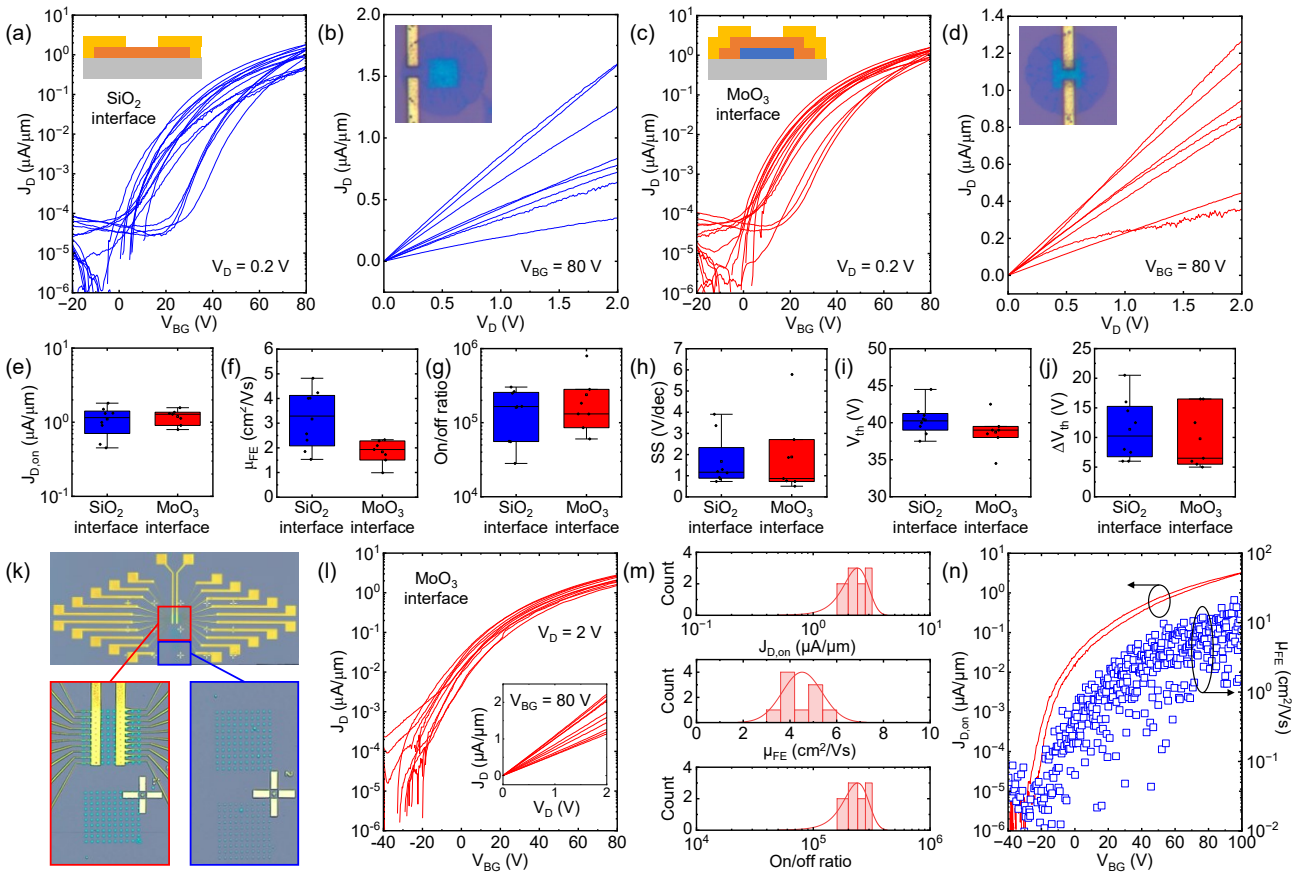


Fig. 2 Comparison of synthetic MoS_2 FETs with SiO_2 and MoO_3 dielectric interfaces. (a-d) Output and transfer characteristics of the MoS_2 FETs with SiO_2 and MoO_3 dielectric interfaces. Insets: The corresponding cross-sectional schematics and microscope images of the devices. (e-j) Statistical analysis of the MoS_2 FETs with SiO_2 and MoO_3 dielectric interfaces, including $J_{D,on}$, μ_{FE}, on/off ratio, SS, V_{th}, and ΔV_{th}. (k-m) Microscope images and transfer characteristics of the wafer-scale MoS_2 FET arrays with MoO_3 dielectric interfaces, and the corresponding statistical analysis including $J_{D,on}$, μ_{FE}, and on/off ratio. Inset of (l): The corresponding output characteristics. (n) The best back-gate MoS_2 FET with MoO_3 dielectric interfaces possesses $J_{D,on}$ of 3 μA/μm, μ_{FE} of 20 cm^2/Vs, and an on/off ratio up to 10^6.

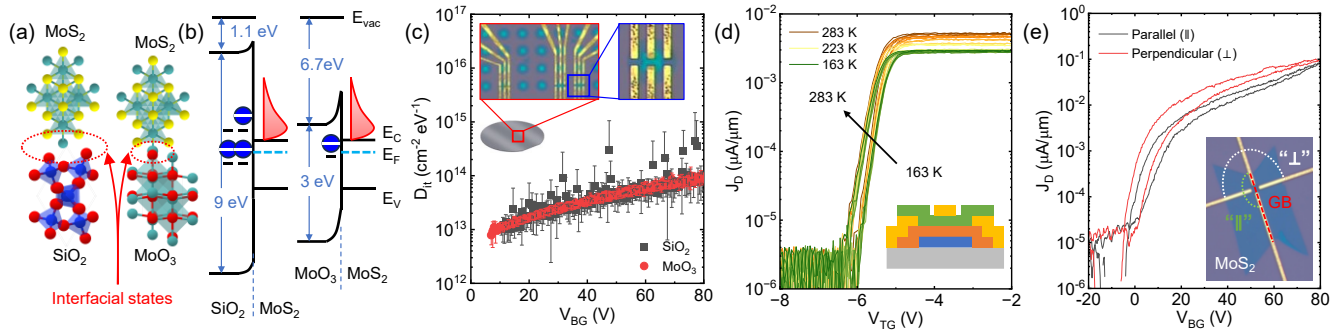


Fig. 3 Impacts of dielectric interfaces and GBs on MoS_2 FET performance. (a-c) Schematics, energy band diagrams, and the extracted D_{it} as a function of V_{BG} for MoS_2/SiO_2 and MoS_2/MoO_3 interfaces. Here E_C, E_V, E_F, and E_{vac} are the conduction band minimum, valence band maximum, Fermi level, and vacuum level, respectively. Inset of (c): Microscope images of wafer-scale MoS_2 FET arrays. (d) Temperature-variable transfer characteristics of a top-gate MoS_2 FET with the MoO_3 bottom dielectric and Al_2O_3 top dielectric. (e) Comparison of MoS_2 FETs with the channels parallel with and perpendicular to a GB.

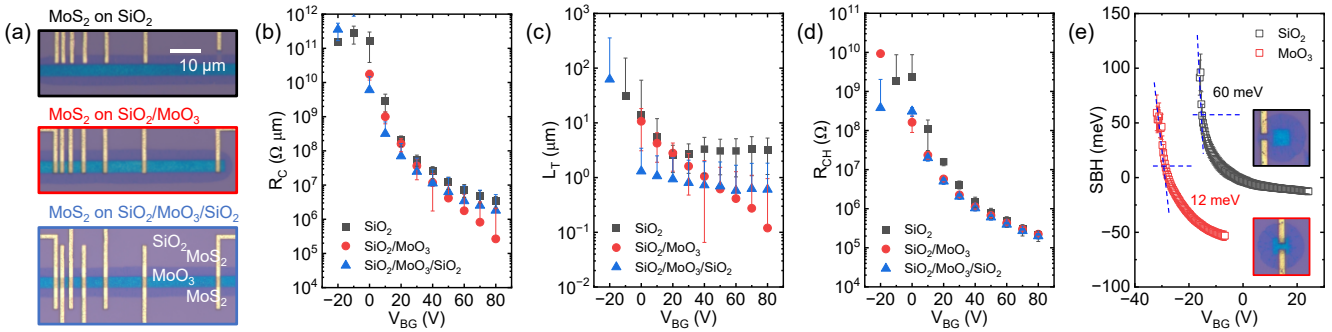


Fig. 4 Metal contacts with SiO_2 and MoO_3 dielectric interfaces. (a-d) Microscope images of MoS_2 TLM devices with different SiO_2 and MoO_3 dielectric interfaces, and the extracted R_C, L_T, and R_{CH}. (e) The extracted SBH as a function of V_{BG}. Inset: Microscope images of the corresponding MoS_2 FETs.

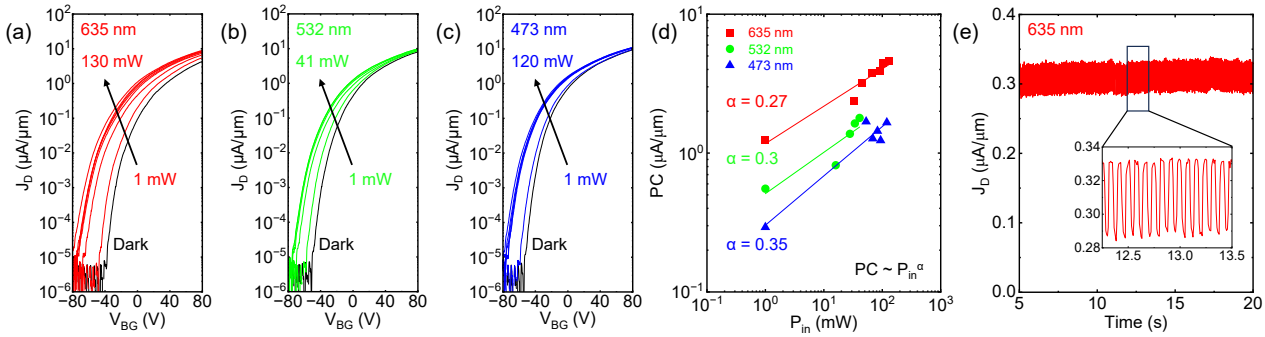


Fig. 5 Photoresponse of MoS_2 FETs as photodetectors. (a-c) Transfer characteristics of a MoS_2 FET under red, green, and blue laser illumination. (d) The extracted PC as a function of the input power (P_{in}). (e) Time-resolved high-speed photo-switching characteristics under the red laser illumination.

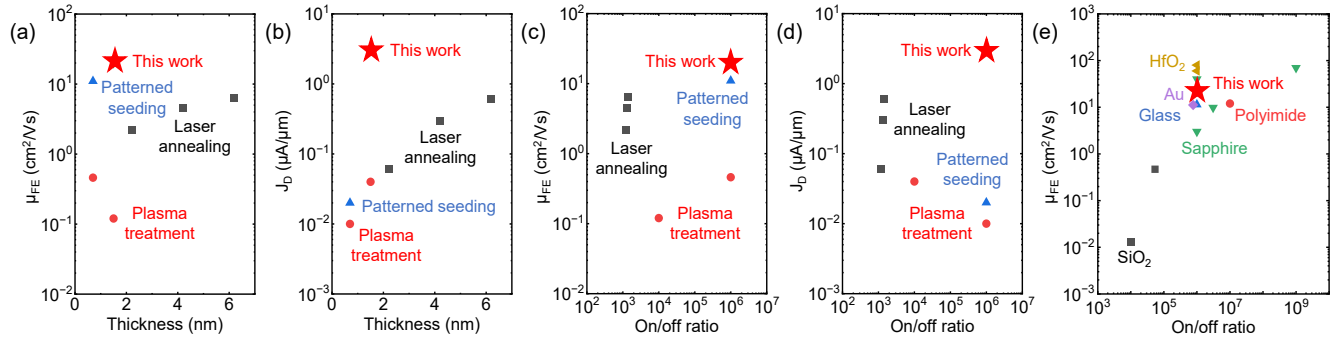


Fig. 6 Benchmarking of the MoS_2 FETs in this work with other state-of-the-art synthetic MoS_2 FETs in terms of $J_{D,on}$, μ_{FE}, on/off ratio, and thickness. (a-d) Benchmarking with other location-on-demand selectively synthesized MoS_2 FETs on SiO_2 substrates. (e) Benchmarking with other MoS_2 FETs grown on special substrates (e.g., SiO_2, HfO_2, sapphire, Au, polymer, and glass) which have no location-on-demand selectivity or require a wet/dry transfer process. The references include: S. Park et al., *ACS Nano*, 14, 8485 (2020); X. Chen et al., *Nanoscale*, 8, 15181 (2016); H. J. Kim et al., *Small*, 13, 1702256 (2017); G. H. Han et al., *Nat Commun.*, 6, 6128 (2015); N. B. Shinde et al., *ACS Appl. Nano Mater.*, 3, 7371 (2020); T. W. Kim et al., *Nanotechnology*, 28, 18LT01 (2017); J. Mun et al., *ACS Appl. Electron. Mater.*, 1, 4, 608 (2019); P. Yang et al., *Nat Commun.*, 9, 979 (2018); H. Yu et al., *ACS Nano*, 11, 12001 (2017); Q. Wang et al., *Nano Lett.*, 20, 7193 (2020); Y. F. Lim et al., *ACS Nano*, 12, 1339 (2018); B. M. Bersch et al., *2D Mater.*, 4, 025083 (2017); P. Yang et al., *ACS Nano*, 14, 5036 (2020); K. S. Kim et al., *Nature*, 614, 88 (2023).

