Threshold Selector and Capacitive Coupled Assist Techniques for Write Voltage Reduction in Metal–Ferroelectric–Metal Field-Effect Transistor

Siddhartha Raman Sundara Raman, S. S. Teja Nibhanupudi, *Graduate Student Member, IEEE*, Atanu K. Saha, *Student Member, IEEE*, Sumeet Gupta, *Senior Member, IEEE*, and Jaydeep P. Kulkarni, *Senior Member, IEEE*

***Abstract*— We propose device and circuit assist techniques to lower the ferroelectric–metal field-effect transistor (FeMFET) write voltage while lowering the effect of depolarizing field. A bipolar threshold selector (TS) is connected between the intermediate node of a FeMFET and V_{SS}, which provides a low-impedance bypass path by triggering an insulator to metallic transition during a write operation. This lowers the voltage drop across MOSFET and reduces the write voltage to ∼1.7 V. For further reduction in write voltage, the MOSFET in the FeMFET is utilized as a circuit assist by repurposing it as a capacitive coupling device. It reduces the write voltage to ∼1.4 V. During the read mode, TS is in insulating state and does not alter the capacitive voltage divider action. Read is followed by a data-dependent write-back stage to bring the polarization close to retention polarization. During the retention mode, TS acts as a weak bleeder resistor, reducing noise coupling and depolarizing field. TS parameter sensitivity for write voltage reduction along with row hammer effect of the proposed read- and write-back operation is also presented.**

***Index Terms*— Aspect ratio, capacitive coupling, depolarizing field, ferroelectrics, threshold selector (TS), write voltage.**

I. Introduction

FERROELECTRIC field-effect transistor (FeFET) is a promising technology for data-centric computing applications as it offers dense one-transistor nonvolatile bitcell, high speed, disturb-free read operation, and use of CMOS process compatible materials and can be integrated in back end of the line with high-performance logic transistors [1]–[4]. Despite the promising attributes, FeFET advancements face significant challenges in the successful adoption in mainstream CMOS. These barriers include: 1) high program/erase voltage (collectively referred to as write voltage, ±4 V) compared to the CMOS logic operating voltage range and 2) large depolarizing field across the ferroelectric capacitor (FeCAP) degrading its retention time. These fundamental challenges if not addressed would render the FeFET technology limited only to niche applications, in spite of having many desired attributes and good material compatibility with CMOS. Multiple FeFET variants are being explored to address these critical challenges. FeFETs are realized with ferroelectric hafnium zirconium oxide (HZO) embedded between the gate and interlayer dielectric (ILD), as shown in Fig. 1(a). The higher permittivity of HZO in contrast to SiO_2 (ILD) incurs a large voltage drop across ILD, which exacerbates reliability issues due to ultrathin ILD experiencing a large electric field. For improving FeFET reliability and to reduce the write voltage by maximizing the drop across HZO, a recessed channel FeFET [R-FeFET, Fig. 1(b)] [5] has been proposed, which increases the electric field across the HZO layer due to the circular geometry of the channel lowering the write voltage to some extent (∼3 V). The application of this technology is limited because of the inherent fabrication difficulties. Ferroelectric–metal field-effect transistor (FeMFET) shown in Fig. 1(c) integrates a dedicated FeCAP on top of a MOSFET gate terminal, thus enabling independent optimization of ferroelectric stack and MOSFET gate-stack [6]. FeMFET can lower the program voltage by reducing the capacitor aspect ratio, thus increasing the voltage drop across the FeCAP and lowering the write voltage to ∼1.8 V for an aspect ratio of 0.1. This device configuration, however, introduces a floating intermediate node between the ferroelectic capacitor and the MOSFET gate-stack, which is susceptible to the noise coupling and can increase the depolarizing field leading to degraded retention time. Also, floating n_1 node having finite voltage can increase MOSFET subthreshold leakage (contribution from unselected bitcells) during a read operation degrading the read sensing margin; hence, the read speed. For further reduction in the FeMFET write voltage, three FeCAPs are connected in parallel to the MOSFET gate terminal forming a multiple-FeMFET has been proposed (Fig. 1) [7]. Write operation is achieved by asserting one wordline at a time and the remaining two ferroelectic capacitors are connected in parallel with the MOSFET, thus effectively increasing the MOSFET capacitance, reducing the drop across the ferroelectic capacitor. This lowers the write

Manuscript received September 19, 2021; accepted October 17, 2021. This work was supported by The University of Texas at Austin of NSF under Grant 1815616. The review of this article was arranged by Editor A. M. P. Anantram. *(Corresponding author: Siddhartha Raman Sundara Raman.)*

Siddhartha Raman Sundara Raman, S. S. Teja Nibhanupudi, and Jaydeep P. Kulkarni are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: s.siddhartharaman@utexas.edu; jaydeep@austin.utexas.edu).

Atanu K. Saha and Sumeet Gupta are with the Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TED.2021.3121348.

Digital Object Identifier 10.1109/TED.2021.3121348

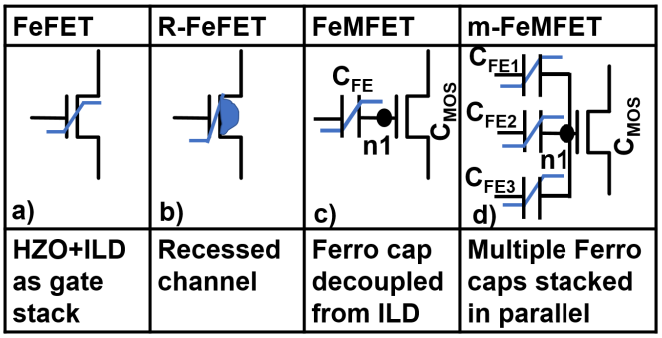

Fig. 1. (a) FeFET variants: FeFET with integrated HZO and ILD. (b) Recessed channel FeFET (R-FeFET). (c) Ferroelectric–metal–ferroelectric (FeMFET) decoupling the ferrocapacitor from the MOSFET gate capacitor. (d) Multiple ferroelectic capacitors (m-MFM) in parallel stacked with MOSFET.

voltage to ∼1.8 V. However, write operation requires three steps resulting in increased write time. In addition, frequent switching of the intermediate (n_1) node can increase depolarizing field across unselected FeCAPs, which can degrade the retention behavior.

