

Enhanced Memory Performance in Ferroelectric NAND applications: The Role of Tunnel Dielectric position for Robust 10-year Retention

Prasanna Venkatesan^{1,^}, Andrea Padovani², Lance Fernandes¹, Priyanka Ravikumar¹, Chinsung Park¹, Huy Tran¹, Zekai Wang², Hari Jayasankar¹, Amrit Garlapati¹, Taeyoung Song¹, Hang Chen³, Winston Chern¹, Zheng Wang⁴, Kijoon Kim⁵, Jongho Woog⁵, Suhwan Lim⁵, Kwangsoo Kim⁵, Wanki Kim⁵, Daewon Ha⁵, Shimeng Yu¹, Suman Datta^{1,6}, Luca Larcher⁷, Gaurav Thareja^{7,#}, Asif Khan^{1,6,\$}

¹School of Electrical and Computer Engineering, Georgia Institute of Technology, GA, USA;

²Engineering Department “Enzo Ferrari”, University of Modena and Reggio Emilia, Italy;

³Institute of Electronics and Nanotechnology, Georgia Institute of Technology, GA, USA;

⁴Micron Technology, USA;

⁵Advanced Device Research Laboratory, Semiconductor Research Center, Samsung Electronics, South Korea;

⁶School of Material Science and Engineering, Georgia Institute of Technology, GA, USA;

⁷Applied Materials, USA.

{^pravindran6,\$akhan40}gatech.edu, #gaurav.thareja@amat.com

Abstract—We report a comprehensive study on memory window (MW) enhancement and retention loss (RL) in ferroelectric (FE) field effect transistors for NAND applications with a dielectric layer inserted within the FE film (next to the gate – GBL devices – or in the middle of the FE layer – TDL devices). Using a temperature-dependent modeling framework, we show that MW and RL properties are controlled respectively by carriers trapping into defects at the dielectric-FE interface and by its fragile balance with FE depolarization. While MW enhancement is achieved whenever the inserted dielectric layer is present, a robust retention is obtained only in TDL FEFETs, thanks to a stabilization of the trapped charge preventing FE depolarization. Simulation results clearly demonstrate that the fragile balance between above phenomena is ultimately controlled by the position of the dielectric layer, which determines i) the field redistribution within the stack, ii) the magnitude of charge de-trapping, and iii) the magnitude of the depolarization field. Our results provide critical guidelines for the improvement of FE 3D-NAND reliability.

Index Terms—FEFET, Retention, 3D-NAND, GinestraTM

I. INTRODUCTION

The exponential growth in data set sizes in omics and AI applications is driving the need for ultra-high density storage solutions. Charge trap flash (CTF) has been the workhorse long-term storage solution over the past decade. CTF NAND became the first memory technology to go 3D. This, in combination with physical scaling (z pitch, x-y spacing) and logic scaling (SLC, MLC, TLC, QLC), has enabled a 10x improvement in storage density (GB/mm²) over the last decade. State-of-the-art CTF NAND solutions feature 200-300 layers boasting over 20 GB/mm² data density [8]. However, further scaling the number of layers without exceeding the package limits has become challenging. This is

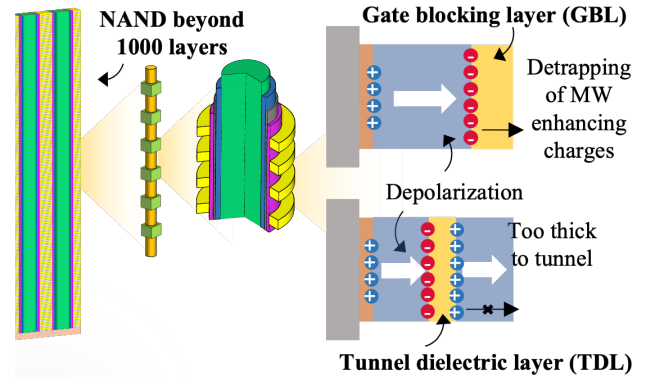


Fig. 1: Ferroelectrics for NAND: Inserting dielectric layers in a FE gate stack increases the MW irrespective of the location of the dielectric (GBL [1]–[4], TDL [5]–[7]) while only TDL FEFETs exhibit robust retention.

mainly driven by the reliability challenges like retention loss and increasing operating voltages associated with lowering the z-pitch [9].