Ultra-low power cryogenic field effect transistor utilising high mobility compressively strained germanium on silicon

M. Myronov[1], P. Waldron [2], and S. Studenikin[2]
[1]Physics Department, The University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK, e-mail: M.Myronov@warwick.ac.uk
[2]National Research Council of Canada, 1200 Montreal Rd., Ottawa, Ontario K1A 0R6, Canada

***Abstract*—**An ultra-low power cryogenic field effect transistor (Cryo-FET), utilizing a high mobility compressively strained germanium on silicon (cs-GoS) material platform, is presented. The device demonstrates beyond state of the art electrical characteristics at cryogenic temperature of 4.2 K, including superior stability, absence of any hysteresis in I-V characteristics, low subthreshold swing, small leakage current, and very low power dissipation in pW range suitable for building large cryo-electronic circuits containing over 1 million transistors dissipating <100μW and still being withing the cooling power budget of regular dilution refrigerators operating down to <100mK. Our findings highlight the potential of cs-GoS Cryo-FETs for advanced and scalable cryogenic applications such as quantum and classical computing, and deep space exploration.

I. Introduction

The growing demand for ultra-low power and high-performance electronic devices at cryogenic temperatures is driven by advancements in quantum computing, space technology, and cryogenic sensing.[1,2] Silicon, the conventional semiconductor on the market, commonly used in electronics faces significant challenges in maintaining required performance and efficiency at these temperatures, in particular, due to carrier freeze-out effect of dopants, charge instabilities, and large heat dissipation, which is not compatible with cryogenic instrumentation limited by the fundamental thermodynamic laws. This paper demonstrates the unique potentials of emerging cs-GoS material platform for Cryo-FET and spin qubit devices applications, building on recent research advancements [3, 4] that highlight the material's superior properties. The cs-GoS grown on full size silicon wafers, up to 200 mm diameter, is fully compatible with silicon foundries and offers the unique combination of properties that are important for quantum and cryogenic electronics applications.[3] The most important properties include the record high carrier mobility of holes and favorable band structure. The compressive strain in germanium enhances its hole mobility up to 4.3 x10^6 $cm^2V^{-1}s^{-1}$ and reduces the hole's effective mass, m*, down to 0.035m_0, making it an ideal candidate for low-temperature electronics.[4] In particular, smaller m* results in larger hole de-Broglie wavelength that makes electronic devices less sensitive to fluctuations and more reproduceable.[3]

II. Materials Synthesis

For the reported research, undoped cs-GoS heterostructures were grown by reduced pressure chemical vapor deposition (RP-CVD) on a relaxed $Si_{0.15}Ge_{0.85}$ buffer on a standard Si(001) wafer of 150 mm diameter. A schematic cross section of the heterostructure and fabricated FET-like gated Hall bar, with its source (S) and drain (D) ohmic contacts and gate (G) stack is shown in Fig. 1. All epilayers were intentionally undoped. Accumulated by the negative gate voltage, holes are confined in the 30 nm thick undoped and compressively strained Ge (cs-Ge) quantum well (QW), acting as an active p-channel for mobile holes, positioned ~300 nm below the surface.

III. Devices Microfabrication

Double-gated Hall bars, enabled to work in a Cryo-FET mode, were fabricated using standard UV lithography, dry etching and thin film deposition techniques. Fig. 2 shows an optical image of the device in a shape of a gated Hall bar with its channel oriented along the <110> in-plane crystallographic direction defined by the mesa structure etched in a Cl_2/Ar plasma. The Hall bar's channel width is 100 μm and the distance between source and drain contacts is 1000 μm. In order to reach an ultra-low power dissipation, we employ a high-quality undoped cs-GoS material stack, which is naturally not conductive at cryogenic temperatures. Therefore, to replace commonly used doping technology, we introduce a specially designed S and D accumulation gates (Figs. 1-2) allowing us to generate mobile carriers in the contact regions next to the active cs-Ge channel, as shown in Fig.1.

The alloyed AlSiGe ohmic contacts were prepared by evaporating a 120 nm thick Al film and annealing it at ~275 °C in N_2 ambient for 30 min. The injection contacts operating in the enhancement mode show low resistivity and excellent linear ohmic behaviour at cryogenic temperatures. Fig. 3 shows activation characteristics of the S and D contacts separately, with all other contacts being grounded. The S-D contacts accumulation gate is isolated from the contact metallisation and the Schottky gate by a 50 nm AlO_3 dielectric deposited by Atomic Layer Deposition at 200 °C. Superior reproducibility of contacts is evident from Fig.3 with a threshold voltage V_{ACC}=−230 mV. Further measurements of Cryo-FET characteristics are performed at V_{ACC} = −350 mV. The top

accumulation Schottky gate is made of 20 nm Ti followed by a 200 nm Au layer.

IV. Electrical Characterization

The Cryo-FET in a shape of a gated Hall bar with additional potential probes (2-3 in Fig.2) allows us to measure intrinsic material's properties like free-carrier mobility, carrier density, their mean free path, etc. [4], which are important for understanding and modeling these new devices. At the same time, we use our device in FET mode in order to measure its input and output characteristics. The electrical performance of the Cryo-FET device was carried out at temperature T=4.2 K.

There is a common issue of undesirable charges at the interface of a semiconductor and a dielectric, which leads to large instabilities and threshold voltage shifts in gated devices, including FETs.[4-6] In order to achieve superior gate control and stability we avoid using dielectrics in the Cryo-FET channel region and, instead, introduce the Schottky gate for cs-GoS devices, for the first time. Fig. 4 shows an I-V characteristic of Schottky gate current I_G versus gate voltage V_G in forward direction. The minimum leakage current is below the detectable limit of the experimental measurement setup. Some small leakage current appears at V_G below −150 mV, within a few pA's, and starts growing exponentially for voltages below −220 mV, following expected Schottky contact behavior in enhancement mode. It should be noted, this is an extremely low value for a relatively large area of 0.1 mm^2 of our test device. Also, we observed very reproduceable sharp peaks, evident in Fig. 4, for 4 back-and-forth sweeps, which may be due to resonant tunneling through deep level states. Very high hole mobility of 1.5×10^6 cm^2 V^{-1} s^{-1} is measured at V_G= −350 mV, as seen in Fig. 5. As a consequence, holes mean free path, shown in Fig. 6, reaches 8 μm. It indicates that a Cryo-FET device with a gate length below this value will operate in a ballistic regime, which may lead to even lower heat dissipation.

A 2D color-map plot of typical input characteristics of the Cryo-FET is shown in Fig. 7. $I_{SD} - V_{SD}$ line-traces at different V_G, extracted from this map, are presented in Fig 8. Due to very high mobility, I_{SD} saturates very fast at very low $V_{SD} \sim -10$ mV. Figure 9 shows a hysteresis check of the Cryo-FET characteristics in semi-logarithmic scales, and Fig. 10 in linear scales. These results indicate the superior performance of the Cryo-FET device: no observable hysteresis, very low threshold voltage, V_{TH} = −15 mV, and very low sub-threshold swing (SS) = ~3 mV/dec. The Cryo-FET reveals reliable gate control with undetectable minimum leakage currents. The off current of the FET is below 1 pA, which is limited by the cryogenic measurements experimental setup including electronics, circuits, cables and wiring. It is necessary to note that no temperature rise was observed during the measurements, discussed above, indicating an ultra-low power dissipation of the Cryo-FET device.

V. Performance Comparison with Conventional Materials

A comparative analysis was conducted between cs-GoS Cryo-FETs and FETs based on traditional semiconductor materials. [2,7] The Cryo-FET shows superior performance thanks to the very high hole mobility and material stack quality resulted in very low subthreshold swing, very low Schottky gate leakage current in the enhancement mode, very low V_{TH}, absolute absence of any hysteretic behavior, and superior stability of all FET characteristics at cryogenic temperatures. The Cryo-FET characteristics were repeatedly obtained during several days and no measurable drift of any characteristic was observed. We estimate that the Cryo-FET can operate in a very low dissipation power regime, with estimated ~50 pW of Joule heat dissipation. It means, an ULSI circuit containing, e.g., one million of such transistors would dissipate just ~50 μW heat power, which is within the cooling power of modern cryogenic-free dilution refrigerators operating down to < 100 mK. These advantages position the cs-GoS as a promising material platform for cryogenic electronic applications. Demonstrated unique properties of the cs-GoS Cryo-FET open up new possibilities for cryogenic classical and quantum electronic systems. Potential applications include low-power quantum computing circuits, cryogenic sensors, and deep space electronics. The high mobility and ultra-low power consumption of cs-GoS FETs are particularly beneficial for high-speed cryogenic electronics where energy efficiency and performance are critical.

VI. Conclusion

This paper presents the development and characterization of ultra-low power and high-performance p-type Cryo-FET based on the new cs-GoS material platform. The exceptional electrical properties of the cs-GoS material stack at cryogenic temperatures make it a very promising candidate for next-generation cryogenic classical and quantum electronics. Future work will focus on further optimization of devices, fabrication technology, and their integration with other cryogenic components such as blocks of qubits performing error corrections or other quantum computing algorithms.

Acknowledgment

The Authors thank the financial support of Quantum Sensing Program of NRC Canada and EPSRC UK.

References

[1] R. Nikandish, E. Blokhina, D. Leipold, and R. B. Staszewski, "Semiconductor Quantum Computing: Toward a CMOS quantum computer on chip," *IEEE Nanotechnology Magazine,* vol. 15, no. 6, pp. 8-20, 2021.

[2] G. Kiene *et al.*, "A 1-GS/s 6–8-b Cryo-CMOS SAR ADC for Quantum Computing," *IEEE Journal of Solid-State Circuits,* pp. 1-12, 2023.

[3] M. Myronov, P. Waldron, P. Barrios, A. Bogan, and S. Studenikin, "Electric field-tuneable crossing of hole Zeeman splitting and orbital gaps in compressively strained germanium semiconductor on silicon," *Communications Materials,* vol. 4, no. 1, p. 104, 2023.

[4] M. Myronov *et al.*, "Holes Outperform Electrons in Group IV Semiconductor Materials," *Small Science,* vol. 3, p. 2200094, 2022 202.

[5] J. Ferrero *et al.*, "Noise reduction by bias cooling in gated Si/Si_x Ge_{1-x} quantum dots," *Applied Physics Letters,* vol. 124, no. 20, 2024.

[6] M. Meyer *et al.*, "Electrical Control of Uniformity in Quantum Dot Devices," *Nano Letters,* vol. 23, no. 7, pp. 2522-2529, 2023/04/12 2023.

[7] X. Xue *et al.*, "CMOS-based cryogenic control of silicon quantum circuits," *Nature,* vol. 593, no. 7858, pp. 205-210, 2021/05/01 2021.

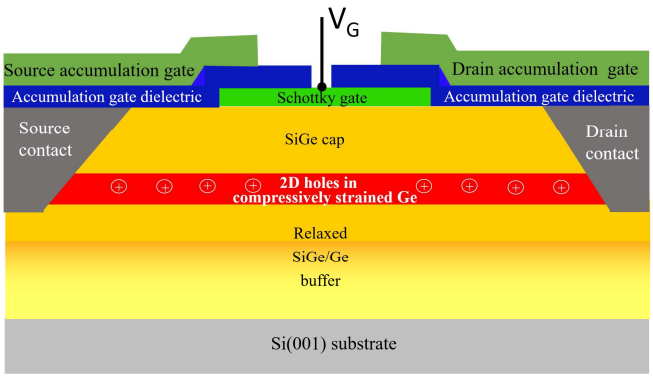


Fig. 1. Cross-section schematic of the Schottky Cryo-FET on undoped cs-GoS epiwafer, employing double-gates technology. The drain and source accumulation gates replace doping or ion-implantation techniques for contact activation, which would deteriorate electrical performance of the device at cryogenic temperatures.

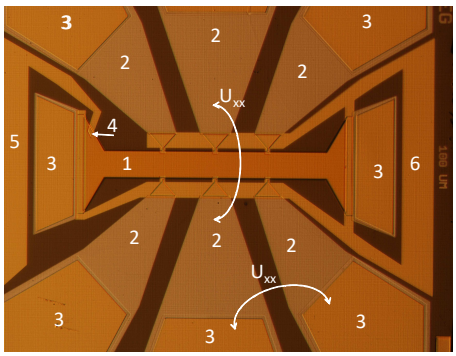


Fig. 2. An optical image of the double-gate Hall bar device which also is used to work as a Cryo-FET. Additional potential probes (2) allow to measure potentials U_{xx} and U_{xy} and extract transport characteristics [4] of the active channel controlled by the Schottky-contact gate (1) in enhancement mode. (3) Ti/Au bonding pads to source-drain and potential probe contacts; (4) via through the gate dielectric to contact Schottky gate; (5) Schottky gate bonding pad; (6) the contacts accumulation gate. The Cryo-FET p-channel width (W = 100 µm) to length (L = 1000 µm) ratio is 0.1.

