Received 13 June 2025; revised 27 September 2025; accepted 4 October 2025.
Date of publication 7 October 2025; date of current version 17 October 2025.

Digital Object Identifier 10.1109/JXCDC.2025.3618883

Benchmarking of FERAM-Based Memory System by Optimizing Ferroelectric Device Model

MOHAMMAD ADNAAN[1] (Member, IEEE), SAEIDEH ALINEZHAD CHAMAZCOTI[2], EMIL KARIMOV[2], MARIE GARCIA BARDON[2], FRANCKY CATTHOOR[3] (Life Fellow, IEEE), JAN VAN HOUDT[4] (Fellow, IEEE), and AZAD NAEEMI[1] (Senior Member, IEEE)

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA
[2]IMEC, 3001 Leuven, Belgium
[3]National Technical University of Athens, 157 72 Zografou, Greece
[4]IMEC, KU Leuven, 3000 Leuven, Belgium
CORRESPONDING AUTHOR: M. Adnaan (madnaan3@gatech.edu)

This work was supported by IMEC.

ABSTRACT We present a framework for design technology co-optimization (DTCO) of the main memory system with one transistor-one capacitor (1T1C) ferroelectric random access memory (FERAM) as an alternative to dynamic random access memory (DRAM). We start with the ferroelectric capacitor device model and perform array-level memory circuit simulation. Then, we map the circuit-level metrics to system-level simulators to analyze the performance enhancement of using FERAM as a main memory. We demonstrate the performance boost and power savings that can be achieved at the system level by improving individual device characteristics and modifying circuit architecture. We have estimated that on average more than 14% improvement in instruction per cycle and 21% reduction in energy consumption can be achieved by substituting DRAM with FERAM equipped with a ferroelectric capacitor having an optimal polarization switching voltage of 1.5 V.

INDEX TERMS Dynamic random access memory (DRAM), ferroelectric memory, ferroelectric random access memory (FERAM), memory circuit, memory circuit benchmarking, nonvolatile memory.

I. INTRODUCTION

FERROELECTRIC memory devices have emerged as a promising memory technology, garnering significant attention from researchers due to their nonvolatile nature and energy-efficient switching characteristics [1]. Specifically with the discovery of ferroelectricity in hafnium oxide, ferroelectrics have once again become a focal point of research, particularly because of the compatibility with CMOS technology [2], [3], [4]. In contrast to earlier use of ferroelectric random access memory (FERAM) technology that was limited to niche applications [5], hafnium zirconium oxide or $Hf_{1-x}Zr_xO_2$ (HZO)-based ferroelectric devices are now being explored for potential applications in embedded memories as the last level of cache [6], as an alternative to dynamic random access memory (DRAM) [7], in conjunction with DRAM [8], and storage class memory [9].

To materialize the milestone of implementing ferroelectric devices as the next generation memory, it is imperative to pay attention at different levels of abstraction encompassing the understanding of device physics [10], fabrication process formulation [11], development of device models [12], [13], and the exploration of circuit architectures [14], [15]. From a materials perspective, HZO has emerged as a top candidate for ferroelectric material due to its exceptional compatibility with the Si CMOS technology [16], setting it apart from alternatives, such as BTO [17], PZT [18], and BFO [19]. Various kinds of ferroelectric memory technologies, such as ferroelectric tunnel junctions [20], FERAM [21], [22], and ferroelectric field-effect transistors [23], employ distinct approaches to leverage the ferroelectric properties and serve as memory devices. Among all these options, FERAM stands out as the most promising choice due to its high endurance [24], low-voltage operation [25], less variability [26], and structural simplicity.

The suitable position of a particular memory technology in the memory system hierarchy is determined by its read/write

latency, energy consumption, density, and cost. Unlike other emerging memory technologies (such as STT-MRAM [27], PCM [28], and RRAM [29]), one transistor-one capacitor (1T1C) FERAM shares a similar memory cell architecture and circuit layout with DRAM, utilizing the same 1T1C structure. Consequently, FERAM can seamlessly replace DRAM without necessitating any new protocols. So, the global system-level macroarchitecture organization does not need to be modified, which we see as an important advantage. Hence, in this article, we will focus on the design technology co-optimization (DTCO)-level changes and exploration. This includes especially the microarchitecture, circuit- and device-level cross-layer design exploration. In addition, as process technology continues to advance, ferroelectric capacitors show promise in reducing their read/write voltage requirements, enabling low-voltage operation [30]. Moreover, when compared to DRAM, FERAM offers distinct advantages due to its nonvolatile nature, eliminating the need for periodic refreshing. Besides, nonvolatility helps FERAM to be used in computing environment with intermittent power supply in applications, such as edge computing [31], [32]. Furthermore, in DRAM, bit lines are always required to be precharged to $(V_{DD}/2)$ to prevent charge leakage from the capacitor. This precharging operation consumes extra energy, which can be saved by using FERAM. All these properties pertaining to device and circuit levels exert influence on the overall performance of the system. Nevertheless, the existing research has predominantly focused on individual abstraction levels, particularly in the examination of FERAM at the levels of individual devices and circuits [33], [34]. Limited research has been conducted on DTCO of FERAM [34], [35] compared to other competitive technologies [36], [37], [38]. In both [34] and [35], diminished system-level performance has been observed when replacing DRAM with the existing technology of FERAM, primarily attributed to the higher access latency of FERAM. However, analyzing results solely at an isolated abstraction level poses challenges in reaching conclusive insights regarding the potential impacts on other layers. Considering all these factors, it is crucial to evaluate the potential improvements at the system-level performance that could arise from the substitution of DRAM with FERAM, encompassing all abstraction levels from the device to the system.