To this end, FEFETs have emerged as a potential solution for z-scaling in vertical NAND devices providing a path for > 1K layers [9]. FEFETs store the memory state in the form of bound ferroelectric polarization rather than trapped charges. This could enable resilience against the reliability challenges faced in CTF NAND. Further, FEFETs also enable lower power consumption and high speed operation. However, FEFETs had to be engineered to meet the commercial NAND process requirements to serve as a drop-in solution: (1) Total gate stack thickness < 20 nm, (2) QLC compatible

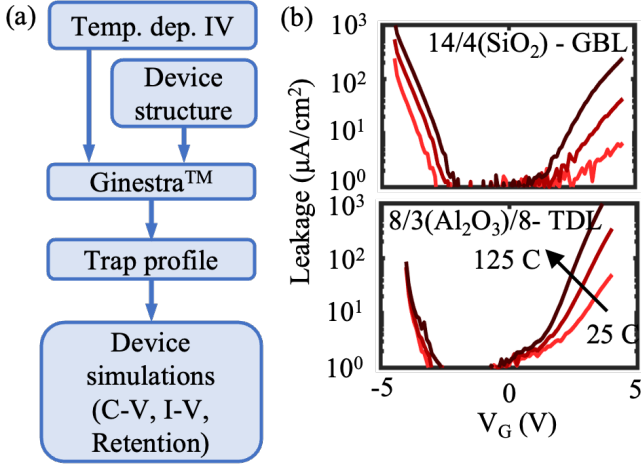


Fig. 3: Modeling framework: (a) Work flow for trap profile extraction. (b) The trap distribution in GBL - 14/4(Si) and TDL - 8/3(Al)/8 FEFETs are extracted from the temperature dependent I-V.

memory window (MW) > 7.5 V and (3) Low operating voltage < 15 V. Over the past two years, several reports have demonstrated low voltage, QLC and TLC compatible operation in FEFETs by inserting a dielectric layer in the ferroelectric gate stack [1], [5], [7], [10], [11] while maintaining the total thickness of the gate stack below 20 nm.

While FEFETs, in general, have been known to have robust retention, retention loss (RL) has been a key concern in these devices. Gate blocking layer (GBL) based FEFETs, where the dielectric layer is inserted next to the gate, shows 25-50% MW loss after 10^4 s at RT [1]–[4]. On the contrary, tunnel dielectric layer (TDL) based FEFETs, where the dielectric is inserted in the middle of the ferroelectric gate stack, demonstrate robust retention while exhibiting similar MWs [5]–[7], [10]. The focus over the last couple of years

has been MW enhancement. Retention, though crucial for non-volatile storage solutions and strongly correlated with the mechanism of MW enhancement, has received little attention.

In this work, we model MW enhancement in GBL and TDL based FEFET (Fig. 1) and perform detailed device simulations to analyze the RL mechanisms. We follow it up with experimental characterization of temperature dependent retention characteristics and capture their behavior using an Arrhenius model for retention loss.

II. SIMULATION FRAMEWORK

The retention study was carried out on four FEFETs with different gate stacks including ref. FE (19 nm HZO), TDL (9 nm HZO/ 1 nm Al_2O_3 / 9 nm HZO, 8 nm HZO/ 3 nm Al_2O_3 / 8 nm HZO), GBL (14 nm HZO/ 4 nm SiO_2) at 25-125°C fabricated using the process flow in [12]–[14]. The choice of dielectric layer has been optimized for maximum MW based on the position of the dielectric [15], [16].

The modeling framework was setup for TDL and GBL based FEFETs in Ginestra™ [17], a defect-based device modeling software. For TDL, 8 nm HZO/ 3 nm Al_2O_3 / 8 nm HZO gate stack is simulated while for GBL, 14 nm HZO/ 4 nm SiO_2 is used. The trap profile was extracted from measured temperature dependent leakage IV data from the corresponding FEFETs (Fig. 2(b)). Simulations consistently account for ferroelectric switching and charge trapping in the ferroelectric-dielectric stacks. The ferroelectric component is modeled using the Landau-Ginzburg model. The enhanced MW in the band-engineered FEFETs is realized by introducing fixed charges Q'_{it} at the FE-GBL/TDL interfaces. In order to calibrate the gate stacks to experiments, $C_G - V_G$ characteristics were simulated using the modeled gate stacks (Fig. 3). The MW is extracted from $C_G - V_G$ based on the shift in the flatband voltage. The MW extracted from $C_G - V_G$