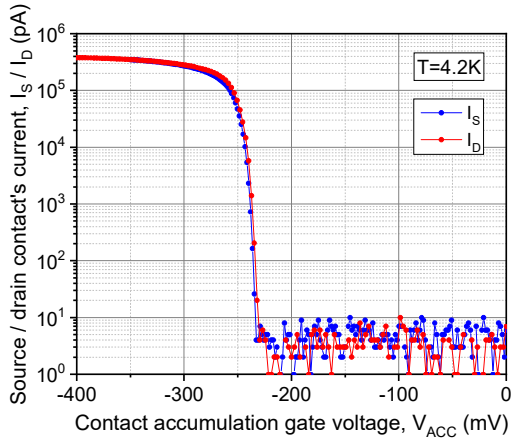


Fig. 3. Source-drain contacts initialization. Individual initialization traces of S and D contacts with all other contacts grounded, Schottky V_G=−100 mV. dielectric thickness 50 nm. In all experiments V_{ACC}= −350 mV.

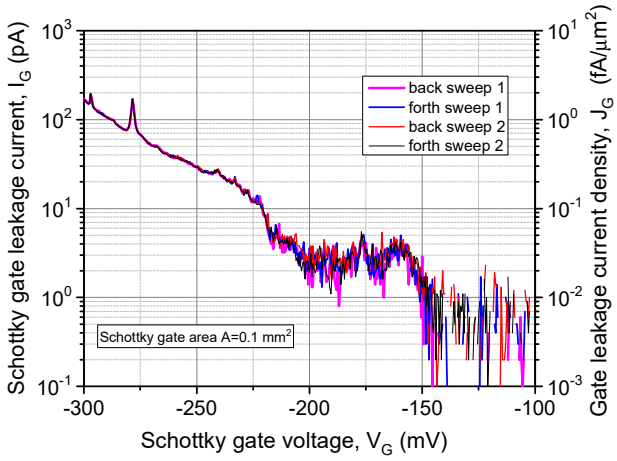


Fig. 4. Schottky-contact gate current in forward bias, so-called the enhancement mode operation. Very small current is detected as an indicator of superior quality of the Schottky contact to the cs-GoS material stack, opening new potentials for cryogenics electronics.

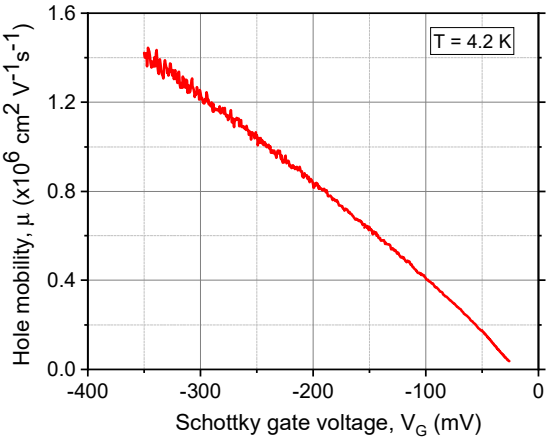


Fig. 5. Hole mobility of free carriers in the p-channel of the Cryo-FET device at T=4.2K.

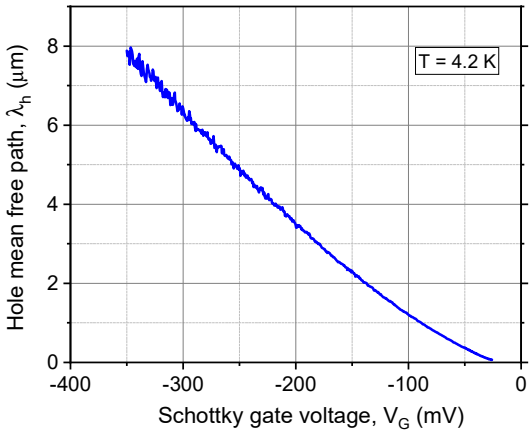


Fig. 6. Mean free path of free carriers as a function of the Schottky gate voltage in forward bias direction, i.e. the enhancement operation mode.

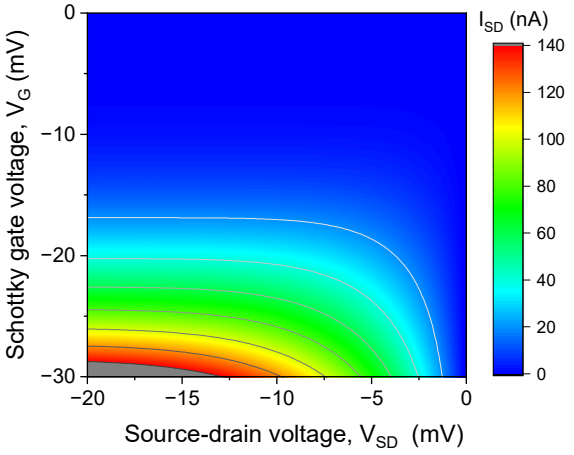


Fig. 7. 2-D color map of the cs-GoS Cryo-FET p-channel enhancement mode source drain current, $I_{SD,}$ characteristics as a function of V_G and V_{SD}, measured at T=4.2K.

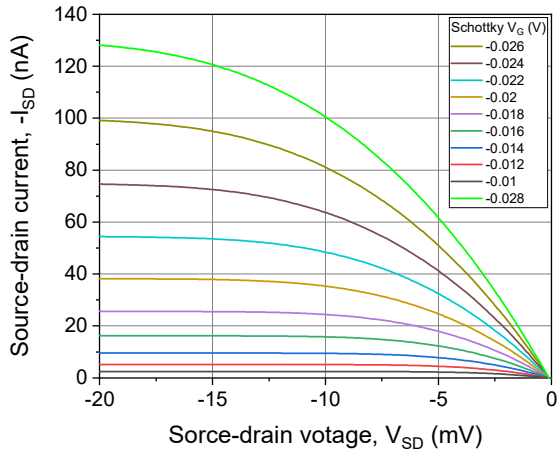


Fig. 8. Typical cs-GoS Cryo-FET p-channel enhancement mode I_{SD} – V_{SD} characteristics at gate voltage, V_G, varied from −10 mV to −28 mV, measured at 4.2 K. I_{SD} increases sharply and saturates at very low V_{SD}, for a given V_G, due to very high hole mobility, shown in Fig. 5.

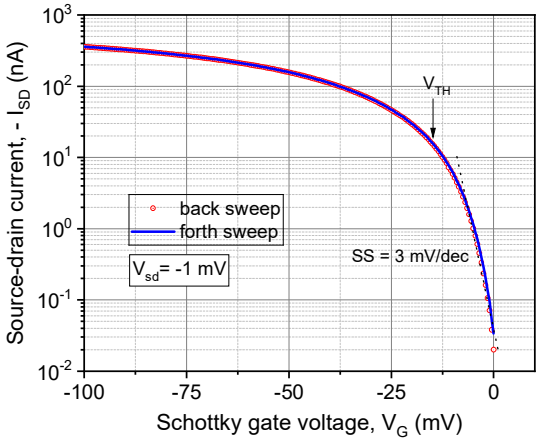


Fig. 9. Typical cs-GoS Cryo-FET p-channel enhancement mode forth and back sweep I_{SD} – V_G characteristics at low drain bias voltage I_{SD} = −1 mV, measured at 4.2 K. I_{SD} is plotted on a logarithmic scale. Both traces lay absolutely on top of each other, i.e., no noticeable hysteresis shifts are detected, indication of the absence of any carrier traps in the whole active Cryo-FET region including bulk and interface. I_{SD} increases sharply due to very high hole mobility, see Fig. 5. The dotted line illustrates the determination of the subthreshold swing SS = 3 mV/dec.

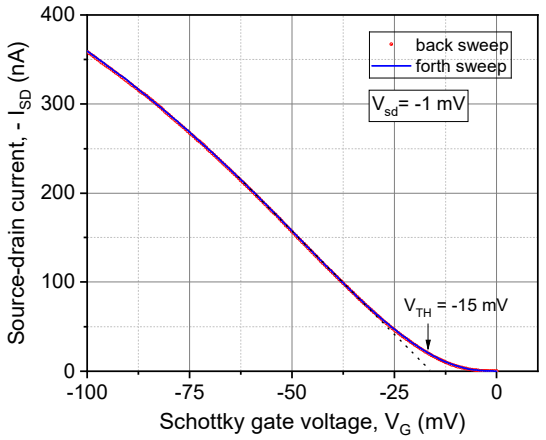


Fig. 10. Typical cs-GoS Cryo-FET p-channel enhancement mode forth and back sweep I_{SD} – V_G characteristics at low drain bias voltage I_{SD} = −1 mV, measured at 4.2 K. I_{SD} is plotted linear scale. No hysteresis is visible. I_{SD} increases sharply due to very high hole mobility, shown in Fig. 5. The dotted line illustrates the determination of the linearly extrapolated threshold voltage V_{TH} = −15 mV. Detailed understanding of the cs-GoS Cryo-FET requires careful modeling and theoretical attention.

2nm Platform Technology featuring Energy-efficient Nanosheet Transistors and Interconnects co-optimized with 3DIC for AI, HPC and Mobile SoC Applications

Geoffrey Yeap, S.S. Lin, H.L. Shang, H.C. Lin, Y.C. Peng, M. Wang, PW Wang, CP Lin, KF Yu, WY Lee, HK Chen, DW Lin, BR Yang, CC Yeh, CT Chan, JM Kuo, C-M Liu, TH Chiu, MC Wen, T.L. Lee, CY Chang, R. Chen, P-H Huang, C.S. Hou, YK Lin, FK Yang, J. Wang, S. Fung, Ryan Chen, C.H. Lee, TL Lee, W. Chang, DY Lee, CY Ting, T. Chang, HC Huang, HJ Lin, C. Tseng, CW Chang, KB Huang, YC Lu, C-H Chen, C.O. Chui, KW Chen, MH Tsai, CC Chen, N. Wu, HT Chiang, XM Chen, SH Sun, JT Tzeng, K. Wang, YC Peng, HJ Liao, T. Chen, YK Cheng, J. Chang, K. Hsieh, A. Cheng, G. Liu, A. Chen, HT Lin, KC Chiang, CW Tsai, H. Wang, W. Sheu, J. Yeh, YM Chen, CK Lin, J. Wu, M. Cao, LS Juang, F. Lai, Y. Ku, S.M. Jang, L.C. Lu

Global R&D Center, Taiwan Semiconductor Manufacturing Company (TSMC), Hsinchu, Taiwan, R.O.C.

Abstract

A leading edge 2nm CMOS platform technology (N2) has been developed and engineered for energy-efficient compute in AI, mobile and HPC applications. This industry-leading N2 logic technology features energy-efficient gate-all-around nanosheet transistors, middle-of-line and backend-of-line interconnects with densest SRAM macro of ~38Mb/mm^2. N2 delivers a full node benefit from previous 3nm node [4] in offering 15% speed gain or 30% power reduction with >1.15x chip density increase. N2 platform technology, equipped with new Cu scalable RDL, flat passivation and TSVs, co-optimizes holistically with 3DFabric™ technology enabling system integration/scaling for AI/mobile/HPC product designs. N2 successfully met wafer-level reliability requirements and passed 1000hrs HTOL qual with high yielding 256Mb HC/HD SRAM, and logic test chip (>3B gates) consisting of CPU/GPU/ SoC blocks. Currently in risk production, N2 platform technology is scheduled for mass production in 2H'25. N2P, 5% speed enhanced version of N2 with full GDS compatibility, targets to complete qualification in 2025 and mass production in 2026.

Introduction

Advanced CMOS technology has been the key enabler for semiconductor product innovations. Since the generative AI break-through moment in Q1'23, AI together with 5G-advanced mobile and HPC have ignited the industry with an insatiable appetite for best-in-class advanced energy-efficient logic technology [1]. Our industry leading 2nm platform technology (N2) is one such advanced logic technology. This paper describes the state-of-art N2 technology successful transition into NS platform technology and acceleration of >140x energy-efficient compute from 28nm to N2 as shown in Fig. 1. We also present system technology co-optimization (STCO) innovation in design rules, standard cell, SRAM and interconnects co-optimization with 3DFabric™. N2 technology has been verified on our development/qual test vehicle. N2 met all the wafer-level reliability requirements and completed the full 1000hours HTOL qualification with high yielding 256Mb HD/HC SRAM and logic test chip (>3B gates). Now in risk production, N2 is on track for mass production in 2H'25. N2P with 5% additional speed and full GDS compatibility targets to complete qual in 2025 and mass production in 2026.