In this article, we propose a unique design that uses a bipolar threshold selector (TS) device as a FeMFET write assist to reduce its write voltage independent of the capacitor aspect ratio in Section III. We also present a capacitive coupling write-assist circuit technique in Section IV, by repurposing the MOSFET in a FeMFET to act as a capacitive coupling write assist, which increases voltage across the FeCAP enabling further reduction in the write voltage. Section V discusses the sensitivity of the FeFET program voltage to different TS parameters. A stepped wordline approach to perform a read operation followed by a data-dependent write back is also presented as part of this article in Section VI. The effect of different retention times affecting the read current margins along with the effect of row hammer on polarization values and read current margins is also presented in Section VII. The impact of TS endurance on different operations is discussed in Section VIII. It is important to mention that the proposed bitcell does not add additional area overhead and TS can be integrated back end of the line.

II. Device Calibration

A. Threshold Selector

Bipolar TS devices utilize phase transition materials, such as oxides like VO_2 and NbO_2 or doped transition metal oxides like Ag-doped HfO_x and Cu-doped HfO_x [8]–[10]. TS undergoes an abrupt resistance switching from insulating to metallic state when the voltage across it is greater than V_{IMT} (insulating-to-metallic transition) threshold and vice versa for voltages lesser than V_{MIT} [metallic-to-insulating transition (MIT)] threshold [Fig. 2(a)]. R_{MET} and R_{INS} scale with length (thickness) (L_{TS}) and cross-sectional area of TS (A_{TS}) similar to a regular resistor [11], [12] as given in (1) and (2) where ρ_{INS} and ρ_{MET} correspond to the resistivity of the device in insulating and metallic state, respectively. The parameters of TS are chosen in accordance with the range of values observed

TABLE I
Parameters of TS

Symbol	Quantity	Chosen values
V_{IMT}	Insulating to metallic transition voltage	0.6V
V_{MIT}	Metallic to insulating tranasition voltage	20mV
R_{MET}	Resistance in metallic state	600 Ω
R_{INS}	Resistance in insulating state	0.12 GΩ
T_{TS}	Switching time	50ps

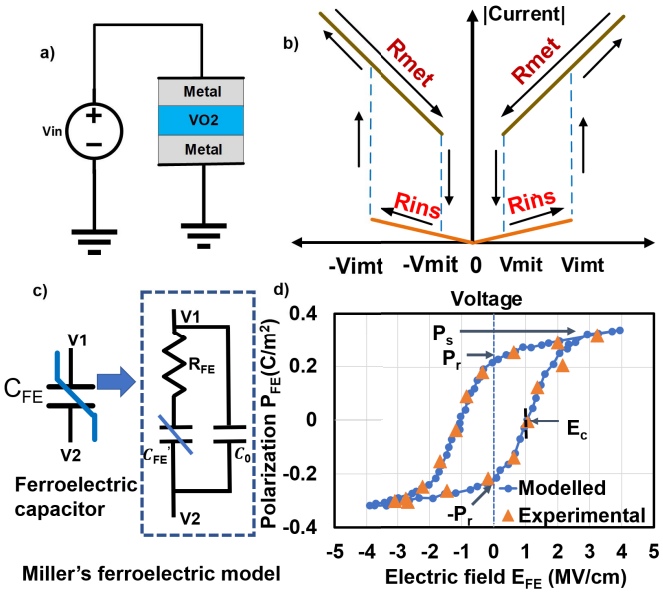

Fig. 2. (a) TS made of metal, transition metal oxides like VO_2. (b) TS I–V characteristics. (c) Miller's ferroelectric model having linear capacitor in parallel with a series combination of R_{FE} and FeCAP. (d) Q–V_{FE} characteristics calibration with the experimental device [14].

experimentally in [12]

$$R_{INS} = \rho_{INS} * L_{PTM}/A_{PTM} \quad (1)$$

$$R_{MET} = \rho_{MET} * L_{PTM}/A_{PTM}. \quad (2)$$

V_{IMT} and V_{MIT} are independent of A_{TS} and are dependent only on L_{TS} as shown in (3) and (4), where J_{CIMT} and J_{CMIT} refer to the current density required to bring about the transition from insulating to metallic and metallic to insulating, respectively. TS parameters are listed in Table I

$$V_{MIT} = I_{MIT} * R_{MET} = J_{CMIT} * \rho_{MET} * L_{PTM} \quad (3)$$

$$V_{IMT} = I_{IMT} * R_{INS} = J_{CIMT} * \rho_{INS} * L_{PTM}. \quad (4)$$

B. Ferroelectric Capacitor

The hysteretic and multidomain polarization switching characteristics of the FeCAP has been modeled based on the Preisach-based Miller's equations [14]–[16]. The equivalent circuit of the FeCAP is shown in Fig. 2(c) that consists of one resistive (R_{FE}) and two capacitive (C_{FE} and C_0) elements. The overall charge density (Q) in the FeCAP incorporates both the contribution from spontaneous polarization (P) and induced displacement due to the background permittivity of the ferroelectric layer (ϵ_r). Thus, $Q = P + \epsilon_0\epsilon_r E_{FE}$, where E_{FE} ($=V_{FE}/T_{FE}$) is the average electric field, V_{FE} is the voltage drop, and T_{FE} is the thickness of the ferroelectric layer. The applied voltage-dependent change in P is captured by C_{FE} and the linear displacement component is modeled as C_0. Here,