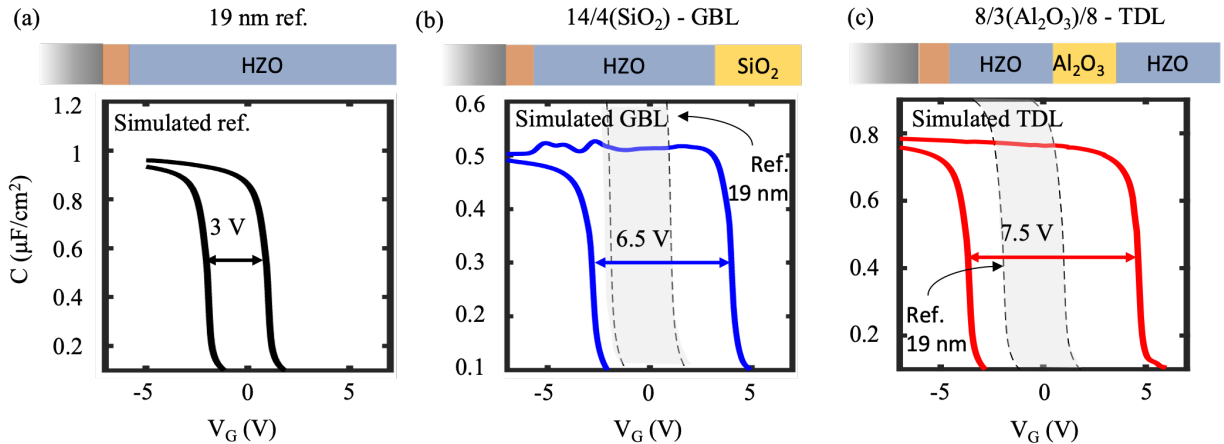


Fig. 2: Simulated $C_G - V_G$ in FE-MOSCAPs with ref. 19 nm, GBL 14/4(Si) and TDL 8/3(Al)/8 gate stacks showing MW enhancement due to interface charge Q'_{it} . The MW from the reference 19 nm FE-MOSCAP is shown as a shaded curve in (b) and (c). The MW is extracted from the shift in flatband voltage V_{FB} between the ERS and PGM states.