N2 NanoFlex™ [3] Technology Architecture

The N2 2nm platform technology is defined and developed to meet PPACt (Power, Performance, Area, Cost, and Time-to-market) [2]. STCO is emphasized with smart scaling features instead of brute-force design rule scaling which drastically increases process cost and inadvertently causes critical yield issues. Extensive STCO coupled with smart scaling of major design rules (e.g., gate, nanosheet, MoL, Cu RDL, passivation, TSVs) was performed in optimizing this 2nm technology to achieve target PPA. This development also involves co-optimization with 3DFabric™ SoIC 3D-stacking and advanced packaging technology (INFO/CoWoS variants) thereby accelerating system integration/scaling for AI/mobile/HPC product designs.

N2 NanoFlex™ [3] standard cell innovation offers not only nanosheet width modulation but also the much-desired design flexibility of the multi-cell architecture. N2 short cell lib for area and power efficiency. Selective use of tall cell lib lifts frequency to meet design target Combining with six-Vt offerings spanning across 200mV, N2 provides unprecedented design flexibility to satisfy a wide spectrum of energy-efficient compute applications at the best logic density. N2 delivers a full node scaling with attractive PPA values at projected cost and time-to-market: ~15% speed gain or ~30% power reduction with >1.15x chip density scaling (Fig. 2-3).

Energy-efficient Nanosheet Transistors, MoL and BEOL Interconnects

Multiple generations of Si FinFet with fin depopulation were in use from 16nm to 7nm (2-fin) node. High mobility channel transistors with industry-first zero-thickness dipole based true multi-Vt (7-Vt), cut metal-gate and gate-contact over-active innovations extended FinFet architecture into N5 node [2]. FinFlex™ DTCO coupled with other key enhancements successfully extracted another full node PPA benefits in N3, last FinFet node [4].

N2 platform technology successfully completes the transition from FinFet into energy-efficient nanosheet technology. Figure 4 shows optimized nominal gate-length NS transistors with excellent DIBL and subs-threshold swings. Long gate-length NS transistors achieve near-ideal 60.1mV/dec swings. Fig. 5 shows the six-Vt's ranging from the extreme low-Vt to standard-Vt in ~200mV span for N2 N/P FETs. Si data is very close to matching ring speed@standby-power at all six Vts. This multi-Vt capability is enabled with 3rd-generation (since N5) dipole-based multi-Vt integration with both n-type and p-type dipoles.

Much process and device enhancements are focused on engineering not just the transistor drive currents through sheet interface/thickness, junction engineering, dopant diffusion/activation and stress engineering, but more on Ceff reduction to drive best-in-class energy efficiency. All these enhancements lead to much improved I/CV speed gain of 70% and 110% respectively for NS N/P FETs. N2 nanosheet technology exhibits substantially better Perf/Watt than FinFET at low Vdd range of 0.5V-0.6V (Fig. 7). Emphasis is placed on low Vdd perf/watt uplift through process and device continuous improvements resulting in 20% speed gain and 75% lower stand-by power at 0.5V operation. N2 NanoFlex coupled with multi-Vt provides unprecedented design flexibility to satisfy a wide spectrum of energy-efficient compute applications at the most competitive logic density.

Overall technology energy efficiency and performance are also critically dependent on MoL, backend and far-backend interconnects. With innovative materials and processing, VG Rc reduces significantly by 55% with barrier-less all-tungsten MoL. The low resistance MoL combined with capacitance reduction features achieve a total of ~6.2% INV D4 ring oscillator speed gain (Figure 8). Optimized M1 with novel 1P1E EUV patterning led to close to 10% std cell capacitance reduction and a saving of multiple EUV masks. Substantial My RC and Vy Rc reductions seen on the tightest 193i 1P1E workhouse metal/via layers (Fig. 12). In summary, N2 MoL and BEOL RC reduce by ~>20% contributing significantly to energy-efficient compute.

Seamless Integration with 3DFabric Tech

This 2nm platform technology, including the new Cu RDL with flat passivation and TSVs, co-optimizes holistically with 3DIC enabling system integration/scaling for AI/mobile/HPC product designs (Fig. 11-12). Attention is paid to optimize materials and processing in backend/far-backend for global warpage and local planarity for robust integration with 3D stacking. N2 also optimizes pTSV/sTSV (for power/signal) in terms CD/pitch/density for F2F/F2B stacking with SoIC bond pitch scaling from 9µm/6µm down to 4.5µm.

SRAM, Logic Test Chip and Qual/Reliability

For advanced nodes, SRAM bitcell scaling has become a challenge. With N2 NanoFlex and improved on-off current, DTCO is employed to maximize #bitcell/bitline, bitline loading and SRAM peripheral layout efficiency resulting in densest 2nm SRAM macro density ~38Mbmm^2 (Fig. 13). N2 HC/HD pull-down Nfet with better Vt-sigma than FinFet resulting in ~20mV lower HC Vmin and 30~35mV lower HD Vmin (Fig. 14). HD 256Mb SRAM shmoo plot in Fig. 15 illustrates full read and write down to ~0.4V. With innovative well engineering and junction isolation, N2 has better latch-up trigger voltage for both logic and SRAM than FinFet (Fig. 15). Higher Vtrig in N2 leads to additional logic density and more effective DVS screening for product quality. N2 test chip demonstrates healthy CPU/GPU functionality and passed the GPU Vmin-power spec shown in Fig. 16.

N2 256Mb HC/HDSRAM consistently demonstrated healthy defect density resulting in >80% / >90% avg/peak yields (w/o repairs). Fig. 20 shows 256Mb SRAM passed 1000hour HTOL qualification with ~110mV margin.

Additional HPC features such as super high performance MiM (SHP-MiM) with ~>200fF/mm2 capacitance density is offered for higher Fmax by minimizing transient drooping voltage. High-speed SerDes test chip also demonstrated fully functioning 14Gb/s LPDDR6 and 10Gb/s HBM3E interfaces.

References:
[1] Y.J. Mii, IEDM, plenary, 2024.
[2] G. Yeap et al., IEDM, 2019.

[3] "TSMC 30th North America Tech Sym. with Innovations Powering AI with Silicon Leadership," TSMC PR, 2024/04/24
[4] S.-Y. Wu et al., IEDM, 2022.

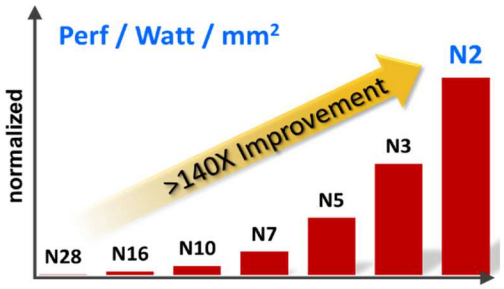


Fig.1 N2 NanoFlex™ accelerates energy-efficient compute

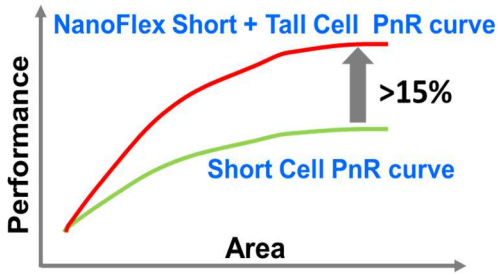


Fig.2 N2 NanoFlex innovation modulates NS width for best PPA. Short cell library for area and power efficiency. Selective use of tall cell library lifts frequency to design target

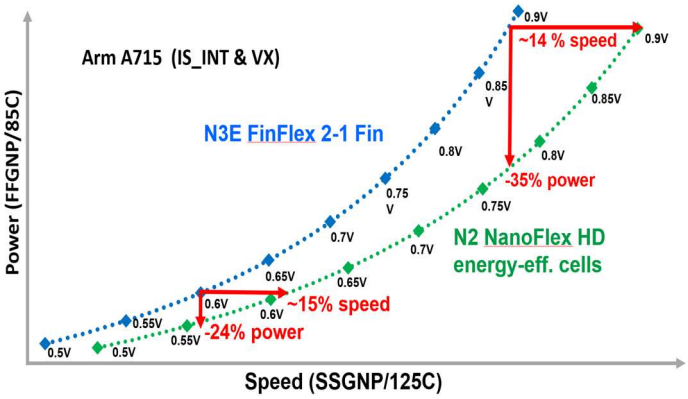


Fig.3 N2 NanoFlex HD cells gain 14~15% speed@power vs. N3E FinFlex 2-1 fin cell across Vdd range: 35% power saving at higher voltage and 24% power saving at lower voltage

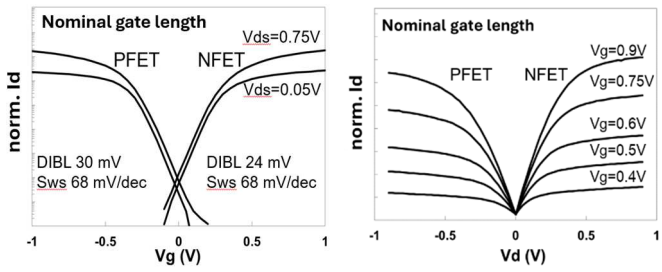


Fig.4 N2 gate & drain characteristics w/ excellent DIBL/Sws

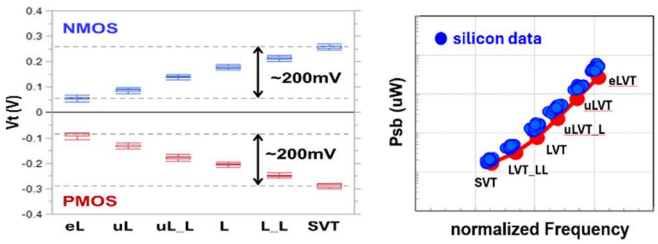


Fig.5 Six-Vt's with ~200mV range for low leakage and high-perf. optimization. Si data closes to matching ring speed @standby-power

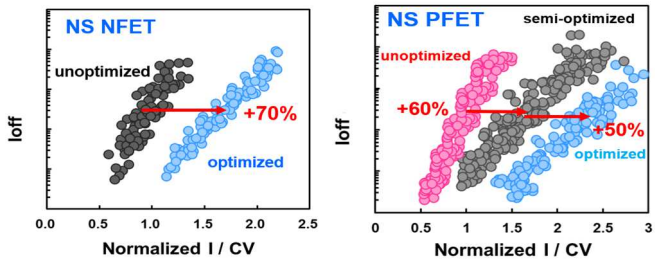


Fig.6 Not only drive current/mobility enhancement, more so on Ceff reduction: N/P +70% and +110% gain in I/CV speed

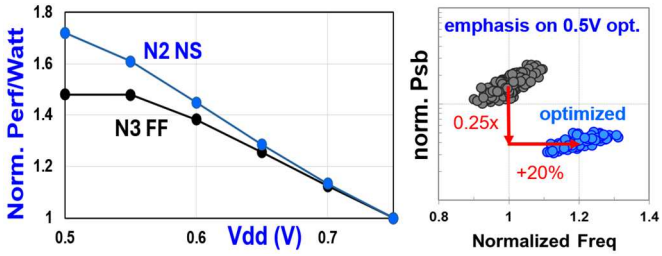


Fig.7 NS vs FF: better Perf./Watt at 0.5V~0.6V. Sp. emphasis at low-Vdd: +20% speed and 75% lower std-by power

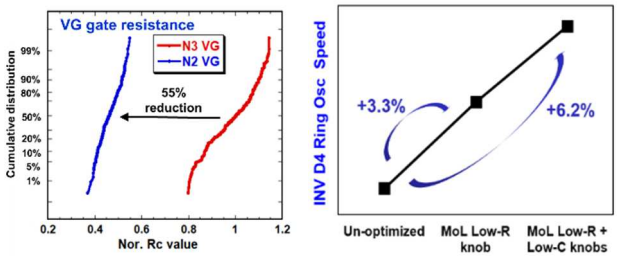


Fig.8 MoL low R (VG, VD and MD) and lower Ceff optimization leading to 6.2% speed gain

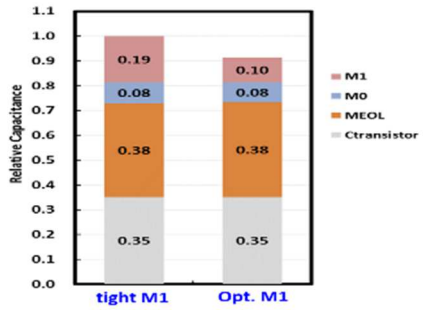


Fig.9 Optimized M1 with novel 1P1E EUV patterning leads to 9% Ceff reduction and a saving of multiple EUV masks