TABLE II
PARAMETERS OF FERROELECTRIC MATERIAL [14]

Symbol	Quantity	Chosen values
τ_{FE}	Time for polarization	1ns
P_r	Remnant polarization of ferro-electric	0.19 C/m2
P_s	Saturation polarization of ferro-electric	0.25 C/m^2
E_c	Coercive of ferroelectric	1.1 MV/cm

the resistance R_{FE} delays the charging of C_{FE} by mimicking the time lag ($\tau_{FE} = R_{FE}C_{FE}$) in the polarization switching with respect to V_{FE} [16] (5). Therefore, the magnitude of effective voltage drop across C_{FE} ($V_{C_{FE}}$) is always less than $|V_{FE}|$. To capture the history effect and the minor loop formation due to the multidomain polarization switching, the calculation of C_{FE} is carried out based on the following set of equations:

$$P_M = P_S \times \tanh\left(\frac{V_{C_{FE}/T_{FE}} \pm E_C}{2\delta}\right) \tag{5}$$

$$\delta = \alpha \times ln\left(\frac{P_s + P_r}{P_s - P_r}\right)^{-1} \tag{6}$$

$$\Gamma = 1 - \tanh\left(\sqrt{\frac{P - P_M}{\pm P_S - P}}\right) \tag{7}$$

$$C_{FE} = \Gamma \times \frac{dP_M}{dV_{C_{FE}}}. \tag{8}$$

Here, $P = \int j_{C_{FE}} dt$, $j_{C_{FE}} = C_{FE} \times (dV_{C_{FE}}/dt)$, E_C is the coercive electric field, P_r is the remanent polarization, P_s is the saturation polarization, and α is a fitting parameter to model the distribution of coercive field in the FeCAP. Furthermore, C_0 is computed based on the following equation:

$$C_0 = \frac{\epsilon_0 \epsilon_r}{T_{FE}}. \tag{9}$$

The model parameters are calibrated (shown in Table II) with experimentally measure Q–V_{FE} characteristics of 10-nm HZO and both the measured and simulated characteristics are shown in Fig. 2(d). A 45-nm high-K planar FET is used for the underlying MOSFET [15].

III. TS DEVICE AS FEMFET WRITE-ASSIST

For the baseline FeMFET with write access transistor (with WWL turned on), the intermediate node (n_1) voltage rises to a finite value based on the capacitive voltage divider ratio between the ferroelectic capacitor and the MOSFET gate capacitance, as shown in Fig. 3(a). Simulation results indicate that significant voltage is dropped across the MOSFET increasing the required write voltage considerably [Fig. 3(b)]. For the proposed TS-assisted FeMFET (TS-FeMFET) device, a backend integrated bipolar TS is connected between the n_1 node and V_{ss}. The proposed TS-FeMFET device structure as shown in Fig. 3(c) lowers the write voltage by decoupling its read and write paths. During a write operation, when the n_1 voltage exceeds (V_{IMT}), the TS transitions into low-impedance metallic state (R_{MET}) after the TS switching time (T_{TS}) and provides a bypass current path to V_{SS}. This discharges the n_1 node toward V_{SS} and the entire write voltage is now dropped across the FeCAP lowering the required write voltage to ~1.7 V. This finite ~0 V at n_1 node brings about an MIT of TS, further resulting in charging of n_1 node via R_{FE}. The

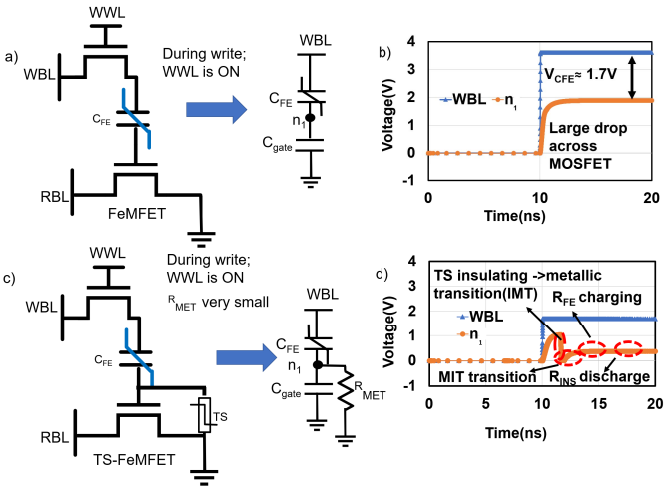

Fig. 3. (a) Baseline FeMFET with floating n_1 node; voltage division between ferroelectic capacitor and gate capacitance of MOSFET during write operation. (b) FeMFET write simulation results showing finite voltage at n_1 node. (c) Proposed TS-FeMFET device structure with a TS connected between n_1 node and V_{SS} and transition to low-impedance metallic state in TS. (d) TS-FeMFET simulation results showing write-bypass mechanism due to metallic state TS transition lowering write voltage, transitioning back into insulating state, thus charging the n_1 node via R_{FE} and discharging the n_1 node via the bleeding path through R_{INS}.

charge buildup at the n_1 node decreases with time because of the weak bleeder path to ground via the R_{INS} of TS, as shown in Fig. 3(d). It is important to note that the above analysis is performed with a unit aspect ratio. This is done to show that without any reduction in aspect ratio, a considerable reduction in write voltage is observed.