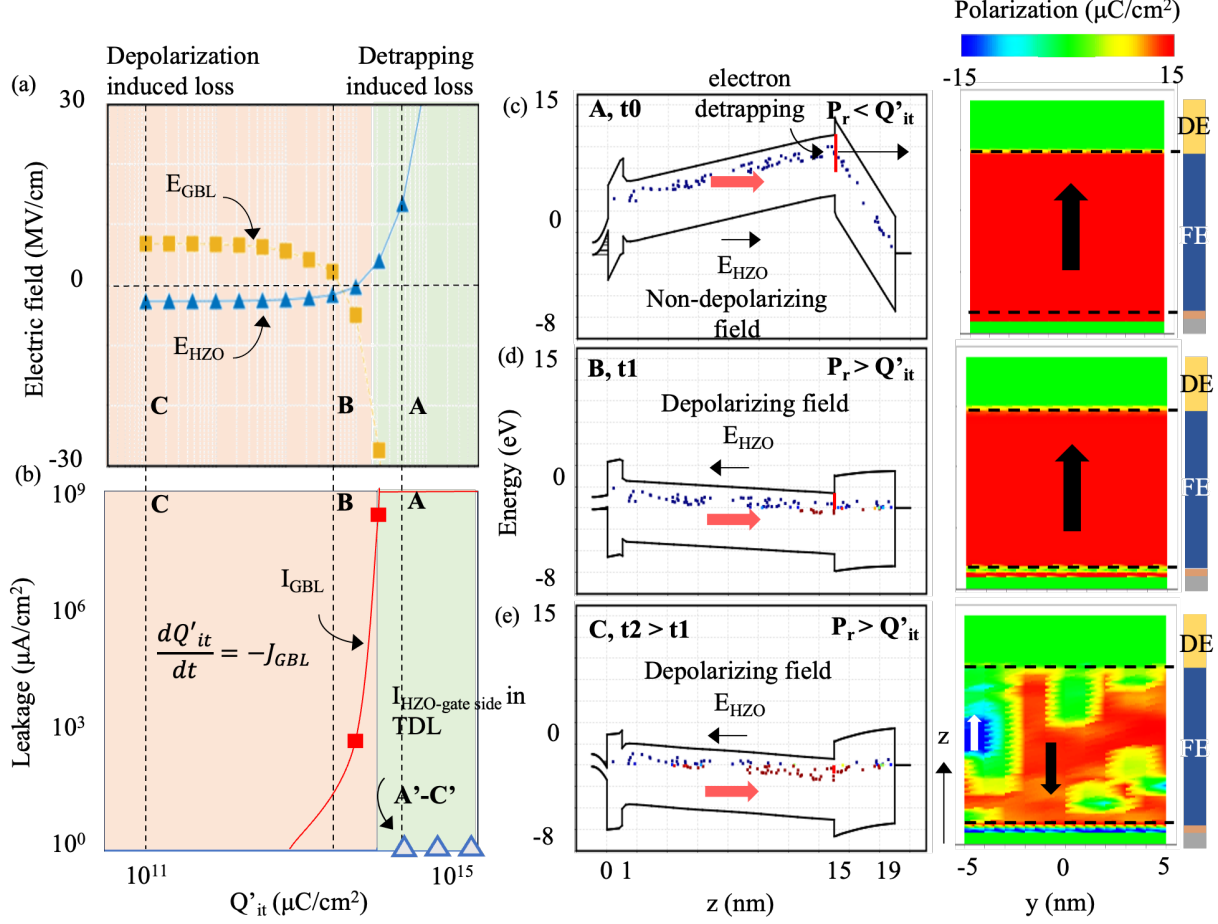


Fig. 4: Retention loss mechanisms: (a) Retention loss in all FEFETs is driven by the depolarization field in the ferroelectric layer. The electric field in the ferroelectric layer is depolarizing when $Q'_{it} < P_r$ and non-depolarizing when $Q'_{it} > P_r$. (b) Leakage in the GBL leads to rapid detrapping of the MW enhancing charges while gate side FE in TDL prevents charge loss. (c) Band diagram(left) and polarization map(right) for different cases: [A] The high field across the GBL drives out the MW enhancing charges ($Q'_{it} > P_r$) leading to loss of the MW enhancement. [B] When Q'_{it} reduces below P_r , the field in the ferroelectric layer becomes depolarizing. [C] The depolarization of the ferroelectric leads to further retention loss.

has been demonstrated to be an accurate measure of FEFET MW with the corresponding gate stacks [5], [10].

III. RETENTION LOSS MECHANISMS

Retention losses in these devices are driven by two mechanisms: (1) Given the charge-trap-like behavior of the MW enhancing charges at the GBL/TDL interfaces, the predominant MW enhancement loss mechanism is leakage/detrapping of these charges, and (2) Depolarization of the ferroelectric layer [18]. To understand the contributions of these mechanisms, the electric fields across different layers and the corresponding leakage current for both GBL and TDL gate stacks during the ERS state are simulated (Fig. 4(a-b)). The detrapping of charge is estimated by integrating the simulated leakage current through the GBL or, in the case of TDL, through gate side-HZO. Using this framework, we identify depolarization of the FE layer as the primary cause of RL in both TDL and GBL devices. The simulated electric

fields across the FE layer and GBL/TDL reveal a transition between the retention loss mechanisms depending on the relative amount of MW enhancing charges Q'_{it} with respect to the ferroelectric polarization P_r . When $Q'_{it} > P_r$, the depolarizing field is screened by the MW enhancing charges. However, when $Q'_{it} < P_r$, field across the ferroelectric becomes depolarizing. Given the large MWs, both device geometries are expected to be overcompensated at the FE-GBL/TDL interface immediately after programming. However, the drastic difference in leakage current between the thick FE-HZO in TDL and the GBL layer indicate the possibility of vastly different retention characteristics between the two configurations.

In GBL FEFETs, retention loss occurs in two stages: (1) when $Q'_{it} > P_r$, MW enhancing charges Q'_{it} detrapp from the FE-GBL interface due to the increased field across the GBL (and the high leakage through GBL), see Fig. 4(c). Under this condition, however, the field across the

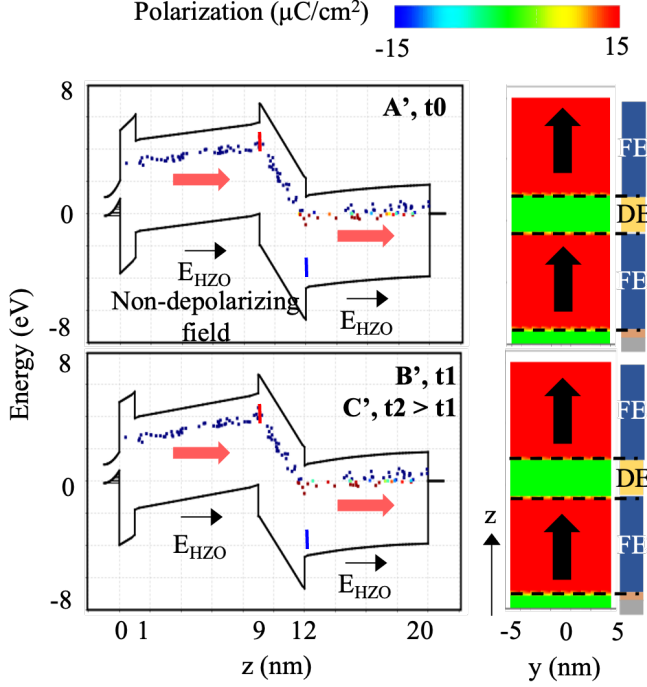


Fig. 5: Robust retention in TDL FEFET: When the dielectric is placed in the middle of the ferroelectric gate stack, the low leakage through the FE-HZO layers secures the MW enhancing charges at the FE-TDL interface. This overcompensated charge density at the interface maintains the non-depolarizing field across the FE leading to robust retention while exhibiting enhanced MW resulting from the charges at the FE-TDL interfaces.

ferroelectric layer is non-depolarizing. (2) Once $Q'_{it} < P_r$, the depolarization field across the FE layer increases due to decreased screening at the FE-GBL interface (Fig. 4(d-e)), leading to RL.