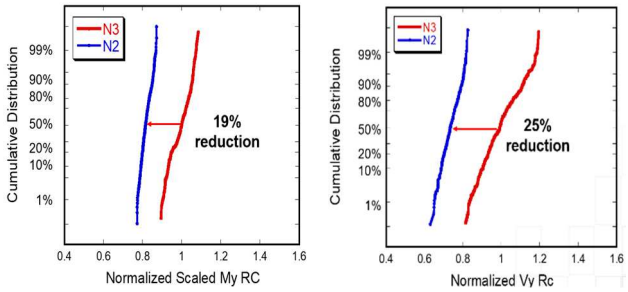


Fig.10 Significant My RC and Vy Rc reduction on N2 tightest 193i 1P1E workhouse metals/vias

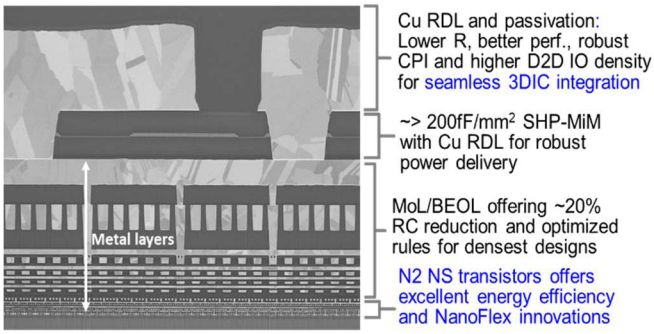


Fig.11 N2 new Cu RDL and passivation provide seamless integration with 3DFabric™ (SoIC, INFO and CoWoS) tech

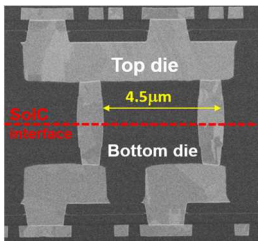


Fig.12 N2 optimizes pTSV and sTSV for F2F/F2B stacking w/ SoIC bond pitch →4.5 µm

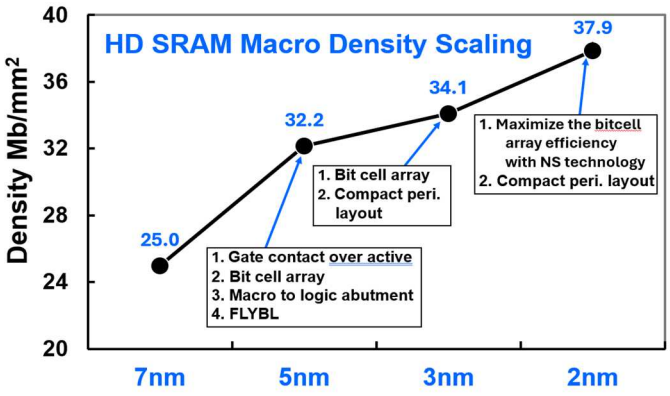


Fig.13 N2 offers highest SRAM macro density ~38Mb/mm²

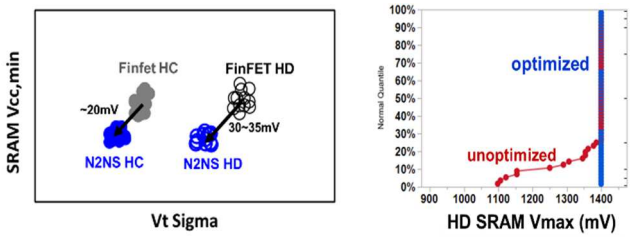


Fig.14 N2NS with better DIBL and Vt-sigma leading to lower Vmin for more energy-efficient compute. HD Vmax>1.4V

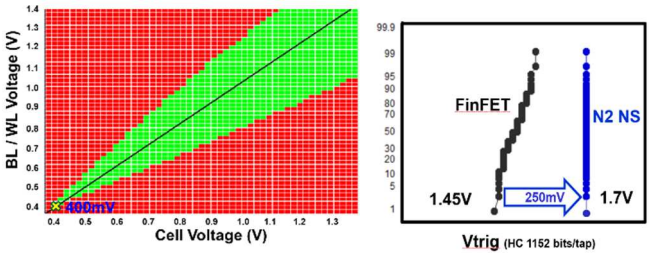


Fig.15 256Mb HD SRAM shmoo to 0.4V. N2 >1.7V Vtrig: higher logic density and effective DVS for product quality

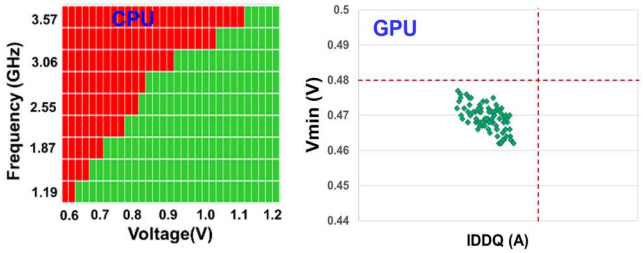


Fig.16 Shmoo plots of CPU block and GPU Vmin vs. IDDQ in the high yielding logic test chip in N2 qual vehicle

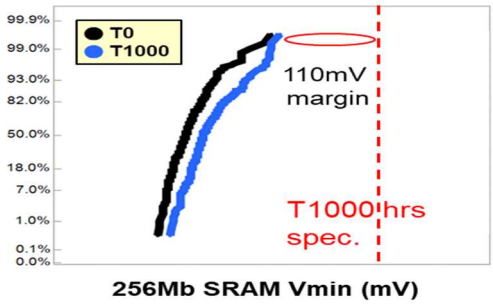


Fig.17 N2 technology met all wafer-level reliability requirements and passed 1000hrs HTOL spec

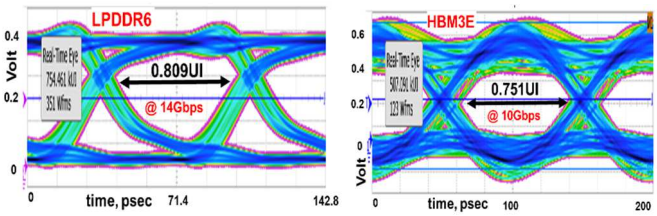


Fig.18 High-speed SerDes test chip in N2 vehicle showing fully functioning LPDDR6 @14Gb/s and HBM3E @10Gb/s

Low-temperature Behavior in Nanowire Transistors by Quantum Transport Simulation

S. Moslemi-Tabrizi[1], P. Schvan[1], U. Kapoor[2], P. Blaise[2], D. Lemus[3], and T. Kubis[3]
[1]Ciena Inc, Ottawa, Canada, email: smoslemi@ciena.com
[2]Silvaco Inc, Santa Clara, CA, USA, [3]University of Purdue, West Lafayette, IN, USA

***Abstract*—** The simulation of a silicon nanowire n-FET is performed using the NEGF quantum transport solver Victory Atomistic down to a temperature of 2 K. Inner spacers offering an optimal electrostatic confinement are used promoting several localized states appearing in the channel close to the bottom of the conduction band. The NEGF simulations allow probing of the NWFET behavior with various localized states at low temperatures, by gradually injecting electrons in the conductive channel and by taking into account the electron-phonon scattering mechanisms.

I. Introduction

The physical dimensions of FET transistors continue to shrink thanks to the VLSI integration and EUV lithography, this is toward a generalized usage of nanosheet and nanowire transistors made of silicon, (NSFET and NWFET), with a typical cross-section of a few nm^2 [1]. At the same time, there is a tremendous regain of interest in using silicon-based transistors at low temperatures for a few Kelvin and below, for high-performance computing, new electronic memories, and quantum computing [2]. This poses difficulties to conventional TCAD simulation of devices, being unable to capture a correct energetic landscape of electrons at a small scale, (i.e. beyond the effective mass approximation due to the strong confinement effect), and hardly converging at low temperatures. In that respect, the non-equilibrium Green's function (NEGF) method offers several advantages and can remedy TCAD difficulties, which is what we showcase in this work.

II. Atomistic Simulation Methodology

A. NEGF code

Throughout this study, we use the VictoryAtomistic tool [3] an evolution of Nemo5 [4] delivering a decisive speed-up in NEGF calculations thanks to a generalization of the modespace approximation technique [5]. Electronic bandstructure is described by a tight-binding $sp^3d^5s^*$ basis offering an accurate description of the electronic levels close to the Si bandgap plus a very good transferability to an NWFET geometry with a nanometric section. The silicon dangling bonds are fully passivated with hydrogen atoms thanks to an ad-hoc self-energy of passivation. An adaptive energy mesh is used with an average resolution of 1 meV for T = 300 K and 60 K, and with a resolution better than 0.2 meV for T = 20 K and T = 2 K. Each simulation is performed using a suitable low-rank approximation matrix [5] that preserves the calculation accuracy and decreases the simulation time by more than two orders of magnitude. This combination of techniques allows one to test many configurations and temperatures, alleviating difficulties previously encountered at low T below a few dozen K in NEGF [6].

B. Physical dimensions and parameters

A typical NWFET structure is visible in Figure 1, showing the essential geometric characteristics and its corresponding atomistic structure. The crystallographic direction along the transport direction is Si <100>. The channel is made of intrinsic Si with a gate-all-round (GAA) structure of 1nm EOT. There are two inner spacers on the source and drain sides. The dielectric thickness of the spacers is 1nm and one uses several dielectric constants. The source and drain are highly doped. All the essential characteristics are summarized in Table 1.

The electron-phonon interactions are included thanks to the self-consistent Born approximation. The physical parameters relative to the acoustic and optical branches are summarized in Table 2. We use silicon-bulk parameters as a first guess to probe the essential effect of phonons on the current characteristics and levels broadening.

C. NEGF self-consistency

Our NEGF simulator solves for a steady state of the electronic device. Its self-consistency is reached when the electronic density solved with the $G^<$ component of NEGF doesn't significantly change when iterated in Poisson's equation solver. Then one can have access to the non-equilibrium occupation numbers and electronic current through the channel, as illustrated in Figure 2. The electron-phonon coupling is naturally taken into account by the self-energy terms. It is worth noting that both the self-energy in the channel and the occupation numbers of each electrode depend directly on temperature.

III. Intermediate Temperatures Results

We study an NWFET of 2x2 nm^2 square cross-section, for which the bandgap becomes direct at the Γ point due to a strong confinement effect. We work at low Vds = 10 mV. The Id(Vg) transfer characteristics are shown in Figure 3 for two temperatures 300 K and 60 K. Thanks to the good electrostatic control offered by the GAA geometry, the subthreshold slopes are obtained at 64 mV/dec and 13 mV/dec for a current of 10^{-10} A at T= 300 K and 60 K respectively, (a bit higher than the predicted theoretical values essentially because of the e-ph coupling). Above the threshold voltage, the saturation current is lower at room temperature due to an increased e-ph coupling. More interestingly, one can see a kind of current fluctuation at

T = 60 K not visible at room temperature. This fluctuation is related to a few quantum localized states that appear below the top of the barrier, as illustrated in Figure 4 for the density of states (DOS) resolved in energy and space obtained at Vg = 0.6 V. Indeed these localized states contribute only very little to the current density as a function of the applied gate voltage. The smearing of levels at T = 300 K on several kB.T doesn't allow these levels to get any weight in the I(V) characteristics, and they start to be hardly perceived at T = 60 K. Moreover a level splitting of 5 meV is visible at T= 60 K for the lowest energy level. These encouraging preliminary results permit us to study the NWFET at even lower temperatures.

IV. Results at T = 20K and T = 2K

A. Electronic levels as a function of Vg at T = 20 K

The confined electronic levels start to detach from the conduction band for a gate potential higher than 0.5 V. Two levels below the top of the barrier are going deeper into the quantum well formed by the channel and the two spacers as illustrated Figure 5 showing the DOS resolved in energy and space. They are followed by other states higher in energy showing a different symmetry with supplemental lobes along the transport direction. Thanks to low Vds applied and the spatial and energetic proximity of the electrodes, in a steady state regime, the lowest energy levels start to be filled by a few electrons, as can be seen in Figure 6 showing the electronic density resolved in energy and space.

B. States occupancy from T = 20 K to T = 2 K

At T = 2 K, the two lowest levels in energy gain a supplementary splitting of 1 meV, see Figure 7. This does not affect the global charge of the levels estimated in Figure 8 for the two temperatures 20 K and 2K. This net charge is obtained by counting the number of electrons in the channel and removing the contribution from any other states than the levels detected in the same energy window and spatial location of interest. Interestingly, the coupling of the charge in the Vg = [0.48, 0.64] V window is almost independent of the temperature and is purely capacitive with a 2 aF proportionality.

C. Confinement function of geometry and dielectric spacers

Figure 9 shows the DOS resolved in energy and space for a slightly larger and rectangular cross-section of 2x3 nm^2. In comparison with the 2x2 nm^2 square cross-section, the splitting of the first energy levels increases from 5 meV to 7 meV. This underlies the importance of describing accurately the atomistic structure and the nontrivial symmetries of the atomic arrangements in the confinement directions.