IV. CAPACITIVE COUPLING CIRCUIT ASSIST TECHNIQUE

During a FeMFET write operation, the MOSFET acts as a capacitive voltage divider. However, in the TS-FeMFET write operation, the MOSFET is bypassed by the TS to achieve low write voltage. We propose to repurpose MOSFET to act as a capacitive coupling write assist to modulate the n_1 node voltage for further reduction in the write voltage.

As shown in Fig. 4(a), after wordline (WWL) is asserted to positive voltage during a write operation, bitline (RBL) and sourceline (SL) nodes are initially at 0 and are transitioned to −0.3 or 0.3 V depending on whether the bitcell is programmed/erased. These rising/falling RBL and SL transitions are coupled onto the n_1 node with MOSFET acting as a coupling capacitor. This increases the effective voltage across the FeCAP, which lowers the required WBL voltage from 4 to ~3.3 V [Fig. 4(b)]. A similar capacitive coupling approach can be utilized for TS-FeMFET as shown in Fig. 4(c), which lowers the required write voltage to ~1.4 V [Fig. 4(d)]. Positive RBL/SL voltage transition can be increased to Vcc but kept at +0.3 V to have a symmetric capacitive coupling effect. For effective capacitive coupling, the RBL/SL should transition from 0 to −0.3 V once the TS device transitions into metallic state first and then back into the insulating state. First TS metallic state transition ensures the n_1 node is initially discharged to V_{SS}. When the voltage across TS goes below V_{MIT}, it transitions back into the high-impedance insulating state. RBL/SL toggling results in n_1 node modulation due to

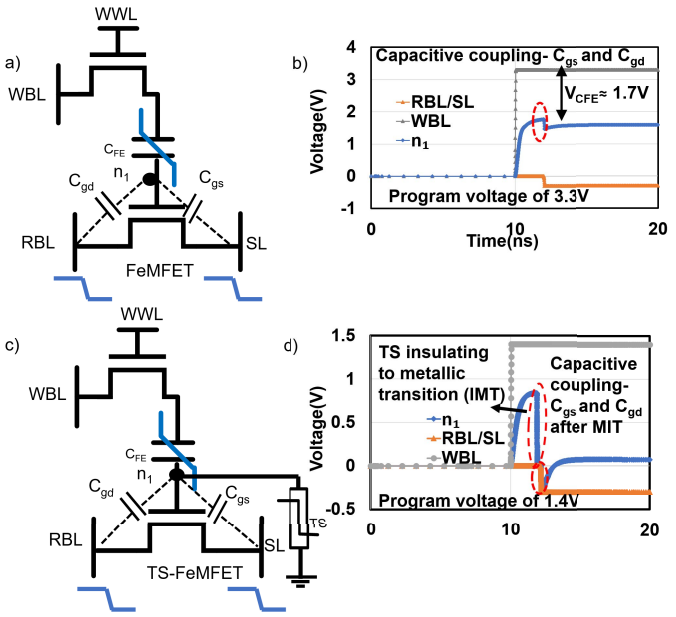

Fig. 4. (a) FeMFET with RBL and SL negative coupling. (b) Simulation waveforms showing capacitive coupling effect in lowering the write-voltage. (c) TS-FeMFET with RBL and SL negative coupling. (d) Simulation waveforms capacitive coupling effect in TS-FeMFET.

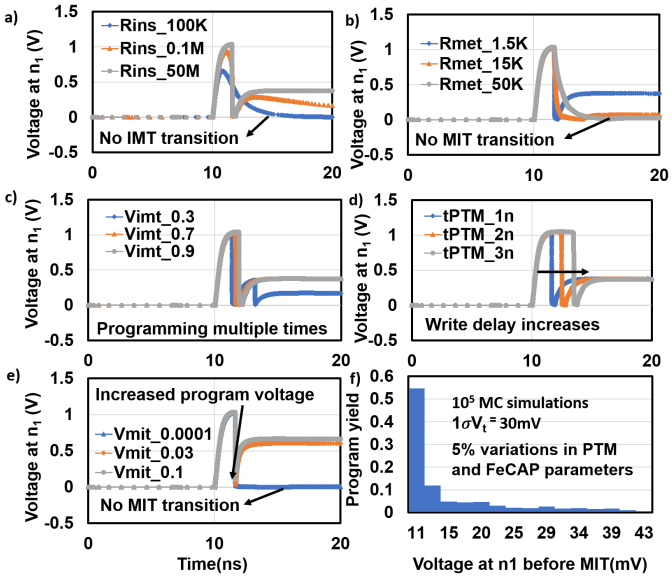

Fig. 5. (a) R_{INS} constraints for successful operation of TS-FeMFET. (b) Write delay increases with a risk of no MIT transition as R_{MET} increases. (c) Reduced V_{IMT} leads to risk of programming multiple times. (d) Increase in TS switching time causes a delayed IMT on n_1 node, increasing the write time. (e) Reduced V_{MIT} leads to risk of no MIT and increased V_{MIT} increases program voltage. (f) Program yield as a function of different voltages at n_1 before MIT in the presence of process variations.

capacitive coupling effect. The TS being in insulating state charges the n_1 node via R_{FE} and decreases back to V_{SS} via R_{INS}. This increases the effective voltage drop across the FeCAP, which aids in lowering the required write voltage to ∼1.4 V [Fig. 4(d)]. In the advanced CMOS technology nodes employing high-κ gate dielectrics with scaled effective oxide thickness, MOSFET gate capacitance will be higher and would result in increased capacitive coupling effect, enabling further reduction in FeMFET write voltage.