Moving the dielectric layer to the middle of the ferroelectric gate stack eliminates the loss of MW enhancing charges due to the low leakage through the gate side HZO (Fig. 5). This in turn preserves the non-depolarizing field across the FE layer (Fig. 5), providing a pathway for robust retention while exhibiting enhanced MW.

IV. EXPERIMENTAL RESULTS AND ARRHENIUS MODEL

Temperature-dependent retention was characterized in four fabricated FEFETs to confirm the hypothesis. The devices characterized have a width of $10 \mu m$ and a length of $3 \mu m$. The pulse scheme for retention characterization is shown in Fig. 6. The FEFETs are programmed to the program (low V_T) and erase (high V_T) states using a $10 \mu s$ write pulse and the transient $I_D - V_G$ is measured after different intervals. The temperature and time-dependent change in MW is modeled using a thermally activated depolarization model. The temperature dependence is captured by the activation energy,

E_a while the time dependence is captured by the exponent, n .

The evolution of $I_D - V_G$ characteristics at RT and the MW of all the FEFETs with time at different temperatures is shown in Fig. 7. The MW evolution is fit with the model and projected out to 10 years. The MW evolution in TDL based FEFETs 9/1(Al)/9 and 8/3(Al)/8 show similar time and temperature dependence to that of the reference FEFET 19. Importantly, the TDL and ref. FEFETs show minimal MW degradation at RT and the MW degradation worsens as temperature is increased. GBL based FEFET 14/4(Si), on the contrary, shows significantly higher MW degradation at RT and similar degradation at higher temperatures. It is also worth noting that the initial MW in GBL FEFETs decreases with increasing temperature while that in ref. and TDL FEFETs stays constant with temperature.

In Fig. 8(a), the MW loss after $1e4$ s is shown to obey the Arrhenius relationship. The GBL FEFET 14/4(Si) exhibits a constant MW loss across temperatures while TDL and ref. FEFETs show increase MW loss with increasing temperature with 8/3(Al)/8 TDL FEFET showing the least retention loss.

The time exponent n , which is indicative of the polarization switching kinetics, is similar across all samples, suggesting similar depolarization dynamics. The MW evolution with time in the samples with GBL shows minimal temperature dependence as confirmed by the near-zero activation energy. On the contrary, the TDL based devices have a strong temperature dependence akin to the standard FEFETs, indicated by the higher activation energies. The difference in temperature dependence is strongly indicative of the different detrapping mechanisms at play in both the samples:

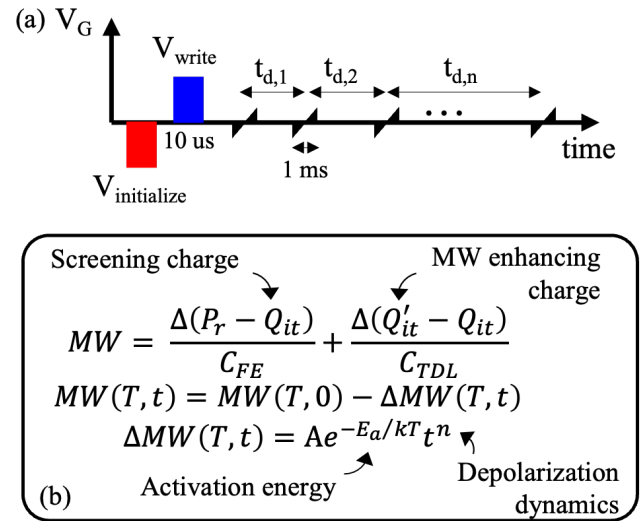


Fig. 6: Experiment and modeling: (a) Pulse scheme for retention characterization. (b) the evolution of the MW over time at different temperatures can be described using a time-dependent Arrhenius model to capture the depolarization dynamics.