Figure 10 illustrates the effect of employing longer dielectric spacers of 14 nm instead of 7 nm with two different values for the dielectric constant $\varepsilon = 12$ and $\varepsilon = 4$. The DOS resolved in energy doesn't show a much better confinement by employing longer spacers. Nevertheless, by decreasing the spacer dielectric constant, the energy confinement is increased by 5 meV, and the spatial localization is only slightly improved. This is explained by a deeper quantum well formed by the presence of the two spacers that are less coupled to the gate when the dielectric constant is lowered.

D. Energy relaxation of the electronic states due to phonons

By taking into account the self-energy of acoustic and optical phonons of silicon in interaction with electrons, the electronic levels of confined states are broadened with increasing temperature, see Figure 11. To resolve the four levels of lowest energy of the confined states, one needs to lower the temperature below 2 K, (note that the energy resolution of Figure 11 is 0.1 meV approx.). For the conduction band of silicon, the inclusion of the spin-orbit coupling interaction doesn't change significantly this result, (calculation not shown).

V. Conclusion

Thanks to an efficient combination of several numerical optimizations, a systematic exploration of the density and density of states resolved in energy and space becomes doable for NWFETs with the NEGF feature of Silvaco's Victory Atomistic. The results obtained at low temperatures show the importance of combining an atomistic resolution with a suitable basis to describe accurately the electronic levels. For the steady state of a Si NWFET oriented <100> and including inner spacers, several confined states appear near the threshold voltage. These states present a complex energetic profile close to the conduction band of silicon depending on several factors and physical parameters. They contribute very slightly to the overall current of the NWFET but can be filled by a few electrons at low Vds voltage. On one hand, the capacitive coupling with the gate can be quantified, on the other hand the short channel combined with the spacers contributes to the efficiency of the quantum well. For better control of low-T devices, due to electron-phonon coupling, a temperature below 2K is required to address each state individually by resolving the level broadening.

Acknowledgment

Part of the calculations were performed using Negishi's machine at Purdue's Rosen Center of Advanced Computation.

References

[1] A. Razavieh, P. Zeitzoff, and E. J. Nowak, "Challenges and Limitations of CMOS Scaling for FinFET and Beyond Architectures," in IEEE Transactions on Nanotechnology, vol. 18, pp. 999-1004, 2019.

[2] K.K. Likharev, "Single-electron devices and their applications," in Proceedings of the IEEE, vol. 87, no. 4, pp. 606-632, April 1999

[3] Victory Atomistic User's Manual by Silvaco International, (2023).

[4] S. Steiger, M. Povolotskyi, H.-H. Park, T. Kubis, and G. Klimeck, "Nemo5: a parallel multiscale nanoelectronics modeling tool", IEEE Trans Nano, vol. 10, p. 1464, 2011

[5] Lemus, D.A., Charles, J. & Kubis, T. Mode-space-compatible inelastic scattering in atomistic nonequilibrium Green's function implementations. J Comput Electron 19, 1389–1398 (2020)

[6] Lavieville R, Triozon F, Barraud S, Corna A, Jehl X, Sanquer M, Li J, Abisset A, Duchemin I and Niquet Y-M 2015 Quantum dot made in metal oxide silicon-nanowire field effect transistor working at room temperature Nano Lett. 15 2958–64

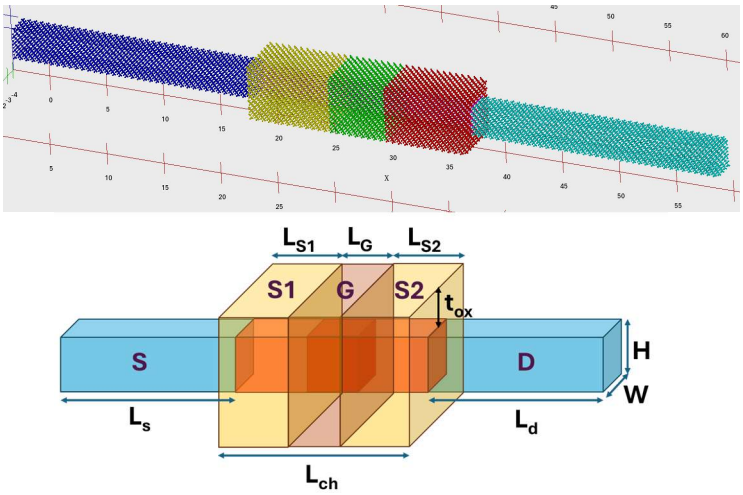


Fig. 1. Si nanowire GAA-FET structure of square cross-section including dielectric spacers. Top: atomistic structure of a 2x2 nm^2 nanowire with the source (blue) and drain (turquoise) of 20 nm length each, gate of 5 nm (green), and dielectric spacers (yellow and red) of 7 nm each. Bottom: schematic of the NWFET geometry with the various characteristic lengths.

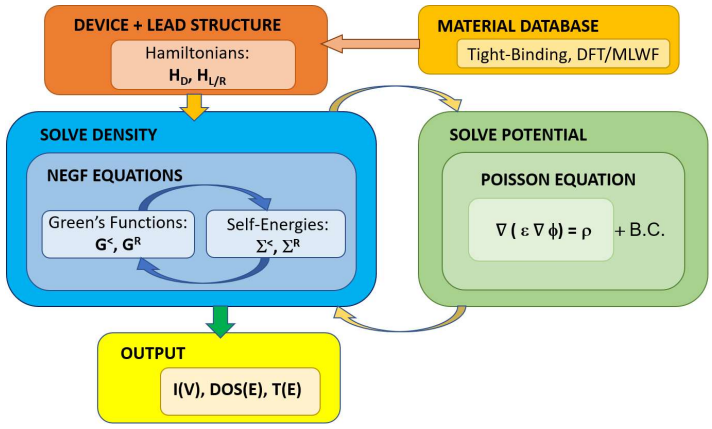


Fig. 2. Victory Atomistic flowchart showing the NEGF solver coupled with the Poisson equation solver. The Hamiltonian of the electrodes and the central device is built within a tight-binding formalism and with open boundaries conditions. The transfer characteristic is calculated when the self-consistency of NEGF and Poisson is reached.

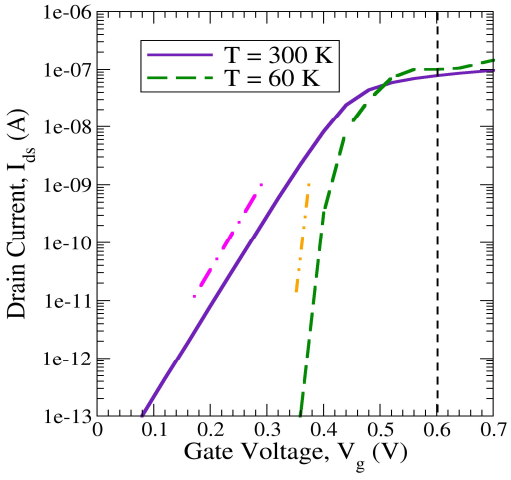


Fig. 3. Id(Vg) current transfer characteristic at T = 300 K and 60 K, Vds = 10 mV, along with the ideal slopes of 60 mV/dec (magenta) and 12 mV/dec (orange).

Param.	Value/range
Cross-section	2x2, 2x3 nm^2
L_S, L_D	20 to 40 nm
L_{CH}	19 to 33 nm
L_{S1}, L_{S2}	7 to 14 nm
L_G	5 nm
t_{OX}	1 nm
$\varepsilon_{S1,S2}$	4 to 20 (ε_0)
ε_G (SiO_2)	3.9 (ε_0)
Doping$_{(CH)}$	10^{15} (e.cm^{-3})
Doping(S,D)	10^{20} (e.cm^{-3})

Table 1. Physical dimensions of the Si <100> NWFET, dielectric properties and doping.

Param.	Value/range
Temperature	2, 20, 60, 300 K
V_{sound} (Si)	8433 m.s^{-1}
ρ(Si)	2.336 g.cm^{-3}
D_{ADP}	8.8 eV
ω_{op}	63 meV
D_{ODP}	110 eV.nm^{-1}
W_F	4.2 eV
V_{DS}	10 mV
V_G	0.0 to 0.7 V

Table 2. Physical parameters employed for the NWFET simulations.

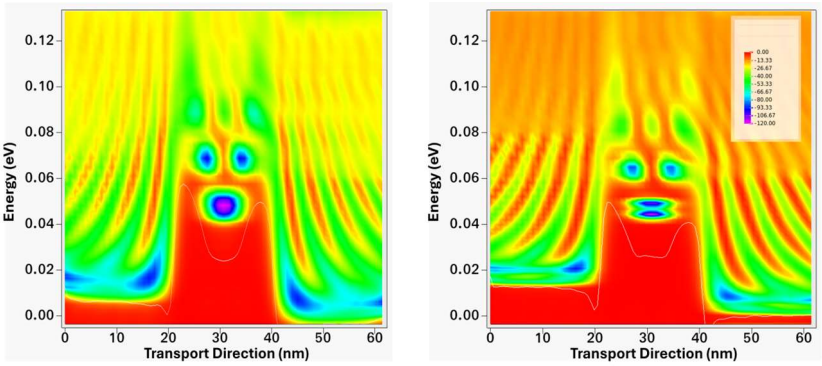


Fig. 4. Density of states resolved in energy and position along the transport axis obtained at T = 60K (left), T = 300K (right) with Vg = 0.6 V, Vds = 10 mV. The color scale is in arbitrary units from low density (red) to high density (violet). A few confined states are visible in the channel region below the gate around 50 meV. The white line is here to visualize the 1D electrostatic profile aligned with the bottom of the conduction band on both leads. At T = 60K a first splitting of 5 meV for the lowest energy level appears.

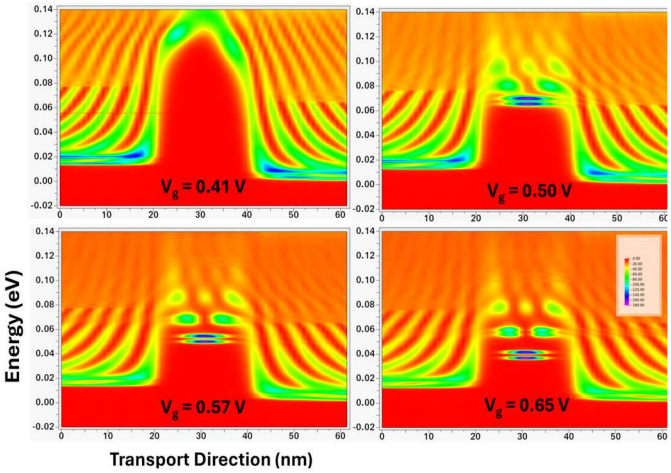


Fig. 5. Density of states resolved in energy along the transport axis at T = 20K, Vds = 10 mV, for four different values of Vg = 0.4, 0.5, 0.57, 0.65 V. Two localized states appear at Vg = 0.5 V slightly below the top of the barrier. These two states are pushed lower in energy as Vg increases, followed by other states showing a different symmetry.

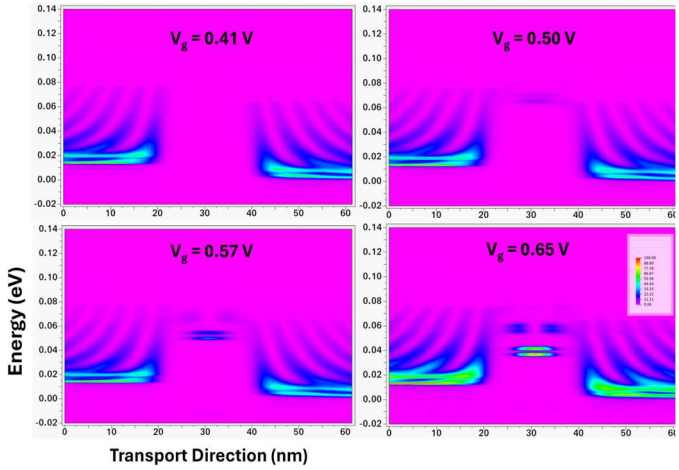


Fig. 6. Electronic density resolved in energy along the transport axis at T = 20K, Vds = 10 mV. The electronic levels are progressively filled using four different values of Vg = 0.4, 0.5, 0.57, 0.65 V, with a total number of electrons approximately equal to one at Vg = 0.57 V, and equal to two at Vg = 0.65 V.

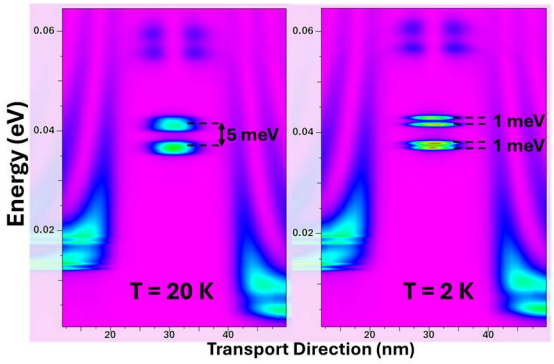


Fig. 7. Electronic density resolved in energy along the transport axis at T = 20K vs 2K, Vds = 10 mV, Vg = 0.65 V. At T = 20 K a level splitting of 5 meV is visualized. At T = 2 K, a second splitting of 1 meV occurs for the lowest energy levels, the progressive filling of these levels being unchanged.