It is important to note that the proposed technique of applying a negative RBL transition of 0.3 V is different from that of applying WBL of 1.7 V with a body voltage of −0.3 V. This is because in the case of former, RBL transition is brought about only after the PTM transitions into insulating state (after being in the metallic state), unlike the latter case where the body voltage has to remain at −0.3 V, throughout the write cycle. Also, in case of unaccessed cells in the same column, voltage developed at n_1 node because of negative RBL transition (C_{gd} coupling) is brought back closer to 0 over a period of time through the R_{INS} of TS.

V. Design Space Exploration

During the TS-FeMFET write operation, the TS should develop sufficient voltage across it for transitioning from the insulating to metallic state. This guides the TS parameter optimization for R_{INS} and V_{IMT}. R_{INS} should be large enough to allow voltage across the TS to exceed V_{IMT} threshold (to transition TS into metallic state) during the initial voltage division between the ferroelectic capacitor and the MOSFET gate capacitance. For an R_{INS} value of 10 K, TS does not transition into metallic state, as shown in Fig. 5(a). After the TS transitions into the metallic state, to maintain the n_1 node voltage close to ∼V_{ss}, R_{MET} of the TS should be chosen as small as possible to ensure that write delay impact is minimal, as shown in Fig. 5(b). Furthermore, large values of R_{MET} pose a potential risk of TS not transitioning into insulating state post write because of increased write delay.

The lower limit of V_{IMT} is set by the risk of programming TS-FeMFET multiple times [Fig. 5(c)] and V_{IMT} should be less than the WBL voltage applied so as to ensure the TS transitions into metallic state. As the TS switching time increases, the transition from insulating to metallic state takes a longer time, thus increasing write delay to polarize the FeCAP [Fig. 5(d)]. Similarly, TS should transition into insulating state, as n_1 node is being discharged toward V_{ss} posing a minimum constraint on V_{MIT}. V_{MIT} values less than 0.1 mV does not ensure transition of TS into insulating state, as shown in Fig. 5(e). Also, as V_{MIT} increases, the voltage drop across FeCAP would decrease, reducing the savings in WBL voltage. Variability study is performed using 10^5 Monte Carlo simulations assuming 1σ variation of MOSFET threshold voltage to be 30 mV with 5% variations in the TS and FeCAP parameters. Voltage at n_1 before TS transitions into the insulating state is a measure of program voltage and program yield is a fraction of the total MC runs with a particular voltage at n_1 node before MIT. It can be inferred from Fig. 5(f) that the program voltage that is the difference between the voltage at WBL and the voltage at n_1 before MIT ranges from 1.69 to 1.67 V even in the presence of process variations. For the capacitive coupling circuit assist technique, the MOSFET capacitance should be large enough in comparison with the FeCAP to ensure a good coupling ratio and hence lower potential drop across MOSFET. The time difference between WBL activation and the RBL/SL transition should be longer than twice the switching time of TS (to transition into metallic state first then into the insulating state) to efficiently couple the RBL/SL node onto the n_1 node.

VI. Read and Retention Analysis

During the retention mode with WBL and RBL = 0, the TS is in insulating state and acts as a weak bleeder resistor connected between the n_1 node and V_{SS} preventing charge build-up and/or noise coupling at the n_1 node. The voltage at n_1 node is dependent on the retention duration and the polarization settles to a value closer to the remnant polarization once the voltage at the n_1 node settles to a value closer to 0.

During a read operation, WWL is asserted, and in general, a smaller WBL (700 mV) voltage (read disturb) is applied resulting in smaller voltage rise at the n_1 node. RBL voltage is applied at the drain of the MOSFET to ensure a drain-to-source current. The drain-to-source current is used to differentiate between the bits stored in the FeMFET. WBL voltage is sufficient enough to partially polarize the FeCAP with the value of C_{FE}' being different based on the direction of polarization embedded in the FeCAP. This difference in FeCAP values results in different capacitive coupling between the gate capacitance of MOSFET and FeCAP, resulting in different voltages at the n_1 node, further causing different read currents. The TS device is optimized such that the voltage rise at the n_1 node during a read operation is smaller than the V_{IMT} threshold, which keeps the TS device in the highly resistive insulating state and does not affect the capacitive voltage divider action. It is important to note that the read operation alters the retention polarization because of the partial polarization in the direction of applied voltage at WBL. Read is followed by a "data-dependent" write-back step to bring the polarization of FeCAP closer to the value during retention state post read. The write-back stage is different from the program or erase operation because TS remains in insulating state, unlike a write operation where TS is in metallic state. The data-dependent write-back step involves using WBL of 0 V for the case of program and −1.7 V in case of erase with RBL = 0, similar to a write operation (Table III). Also, the WWL is kept ON during the write-back operation as well. The write-back voltage on WBL can be tuned postsilicon and ensures that FeCAP does not fall into minor loops. It is important to note that the write-back voltage is different for the program and erase operations. This is because the read disturb voltage polarizes the FeCAP by different amounts when the FeCAP polarization is in the +P- or −P-direction. From Fig. 2(d), post-retention operation, the polarization is close to +Pr or −Pr depending on the bitcell content. The same read disturb voltage polarizes the FeCAP by different amounts resulting in different polarizations during read (the slope on the +P side is different from the slope on the −P side). To counteract this difference in polarization brought about by the read operation, the write-back voltages are different between program and erase. The read operation can be preceded by: 1) a write followed by a short retention period or 2) a write followed by long retention.