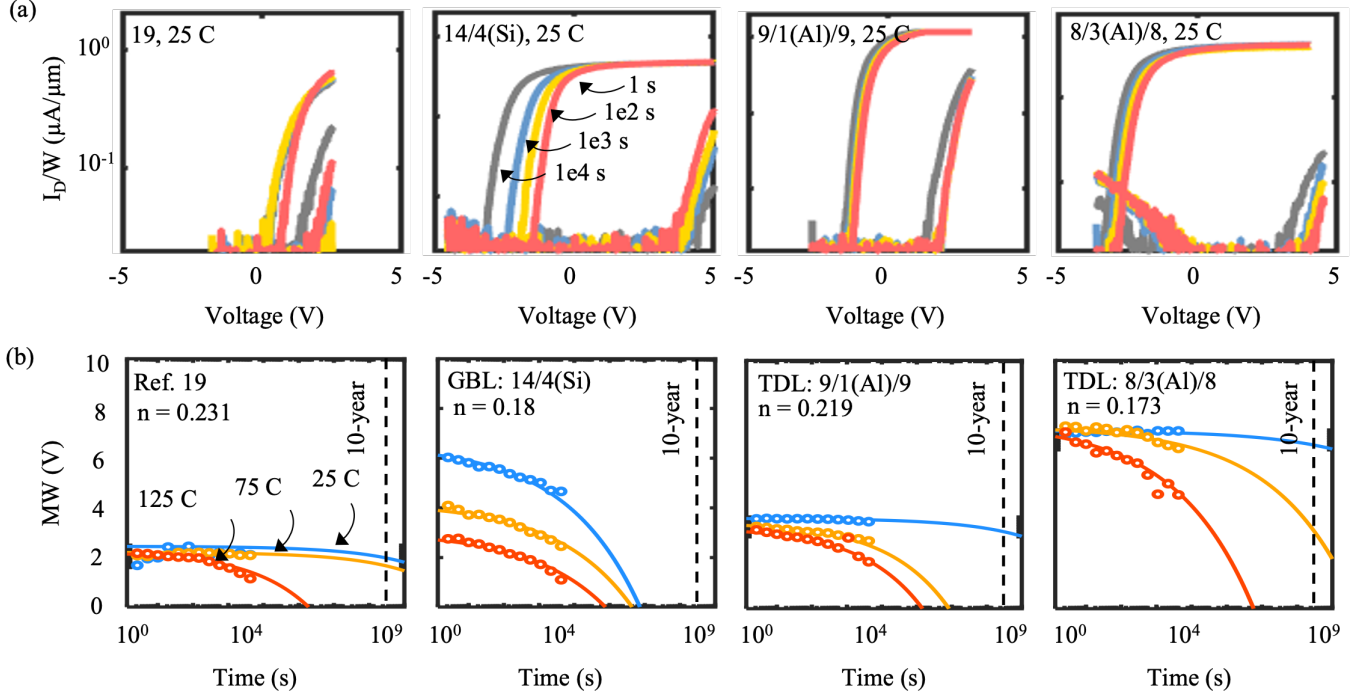


Fig. 7: (a) Evolution of transfer characteristics $I_D - V_G$ with time at 25°C for ref. FE (19 nm HZO), TDL (9 nm HZO/ 1 nm Al_2O_3 / 9 nm HZO, 8 nm HZO/ 3 nm Al_2O_3 / 8 nm HZO), GBL (14 nm HZO/ 4 nm SiO_2). The write voltage for the devices are 11, 14, 12 and 13 V, respectively. (b) Evolution of MW with time at 25-125°C. The Arrhenius model is used to fit the data and MW evolution is projected to 10 years.

temperature independent direct tunneling through the GBL dominates in the case of GBL-based FEFETs while trap-assisted tunneling, which has a strong temperature dependence, dominates in the case of TDL-FEFETs. Thus, in TDL-FEFETs, the depolarization fields are screened by the MW enhancing charges, enabling large MW, QLC compatible operation while exhibiting robust retention in accordance with the predication from the simulations. At higher temperatures, leakage across the gate side HZO increases, resulting in similar depolarization of HZO and related RL at rates comparable to GBL-FEFETs.

V. BENCHMARK

The benchmark of the MW after 10^4 s at RT compared to the initial MW indicates that GBL FEFETs, in general, exhibit significant RL (25-50%). TDL FEFETs, on the other hand, achieve similar enhanced MWs while exhibiting robust retention ($< 3\%$ RL) (Fig. 9). TDL FEFETs not only resolve retention loss for NAND applications but also provide a pathway for optimizing retention characteristics in standard FEFETs as evidenced by their superior retention characteristics compared to standard FEFETs. Importantly, the 8/3(Al)/8 TDL based FEFET reported in this work shows