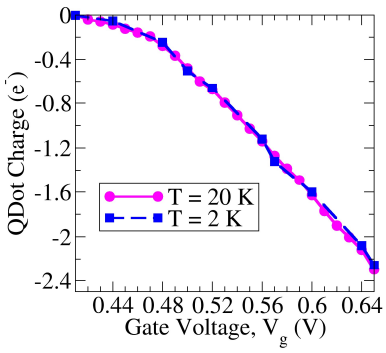


Fig. 8. Total electronic charge captured by the localized states of lowest energy function of Vg at T = 2K and 20K.The total charge follows a quasi-linear law for Vg=[0.48, 0.64] V.

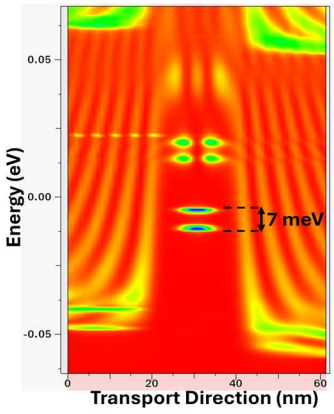


Fig. 9. Density of states resolved in energy along the transport axis at T = 20 K, Vds = 10 mV, Vg = 0.65 V for a rectangular NW of 2x3 nm^2 showing a level splitting of 7 meV.

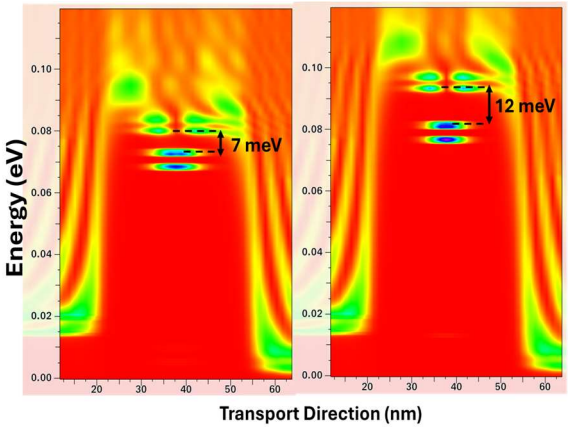


Fig. 10. DOS(E, x) along the transport axis at T = 20K, Vds = 10 mV, Vg = 0.57 V using long spacers of length 14 nm, and with two different dielectric constants ε = 12 (left), ε = 4 (right). A lower dielectric constant helps in confining the electronic levels.

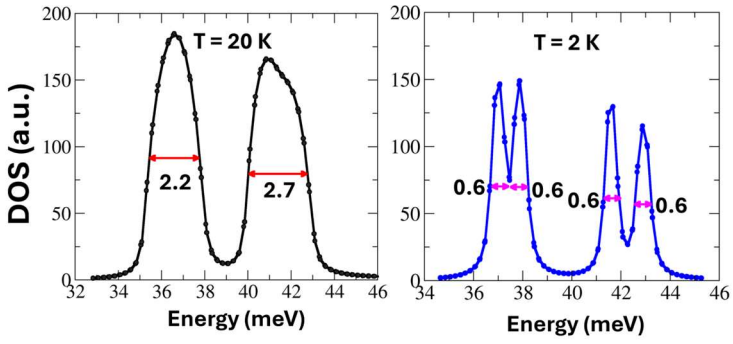


Fig. 11. Energy relaxation of the electronic states of lowest energy in the transistor channel due to the electron-phonon coupling for two temperatures at T = 20 K (left panel) and T = 2K (right panel), Vds = 10 mV, Vg = 0 .65V. Each level broadening is estimated with the full width at half maximum. Acoustic and optical phonon branches are taken into account in NEGF simulation within the self-consistent Born approximation.

Ultrafast Charge Trap-based Volatile Memory Cell with Schottky Barrier S/D and Thin Tunnel Oxide

Hyeongyu Kim[1*], Dabok Lee[2*], Hyun-Sik Choi[3], Yoojin Seol[1], Jonghyeon Ha[2], Sung-Il Chang[4], Kihyun Kim[1], Jungsik Kim[2], Won-Ju Cho[3], Zvi Or-Bach[4]

[1]Division of Electronic Engineering, Jeonbuk National University, Jeonju, Jeonbuk, 54896, Rep. of Korea

[2]Department of Electrical Engineering, Gyeongsang National University, Jinju, Gyeongnam, 52828, Republic of Korea

[3]Electronic Materials Engineering, Kwangwoon University, Seoul, 01897, Rep. of Korea

[4]MonolithIC3D, Klamath Falls, OR 97601, United States

*: Equal contribution; Corresponding E-mail: sung.il@monolithic3d.com

***Abstract*—**In this paper, we propose a novel 3D Charge Trap-based DRAM (CT DRAM) having great advantages on power consumption and heat dissipation. Schottky barrier-induced hot carrier injection and ultrathin tunnel oxide were used for fast write operation. Microwave anneal (MWA) shows better silicide formation (low resistance and high diode on/off ratio) than conventional rapid-thermal annealing (RTA) and its volume heating is quite effective for 3D integrated devices. We designed key parameters (Schottky barrier height, ONO thickness, and trap characteristics) with 3D TCAD simulation and fabricated the planar device having poly-Si channel to explore 3D integration. The device is expected to have a program/erase window larger than 2 V below 10 ns write pulse, and retention time longer than 10s@85°C, and >0.22V window after 10^{15} cycles.

I. Introduction

The current technology of 10-nm saddle-fin-based DRAM has reached the limit of miniaturization due to various reasons, such as the limitation of meeting the minimum line resistance and capacitance values and the reliability problem of bit flip due to row hammer [1-2]. Although various attempts are also being made to develop next-generation 3D DRAMs, 1T1C 3D DRAM is extremely hard to integrate, and 2T IGZO 3D DRAM is not cost-effective for over several hundreds of layers [1, 3]. Therefore, developing a DRAM structure with a straightforward 3D integration scheme like 3D NAND is essential to achieve higher bit density [4].

In this paper, we present a novel 3D charge trap-based memory cell structure as a next-generation DRAM, as shown in Fig. 1. It uses an ONO structure instead of a capacitor and a metal silicide source/drain (S/D) for hot carrier injection (HCI) [5]. Based on parameters optimized by TCAD simulation, 2D device was fabricated and their operating speed and reliability were measured to verify the feasibility of 3D memory cells.

II. Device Design of 3D DRAM

A. Structure design

Fig. 2. summarizes the main process flow of the 3D CT DRAM devices. The key processes are a thin tunnel oxide, donut-type poly-Si channel and metal silicidation. We deposit O/N/O and poly-Si sequentially on the nitride-indented mold with holes. Following poly-Si etch results in the formation of the separated donut-type poly-Si channels in the recessed region. The process sequence used to form the poly-Si islands is already verified through the mass production of the floating-gate type 3D NAND devices from Intel and it is a highly reliable process [6-7]. Metal silicide is formed at the both ends of the poly-Si channel in the SL and BL sides. The proposed structure has strong advantages in power consumption and heat dissipation, which are the most important aspects for high-performance DRAM like HBM. Fig. 3. (a) shows its power advantage. It consumes less power than the conventional 1T1C DRAM cell because it does not lose charges during the read operation and does not need to be restored. The CT DRAM without capacitors can achieve extremely high integration density and metal S/Ds is very effective to dissipate heat from the underlying logic die as shown in Fig. 3 (b).

Fig. 4. (a) shows the cross-sectional device structure of the 3D CT DRAM and Fig 4. (b) shows its program operation. It is very hard for the poly-Si channel transistors to make hot electrons at the drain side because of many grain boundaries and trap sites. But if we utilize the SB at the source side, we can make hot carriers easily even with the poly-Si channel. The electrons injected through the SB gain additional energy and result in impact ionization by the steep E-field on the source side to form electron-hole pairs (EHPs).

B. Optimization of device parameters

We optimized the process/device parameters of the 3D CT DRAM with 3D TCAD simulation. Fig. 5. (a) and (b) show the simulated I_d-V_g curves and the energy band diagrams along with various SB heights, respectively. The larger SB height is, the steeper band bending occurs on the source side but on-state current is decreased because of decreased carrier tunneling. Considering the program characteristics and the on-current, SB of around 0.25 eV is good for the device. For T_{oxb} between 0.9 nm and 1.5 nm, the variation in threshold voltage (V_{th}) and electrostatic potential is small as shown in Fig. 5. (c), which means that the difference in vertical and lateral E-fields required for HCI is also negligible as shown in Fig. 5. (d).

Fig. 6. (a) shows the retention time as a function of the trap density of charge trap layer (CTL). A high trap density increases the number of electrons trapped in the CTL, resulting in a higher initial V_{th}. Fig. 6. (b) shows the retention time as a function of trap energy level. High trap energy level improves

retention time because a deep trap is formed, requiring high energy for electrons to escape from the trap. In Fig. 6. (c), we can see that the thicker T_{oxb} leads to an increase in the energy to escape to the channel beyond the tunnel oxide, which improves the retention time. T_{oxb} around 1 nm is good enough considering the proper retention characteristics for DRAM. Fig. 6. (d) summarizes the mechanism of electron leakage.

The process target values were set based on the simulation result, and the detailed parameters are shown in Table 1.

III. Device Fabrication

The feasibility of the 3D CT DRAM was verified by fabricating and measuring CT DRAM devices with planar structures as shown in Fig. 7.

A. Schottky barrier silicide formation

In this study we formed silicides by using Microwave annealing (MWA). MWA utilizes electromagnetic radiation to heat the materials, offering the advantage of achieving uniform volume heating [8]. The sheet resistance (R_s) and junction characteristics of the silicide were examined as a function of microwave power. Fig. 8 depicts the sheet resistance of Ni and Co-silicide according to MWA conditions. Overall, Ni-silicide exhibits lower R_s compared to Co-silicide. Fig. 9. (a) and (b) show the on/off current ratio and SB heights (ϕ_b) of the SB diodes, respectively. On/off currents were extracted from 5 V and averaging reverse current, respectively. ϕ_b was extracted using the following equation (1).

$$\boldsymbol{\phi_b} = \frac{kT}{q} \ln\left(\frac{A^{**}T^2}{J_0}\right) \quad (1)$$

Where q, k, T, A^{**}, and J_0 are the unit electron charge, Boltzmann's constant, absolute temperature, equivalent Richardson's constant, and reverse saturation current density, respectively. In this study, we applied Ni, which exhibits a relatively higher on/off current ratio, SB heights, and lower R_s. SB height can be further adjusted by optimizing the stoichiometry of Ni silicide [9-10].

B. Device integration

Fig. 10 illustrates the fabrication of the 2D CT DRAM device. Buried oxide and active poly-Si layers were formed on the substrate. After the active region was formed, N/O layers and n+ poly-Si gate were formed. Then, ON spacer was formed to isolate gate and S/D regions during following self-aligned silicidation process. Finally, nickel was deposited using an Electron Beam Evaporator, and MWA was performed to form a silicide with low contact resistance. The unreacted Ni was removed.

IV. Results and Discussions

Fig. 11. (a) shows the cross-sectional TEM image of the fabricated device. NiSi was formed successfully at the gate, source, and drain. The S/D regions were fully silicided and its thickness is 20 nm. The gate spacer of 15 nm can be further reduced to decrease the non-overlap length between the gate and the NiSi S/D. Fig. 11. (b) illustrates the thickness of the ONO layer, with the tunnel oxide, trap layer, and blocking oxide layers of 1, 2.3, and 4 nm, respectively.

Fig. 12 illustrates the I_d–V_g curves of the 2D CT DRAM devices utilizing MWA at different time and power levels. As mentioned earlier, we used 600 W/1 min of MWA condition considering the lateral growth of NiSi as well as on-current of the devices. Fig. 13. (a) compares the program characteristics of the device with a tunnel oxide thickness of 1 nm to that with a thickness of 2.9 nm. The device with T_{oxb} of 1 nm showed as fast program speed as V_{th} shift of 2 V even at the program time of 20 ns@V_{GS} = 10 V. Fig. 13. (b) shows the erase characteristics of the device as a function of erase time and voltage. The device with a T_{oxb} of 1 nm demonstrated a fast erase speed with a V_{th} change of approximately -0.8 V at an erase time of 20 ns@V_{GS} = -6.5 V. Fig. 14. shows the program characteristics utilizing hot carrier injection at a short program time of 20 ns and lower gate biases. When programming using FN tunneling without applying drain voltage, the threshold voltage shift was very small, approximately 0.5 V at a program voltage of 9 V. As the drain voltage increases, a larger E-field is generated at the source end, resulting in significantly faster program characteristics.

Fig. 15 shows the retention characteristics at 85°C of the 2D CT DRAM device with a thin tunnel oxide. Even with the tunnel oxide as thin as 1 nm, more than 50% of the initial Vth window was retained after 0.2 seconds, which is sufficient for DRAM applications.