A. Case 1: Write–Short Retention–Read

A program operation involves polarizing the FeCAP in the positive direction and an erase operation involves polarizing the FeCAP in the negative direction. A retention operation decreases the magnitude of polarization in both cases and

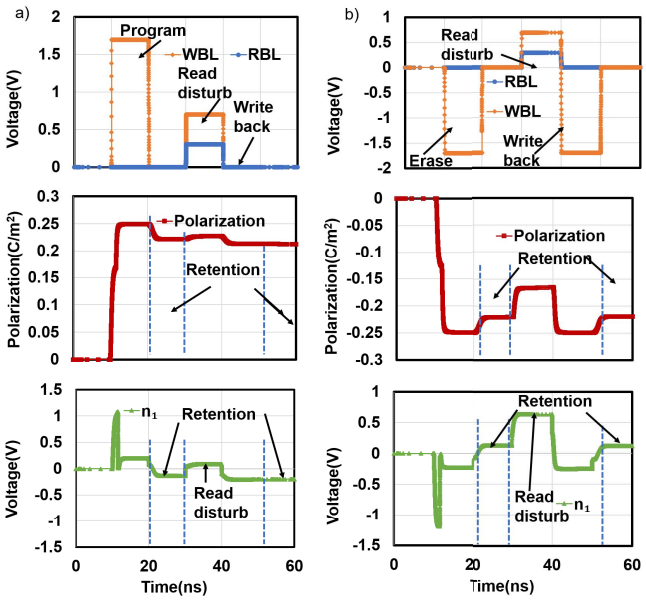

Fig. 6. (a) Variation of voltage at n_1 node, polarization, WBL, and RBL for programming with a small retention time. (b) Variation of voltage at n_1 node, polarization, WBL, and RBL for erasing with a small retention time.

TABLE III
Operation and Voltages

Operation	WBL	RBL	SL
Program	1.7V	0V	0V
Erase	-1.7V	0V	0V
Read	0.7V	0.3V	0V
Writeback - program	0V	0V	0V
Writeback - erase	-1.7V	0V	0V

the magnitude of polarization depends on the voltage at n_1, which further depends on the retention time. The read disturb voltage disturbs the polarization further and the write-back voltage is necessary to bring the polarization back closer to the polarization present during the retention operation (once the WBL and RBL are turned off post-read and write-back operation). As shown in Fig. 6(a) and (b), the read current (drain-to-source current for nMOS) for program is in the range of nanoamperes and the read current for erase is in the range of hundreds of microamperes, sufficient enough to distinguish between the values stored.

B. Case 2: Write–Long Retention–Read

The polarization of FeCAP reaches $+P_r$ in case of program and $-P_r$ in case of erase. This polarization is disturbed by a small voltage at the WBL during the read operation. Furthermore, the write-back step helps in bringing the n_1 node closer to the voltage during retention Since the voltage at the n_1 node reaches closer to 0 due to MOSFET gate leakage [17] irrespective of whether program or erase is program, there is no any difference in the node n_1 voltage before the read operation. This results in a slightly decreased read margin in contrast to case 1). As shown in Fig. 7(a) and (b), the read current for program is in the range of tens of nanoamperes, and in the case of erase operation, the read current is in the range of hundreds of microamperes.

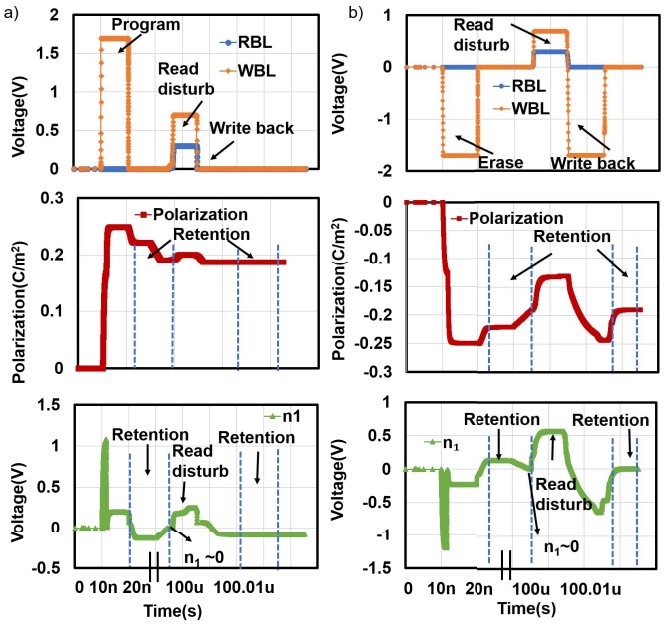

Fig. 7. (a) Variation of voltage at n_1 node, polarization, WBL, and RBL for programming with a long retention time. (b) Variation of voltage at n_1 node, polarization, WBL, and RBL for erasing with a long retention time.

VII. Row Hammer Effect

The variation of polarization as a function of number of times a bitcell has been accessed gives a measure of whether the polarization is retrieved back to the retention state after every bitcell access. Fig. 8 shows the variation of retention polarization as a function of cycle number. The variation in case of program and erase for long retention [Fig. 8(c) and (d)] is negligible, whereas the variation in case of small retention [Fig. 8(a) and (b)] is a bit larger than long retention. This is because of inaccuracy in determining the polarization values for short retention period as the voltage at n_1 node becomes closer to 0 V. The retention polarization value stabilizes after cycle count of around 50 in the case of program and erase because the changes in n_1 node voltage become negligible after 50 cycles.