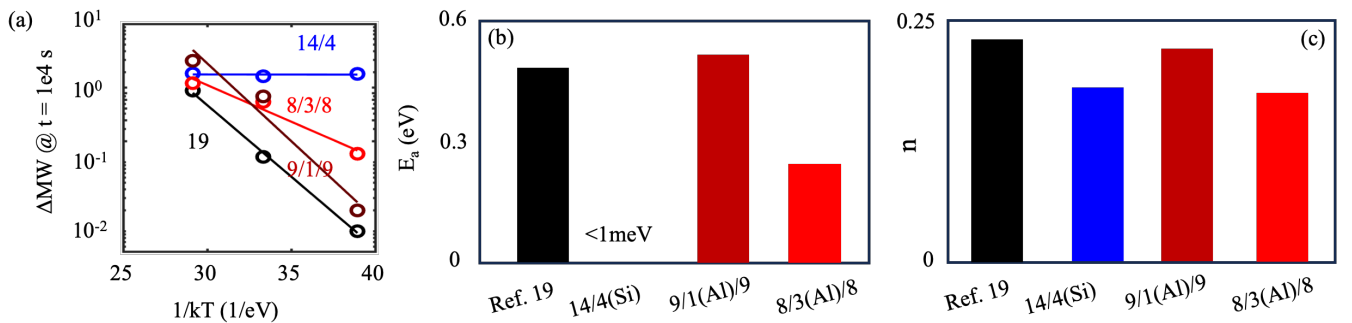
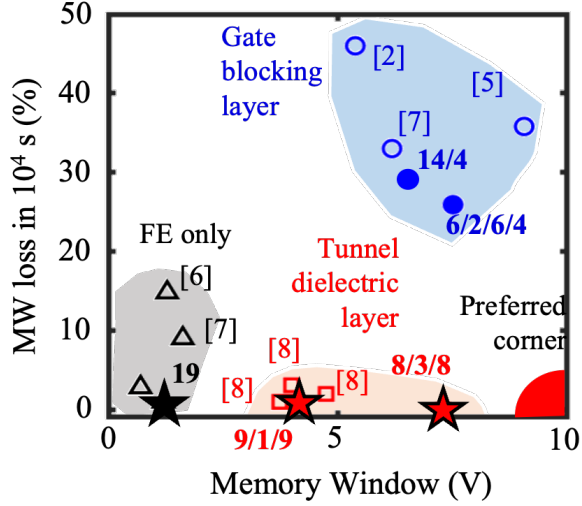


Fig. 8: (a) Arrhenius relationship describes the maximum MW degradation ($\Delta MW @ t = 1e4 s$) in FEFETs (b&c) The very low E_a of 14/4(Si) GBL FEFET indicates that the retention loss is independent of temperature. The comparable time exponent across all samples indicates similar time dependent behavior due to similar depolarization dynamics.



Department of the University of Modena and Reggio Emilia, Italy for funding.

Fig. 9: MW vs Retention loss benchmark of GBL, TDL and standalone FE gate stacks.

< 1% retention loss while providing a QLC compatible, 7.5 V MW.

While the higher thickness of the TDL is critical to achieving large MW [5], robust retention has also been demonstrated with even thinner Al_2O_3 TDL layers (2.4 Å) by Lee et al. highlighting the intrinsic retention advantage of TDL FEFETs [6].

VI. CONCLUSION

In a first-of-its-kind comprehensive temperature-dependent retention study of FEFETs, we show that depolarization of the FE layer is the main cause of retention loss. Through modeling and experiments, we showed that placing the dielectric insert for MW enhancement in the middle of the FE gate stack reduces depolarization significantly by preventing the detrapping of the MW-enhancing charges. Thereby, energetically stabilizing the enhanced MW. On the contrary, GBL based FEFETs are shown to undergo significant retention loss even at RT due to detrapping of the MW enhancing charges and resulting depolarization of HZO. The minimal temperature dependence of retention characteristics in GBL based FEFETs reveals that the MW loss is driven by temperature-independent direct tunneling of the MW enhancing charges through the GBL while the loss of MW enhancing charges in TDL based FEFETs is driven by temperature-dependent trap-assisted tunneling.

ACKNOWLEDGEMENT

This work was supported by Samsung Electronics and SUPREME, one of the seven SRC-DARPA JUMP 2.0 centers. Fab was done at the IEN, supported by the NSF-NNCI program (ECCS-1542174). Padovani acknowledges the FAR 2023-2024 project of the “Enzo Ferrari” Engineering