Fig. 16 shows the endurance (>$5x10^5$ Cycles) of the fabricated device as a function of pulse cycles. The fabricated 2D CT DRAM cell shows a V_{th} window of 0.22 V, ensuring sufficient reliability after 10^{15} P/E cycles.

V. Conclusions

This study investigated the CT device, its potential as a next-generation 3D memory solution. We showed that CT DRAM was promising when it was combined with ultrathin tunnel oxide and HCI from metal silicide S/D. Moreover, the 1T CT DRAM is one of the most promising candidates for 3D DRAM due to its simple process integration and easy heat dissipation through the metal S/D.

Acknowledgment

The authors would like thank H.-C. Hwang/B.-R. Yoon at NNFC for helping the device fabrication.

References

[1] J. W. Han, et. al., IEEE VLSI Tech., 2023
[2] M. Suh, et. al., IEEE TED, 2024
[3] A. Belmonte, et. al., IEEE IEDM, 2021
[4] J. Kim, et. al., IEEE IEDM, 2023
[5] S. -J. Choi, et al., IEEE VLSI Tech., 2009
[6] A. Khakifirooz, et. al., IEEE ISSCC, 2023
[7] K. Parat, et. al., IEEE IEDM, 2015
[8] Min, J.G., et. al., Nanomaterials, 2022
[9] K. Takahashi, et al., IEEE IEDM, 2004
[10] A. Lauwers, et. al., IEEE IEDM, 2005

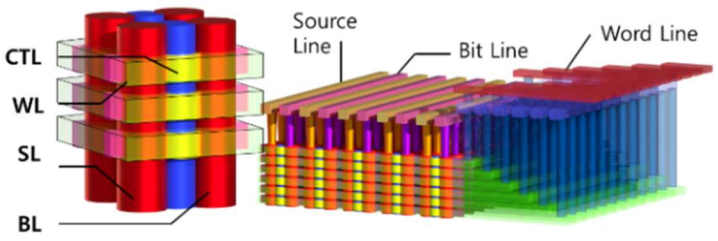


Fig. 1. Schematic illustration of the 3D CT DRAM

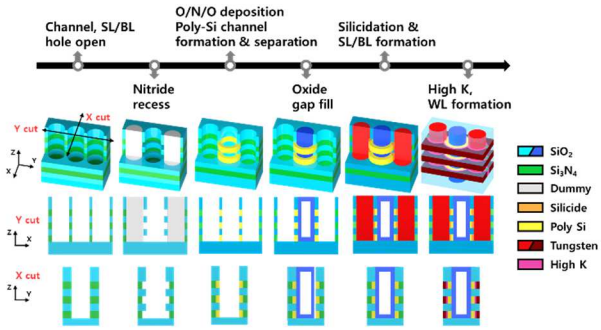


Fig. 2. Process integration of the 3D CT DRAM with SB S/D.

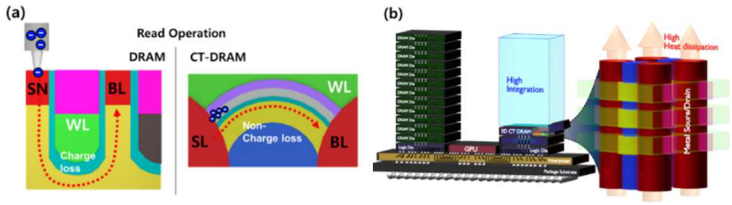


Fig. 3. Advantages of the CT DRAM compared to the conventional DRAM in view of (a) power and (b) 3D integration/heat dissipation.

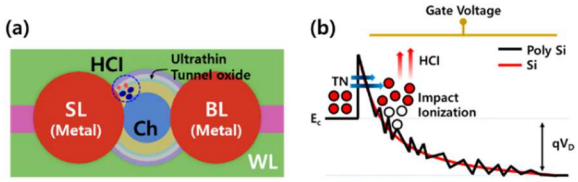


Fig. 4. (a) Cross-sectional view of the 3D CT DRAM with Schottky barrier (SB) S/D and ultrathin tunnel oxide (b) HCI at the source-side was used for the program operation at the device having a poly-Si channel.

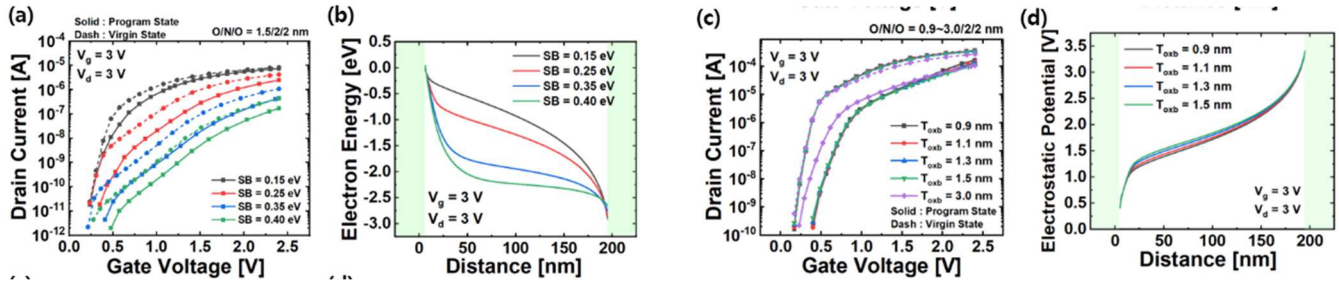


Fig. 5. I_d–V_g characteristics at different (a) SB (0.15 eV ~ 0.40 eV). The SB characteristics that vary with (b) programmed energy band. I_d–V_g characteristics at the different tunnel oxide thickness (0.9 nm ~ 1.5 nm) shown in (c). (d) The electrostatic potential hardly changes with tunnel oxide thickness.

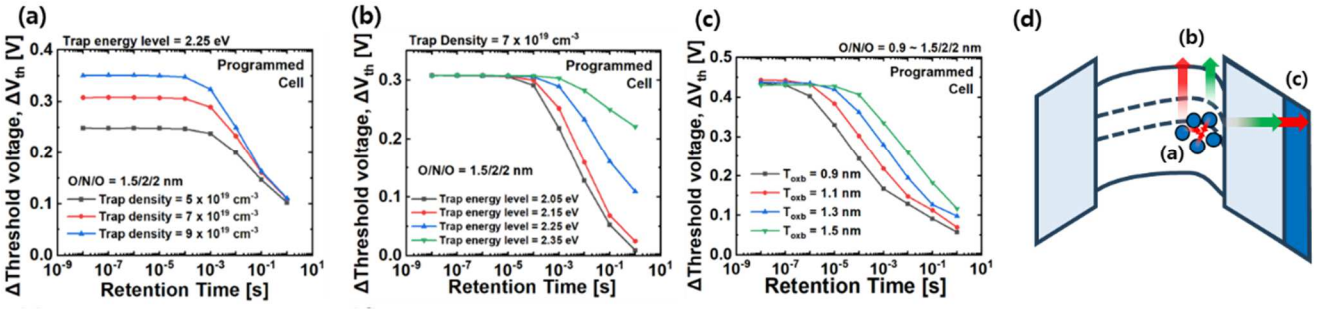


Fig. 6. The retention characteristics at (a) different trap density, (b) trap energy level, and (c) tunnel oxide thickness. (d) Mechanism of electron leakage.

Table 1. Optimized Device parameters for the 3D CT DRAM

Description (parameter)	Values
Blocking oxide thickness (T_{oxt})	2 nm
CTL(Si_3N_4) thickness (T_{CTL})	2 nm
Tunnel oxide thickness (T_{oxb})	1 nm
Channel length (L_g)	150~300 nm
Channel thickness (T_{ch})	7 nm
Schottky barrier height (SB)	0.25 eV
Nitride Trap density (N_{trap})	7×10^{19} cm^{-3}
Substrate doping concentration	Undoped
Gate work function	4.65 eV
Trap energy level	2.25 eV

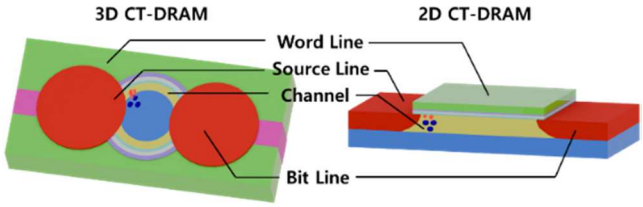


Fig. 7. Schematic of a planar CT DRAM to verify the feasibility of the 3D CT DRAM.

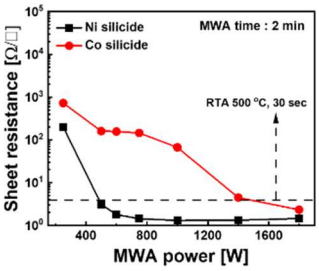


Fig. 8. Sheet resistance of Ni and Co silicide according to MWA power.

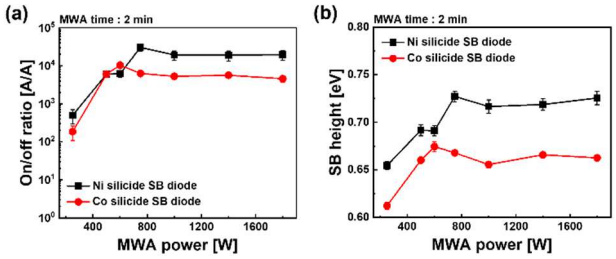


Fig. 9. (a) On/off current ratio, and (b) SB heights of Ni and Co-silicide SB diodes.

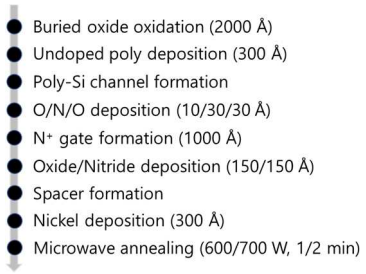


Fig. 10. Process integration of the 2D CT DRAM with SB S/D

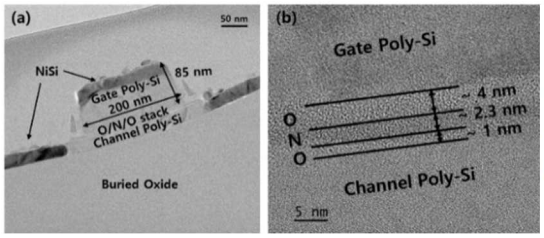


Fig. 11. TEM photographs of the fabricated 2D CT DRAM with SB S/D (a) Cross-sectional TEM (b) Magnified TEM of ONO layers.

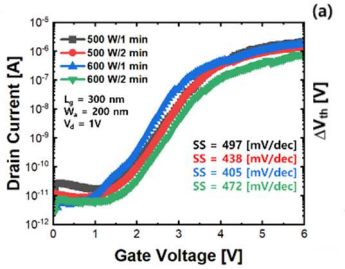


Fig. 12. Measured initial I-V characteristics of 2D CT DRAM device with L_g/W=300 nm/300 nm.

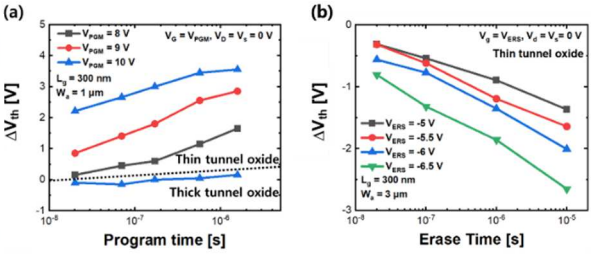


Fig. 13. Measured (a) program and (b) erase characteristics of the devices with thick and thin tunnel oxide.

\

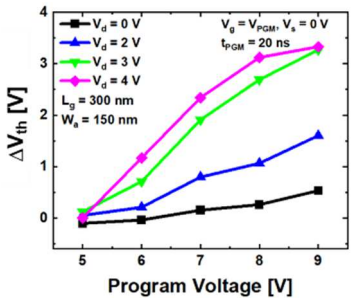


Fig. 14. Measured program operation characteristics using hot carrier injection with varying drain voltages.

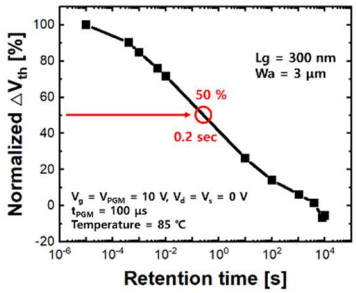


Fig. 15. Measured retention characteristics of the device with thin tunnel oxide at 85 ℃.

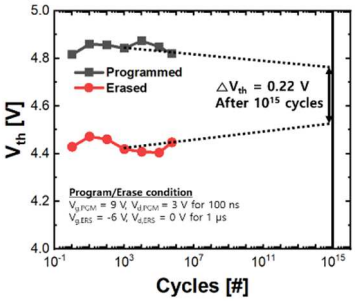


Fig. 16. Measured endurance characteristics of the device with thin tunnel oxide