The variation in polarization further affects the read current because of different C_{FE}', resulting in different voltages at the n_1 node. The read current decreases from 600 to 400 μA [Fig. 9(a)] in case of short retention for erase as the cycle count increases. This is because of increase in magnitude of retention polarization, resulting in decreased voltage at n_1 node, further decreasing the drain-to-source current. Furthermore, the read current increases from 4 to 20 nA [Fig. 9(b)] with cycle number in case of program because the magnitude of retention polarization starts decreasing, further resulting in an increased voltage at the n_1 node. The difference in voltage at the n_1 node across different cycles in case of long retention [Fig. 9(c) and (d)] is ~0 V resulting in decreased read current margin as opposed to a read preceded by short retention period.

VIII. Effect of TS Endurance

The endurance of TS plays an important role in determining the feasibility of the TS-FeMFET bitcell in a large-scale array design. The experimental analysis of TS has shown endurance

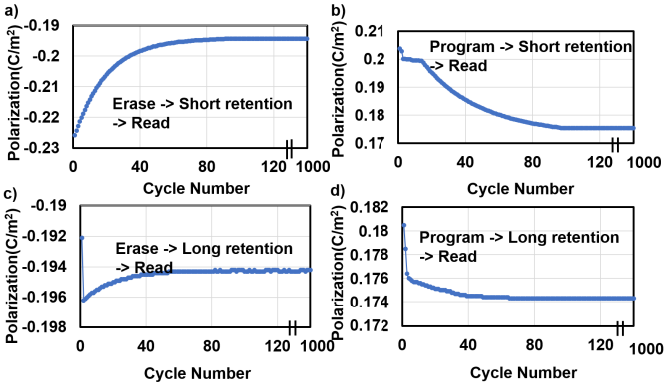

Fig. 8. (a) Variation in retention polarization as a function of cycle count saturates after 50 cycles for erase followed by short retention followed by read. (b) Variation in retention polarization as a function of cycle count saturates after 90 cycles for program followed by short retention followed by read. (c) Variation in retention polarization as a function of cycle count is minimal for erase followed by long retention followed by read. (d) Variation in retention polarization as a function of cycle count is minimal for program followed by long retention followed by read.

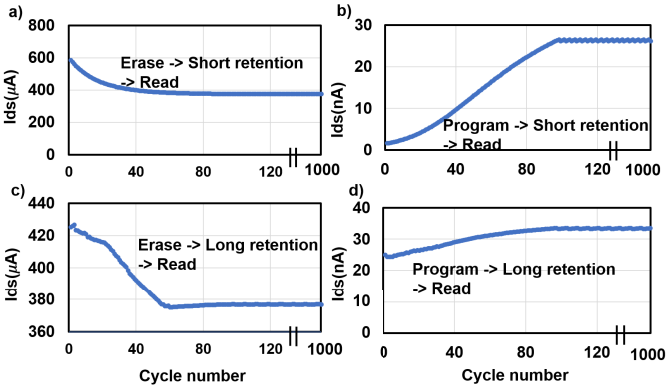

Fig. 9. (a) Variation in read current for erase followed by short retention followed by read as a function of cycle count and saturates at 50 cycles. (b) Variation in read current for program followed by long retention followed by read as a function of cycle count and saturates at 90 cycles. (c) Variation in read current for erase followed by long retention followed by read as a function of cycle count. (d) Variation in read current for program followed by long retention followed by read as a function of cycle count.

of around 10^9 [18] cycles. In case of FeMFET, the charge trapping issue observed in FeFET is eliminated, leading to a better endurance of 10^{10} cycles [6], which is in a similar range of the TS endurance. Both these devices are being researched considerably so as to improve the endurance cycles of both these technologies. In the worst case scenario of TS failing before the FeMFET, TS failure can happen either in metallic state or insulating state and during read or write or retention operation. This section analyzes the effect of TS failure in each of the operations.

A. Effect of Endurance on Write Operation

If TS is stuck in the metallic state during a write operation, there is no increase in the write voltage because the FeFET is programmed/erased when the TS is in the metallic state. If TS is in the insulating state, the write voltage needs to be high enough to program/erase because the voltage at n_1 node

TABLE IV
EFFECT OF TS ENDURANCE

Operation	TS state	Failure impact
Write	Metallic	No impact - Program voltage $\sim 1.7V$
	Insulating	Behaves like FeMFET
Read	Metallic	Reduced read margin
	Insulating	No impact
Retention	Metallic	Reduced noise coupling issues
	Insulating	No impact
Writeback	Metallic	Risk of FeFET entering minor loop
	Insulating	No impact

is weakly connected to 0, resulting in a scenario similar to that of baseline FeMFET.

B. Effect of Endurance on Read Operation

If TS is stuck in the low-resistance metallic state during a read operation, the voltage at n_1 node will be pulled closer to 0 irrespective of whether FeCAP is programmed or erased. This results in reduced read margin (50% reduction) for read performed after both short and long retention. In case of TS in high-resistance state before performing a read, there is no impact on case of both read after short retention/long retention because the TS needs to be in insulating state for the proposed TS-FeMFET operation (Table IV).

C. Effect of Endurance on Retention Operation

TS in metallic state during retention results in the voltage at n_1 node becoming close to 0, thus reducing potential noise coupling issues. TS in insulating state results in no impact on retention because TS-FeMFET needs to be in the insulating state for a weak bleeder operation during retention.

D. Effect of Endurance on Write-Back Operation

TS in insulating state has no impact as the write-back operation necessitates TS to be in insulating state to bring the polarization close to the retention polarization. TS in metallic state results in the voltage at n_1 node going closer to 0, resulting in a risk of FeCAP going into minor loops. This makes it difficult to bring back the retention polarization values post a read operation.