REFERENCES

- [1] S. Lim, T. Kim, I. Myeong, S. Park, S. Noh, S. M. Lee, J. Woo, H. Ko, Y. Noh, M. Choi, *et al.*, “Comprehensive design guidelines of gate stack for qlc and highly reliable ferroelectric vnand,” in *2023 International Electron Devices Meeting (IEDM)*, pp. 1–4, IEEE, 2023.
- [2] S.-H. Kuk, K. Ko, B. H. Kim, J.-H. Han, and S.-H. Kim, “Unstable retention behavior in mifis fefet: Accurate analysis of the origin by absolute polarization measurement,” *arXiv preprint arXiv:2406.19618*, 2024.
- [3] T. Hu, X. Sun, M. Bai, X. Jia, S. Dai, T. Li, R. Han, Y. Ding, H. Fan, Y. Zhao, *et al.*, “Enlargement of memory window of si channel fefet by inserting Al_2O_3 interlayer on ferroelectric $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$,” *IEEE Electron Device Letters*, 2024.
- [4] T. Hu, X. Shao, M. Bai, X. Jia, S. Dai, X. Sun, R. Han, J. Yang, X. Ke, F. Tian, *et al.*, “Impact of top SiO_2 interlayer thickness on memory window of si channel fefet with $\text{tin}/\text{SiO}_2/\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2/\text{SiO}_2/\text{Si}$ (mifis) gate structure,” *IEEE Transactions on Electron Devices*, 2024.
- [5] D. Das, H. Park, Z. Wang, C. Zhang, P. V. Ravindran, C. Park, N. Afroze, P.-K. Hsu, M. Tian, H. Chen, *et al.*, “Experimental demonstration and modeling of a ferroelectric gate stack with a tunnel dielectric insert for nand applications,” in *2023 International Electron Devices Meeting (IEDM)*, pp. 1–4, IEEE, 2023.
- [6] H. J. Lee, S. Nam, Y. Lee, K. Kim, D.-H. Choe, S. Yoo, Y. Park, S. Jo, D. Kim, and J. Heo, “Laminated ferroelectric fet with large memory window and high reliability,” *IEEE Transactions on Electron Devices*, 2024.
- [7] P. Venkatesan, C. Park, T. Song, L. Fernandes, D. Das, N. Afroze, P. G. Ravikumar, M. Tian, H. Chen, W. Chern, *et al.*, “Disturb and its mitigation in ferroelectric field-effect transistors with large memory window for nand flash applications,” *IEEE Electron Device Letters*, 2024.
- [8] A. Goda, “Recent progress on 3d nand flash technologies,” *Electronics*, vol. 10, no. 24, p. 3156, 2021.
- [9] J. Han, S. Kang, K. Kim, J. Jang, and J. Song, “Fundamental issues in vnand integration toward more than 1k layers,” in *2023 International Electron Devices Meeting (IEDM)*, pp. 1–5, IEEE, 2023.
- [10] D. Das, H. Park, Z. Wang, C. Zhang, P. V. Ravindran, C. Park, N. Afroze, P.-K. Hsu, M. Tian, H. Chen, *et al.*, “Ferroelectric gate stack engineering with tunnel dielectric insert for achieving high memory window in fehets for nand applications,” in *2024 8th IEEE Electron Devices Technology & Manufacturing Conference (EDTM)*, pp. 1–3, IEEE, 2024.
- [11] S. Yoon, S.-I. Hong, D. Kim, G. Choi, Y. M. Kim, K. Min, S. Kim, M.-H. Na, and S. Cha, “Qlc programmable 3d ferroelectric nand flash memory by memory window expansion using cell stack engineering,” in *2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, pp. 1–2, IEEE, 2023.
- [12] N. Tasneem, Z. Wang, Z. Zhao, N. Upadhyay, S. Lombardo, H. Chen, J. Hur, D. Triyoso, S. Consiglio, K. Tapily, *et al.*, “Trap capture and emission dynamics in ferroelectric field-effect transistors and their impact on device operation and reliability,” in *2021 IEEE International Electron Devices Meeting (IEDM)*, pp. 6–1, IEEE, 2021.
- [13] N. Tasneem, Z. Wang, H. Chen, S. Yu, W. Chern, and A. Khan, “Immediate read-after-write capability in p-type ferroelectric field-effect transistors and its evolution with fatigue cycling,” *IEEE Transactions on Device and Materials Reliability*, vol. 23, no. 1, pp. 142–146, 2023.
- [14] C. Park, P. V. Ravindran, D. Das, P. G. Ravikumar, C. Zhang, N. Afroze, L. Fernandes, Y. H. Kuo, J. Hur, H. Chen, *et al.*, “Plasma-enhanced atomic layer deposition based ferroelectric field-effect transistors,” *IEEE Journal of the Electron Devices Society*, 2024.
- [15] L. Fernandes, P. V. Ravindran, T. Song, D. Das, C. Park, N. Afroze, M. Tian, H. Chen, W. Chern, K. Kim, *et al.*, “Material choices for tunnel dielectric layer and gate blocking layer for ferroelectric nand applications,” *IEEE Electron Device Letters*, 2024.
- [16] L. Fernandes, P. V. Ravindran, J. Chen, M. Tian, D. Das, H. Chen, W. Chern, K. Kim, J. Woo, S. Lim, *et al.*, “Optimizing memory window for ferroelectric nand applications: An experimental study on dielectric material selection and layer positioning,” *IEEE Transactions on Electron Devices*, 2024.
- [17] “<https://www.appliedmaterials.com/sg/en/semiconductor/ginestra-software.html>,”
- [18] M. Pešić, B. Beltrando, A. Padovani, T. Miyashita, N.-S. Kim, and L. Larcher, “Electron-assisted switching in fehets: Memory window dynamics—retention—trapping mechanisms and correlation,” in *2022 IEEE International Reliability Physics Symposium (IRPS)*, pp. 4A–1, IEEE, 2022.