IX. CONCLUSION

This article presents a TS-assisted FeMFET (TS-FeMFET) device to lower the write voltage to $\sim$1.7 V and using capacitive coupling circuit assist to further reduce it to $\sim$1.4 V. Addition of a TS provides a write-bypass mechanism and decouples the read and write paths enabling write-voltage reduction independent of capacitor aspect ratio, improving the device reliability of the MOSFET. A stepped WBL approach for read followed by data-dependent write back to bring the polarization closer to retention polarization is also proposed in this article. The variation of retention polarization coupled with variation of read current margins as a function of cycle count for program and erase is also presented. Design space exploration of TS parameters with Monte Carlo simulations for variability study for efficient write operation is also discussed. The impact of endurance of TS on the different operations is also discussed in detail.

ACKNOWLEDGMENT

The authors would like to thank Prof. Sanjay Banerjee, Chunguang Wang, Sandeep Thirumala, and Saikat Chakraborty for useful technical discussions on FeFET device characteristics and proposed design.

REFERENCES

[1] A. I. Khan, A. Keshavarzi, and S. Datta, "The future of ferroelectric field-effect transistor technology," *Nature Electron.*, vol. 3, no. 10, pp. 588–597, Oct. 2020.
[2] J. Muller, T. S. Boscke, U. Schroder, R. Hoffmann, T. Mikolajick, and L. Frey, "Nanosecond polarization switching and long retention in a novel MFIS-FET based on ferroelectric HfO_2," in *IEEE Electron Device Lett.*, vol. 33, no. 2, pp. 185–187, Feb. 2012.
[3] K. Ni, S. Dutta, and S. Datta, "Ferroelectrics: From memory to computing," in *Proc. 25th Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jan. 2020, pp. 401–406, doi: 10.1109/ASP-DAC47756.2020.9045150.
[4] J. Müller *et al.*, "Ferroelectricity in simple binary ZrO_2 and HfO_2," *Nano Lett.*, vol. 12, no. 8, pp. 4318–4323, Aug. 2012.
[5] K. Lee, J.-H. Bae, S. Kim, J.-H. Lee, B.-G. Park, and D. Kwon, "Ferroelectric-gate field-effect transistor memory with recessed channel," *IEEE Electron Device Lett.*, vol. 41, no. 8, pp. 1201–1204, Aug. 2020.
[6] K. Ni *et al.*, "SoC logic compatible multi-bit FeMFET weight cell for neuromorphic applications," in *IEDM Tech. Dig.*, Dec. 2018, pp. 13.2.1–13.2.4, doi: 10.1109/IEDM.2018.8614496.
[7] M.-H. Yan *et al.*, "BEOL-compatible multiple metal-ferroelectric-metal (m-MFM) FETs designed for low voltage (2.5 V), high density, and excellent reliability," in *IEDM Tech. Dig.*, Dec. 2020, pp. 75–78.
[8] K. Kosuge, "The phase transition in VO_2," *J. Phys. Soc. Jpn.*, vol. 22, no. 2, pp. 551–557, 1967.
[9] T. Sakata, K. Sakata, and I. Nishida, "Study of phase transition in NbO_2," *Phys. Status Solidi*, vol. 20, no. 2, pp. K155–K157, 1967.
[10] W. H. Jeong, J. H. Han, and B. J. Choi, "Effect of ag concentration dispersed in HfO_x thin films on threshold switching," *Nanosc. Res. Lett.*, vol. 15, no. 1, p. 27, Dec. 2020.
[11] Y. Zhou, X. Chen, C. Ko, Z. Yang, C. Mouli, and S. Ramanathan, "Voltage-triggered ultrafast phase transition in vanadium dioxide switches," *IEEE Electron Device Lett.*, vol. 34, no. 2, pp. 220–222, Feb. 2013.
[12] E. Cha, J. Park, J. Woo, D. Lee, A. Prakash, and H. Hwang, "Comprehensive scaling study of NbO_2 insulator-metal-transition selector for cross point array application," *Appl. Phys. Lett.*, vol. 108, no. 15, 2016, Art. no. 153502.
[13] W.-Y. Tsai *et al.*, "Enabling new computation paradigms with HyperFET—An emerging device," *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 2, no. 1, pp. 30–48, Jan. 2016.
[14] S. L. Miller, J. R. Schwank, R. D. Nasby, and M. S. Rodgers, "Modeling ferroelectric capacitor switching with asymmetric nonperiodic input signals and arbitrary initial conditions," *J. Appl. Phys.*, vol. 70, no. 5, pp. 2849–2860, Aug. 1991.
[15] A. K. Saha and S. K. Gupta, "Modeling and comparative analysis of hysteretic ferroelectric and anti-ferroelectric FETs," in *Proc. 76th Device Res. Conf. (DRC)*, Jun. 2018, pp. 1–2, doi: 10.1109/DRC.2018.8442136.
[16] A. K. Saha and K. Suman Datta Sumeet Gupta, "'Negative capacitance' in resistor-ferroelectric and ferroelectric-dielectric networks: Apparent or intrinsic?" *J. Appl. Phys.* vol. 123, no. 10, 2018, Art. no. 105102.
[17] S. K. Thirumala and S. K. Gunta, "Gate leakage in non-volatile ferroelectric transistors: Device-circuit implications," in *Proc. 76th Device Res. Conf. (DRC)*, Jun. 2018, pp. 1–2.
[18] J. S. N. Frougier *et al.*, "Phase-transition-FET exhibiting steep switching slope of 8 mV/decade and 36% enhanced ON current," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2016, pp. 1–2, doi: 10.1109/VLSIT.2016.7573445.