The main contributions of this article are listed as follows.

1) A comprehensive DTCO of FERAM spanning from the device to the system level has been conducted.
2) A framework to assess the performance of memory system with 1T1C FERAM and compare it with DRAM-based systems has been formulated.
3) Demonstrated the impact of reducing polarization switching voltage; hence, the impact of the read/write voltage on the system-level performance.
4) Optimized the ferroelectric capacitor device model and showcased the voltage range that, when utilized in memory circuits, dissipates less power as compared to DRAM.
5) Highlighted the path of reducing the read/write latency and subsequently enhancing system speed by modifying the memory block architecture.

This article is organized as follows. In Section II, we explain the formulation of device models for ferroelectric capacitors. Then, in Section III, we describe the FERAM circuit architecture, read/write scheme, memory cell layout, and circuit-level simulation results. Finally, in Section IV, we show the work flow of our system-level simulation and mapping of the circuit-level results to the system-level simulators and the system-level results. Section V summarizes the conclusion.

II. BACKGROUND ON FERROELECTRIC CAPACITOR MODEL

The HZO ferroelectric capacitor (9.9 nm thick, 200 × 200 nm footprint) considered for circuit simulations has been modeled with multidomain Preisach-based approach [39] and calibrated with the experimental data from IMEC. The circuit model for each device consists of 100 ferroelectric domains with individual domain switching time of $\tau = 5$ ns. In this model, if the applied electric field $E_{fe} > E^{+}_{c,i}$ (positive coercive electric field of the ith ferroelectric domain) or $E_{fe} < E^{-}_{c,i}$ (negative coercive electric field of the ith ferroelectric domain), then the fractional change of polarization of that individual domain within a duration of time Δt is $\delta P_i = \pm(\Delta t)/\tau e^{-(\Delta t)/\tau}$, where τ is individual domain switching time. Then, the polarization value of ith domain at any time $t + \Delta t$ becomes $P_{i,t} + \delta P_i$. The polarization value of each domain P_i is capped at ± 1. Next, the overall polarization of ferroelectric capacitor at any time t is calculated as $P_t = P_r \times (\Sigma P_{i,t})/N$, where P_r is the remanent polarization and N is the total number of ferroelectric domains. Finally, the displacement current of the ferroelectric capacitor is calculated as follows:

$$I = \frac{d}{dt}\left(P_t + \epsilon_0 \epsilon_{fe} E_{fe}\right) A \tag{1}$$

where A is the area and $\epsilon_{fe} = 35$ is the background dielectric permittivity of the ferroelectric capacitor. A bivariate Cauchy–Lorentz distribution has been considered for coercive electric field of ferroelectric [39], [40]. The mean value and standard deviation of coercive field distribution have been calibrated with experimental data to get the original model shown in Fig. 1(a). The original model is based on experimental data and requires -2.5 V to fully switch polarization to a negative state. It has been shown that there are multiple pathways to reduce the coercive voltage/polarization switching voltage of HZO keeping the remanent polarization same by stress-induced crystallization [41] and thickness scaling [25].Therefore, to analyze system-level performance, we set two other target polarization switching voltages of -2.0 and -1.5 V. To decrease the polarization switching voltage below -2.0 V, we adjust the parameters of the coercive field distribution accordingly, which gives us the modified model 1 (MM1). Next, we further modify those

TABLE 1. Ferroelectric capacitor models.

	Mean value of positive coercive electric field $E_{c+,\mu}$	Mean value of negative coercive electric field $E_{c-,\mu}$	Standard deviation of positive coercive electric field $E_{c+,\sigma}$	Standard deviation of negative coercive electric field $E_{c-,\sigma}$	Polarization switching voltage
Original model	0.7 MV/cm	-1.35 MV/cm	0.35 MV/cm	0.38 MV/cm	2.5V
Modified model 1 (MM1)	0.7 MV/cm	-0.7 MV/cm	0.35 MV/cm	0.35 MV/cm	2.0V
Modified model 2 (MM2)	0.53 MV/cm	-0.53 MV/cm	0.28 MV/cm	0.28 MV/cm	1.5V

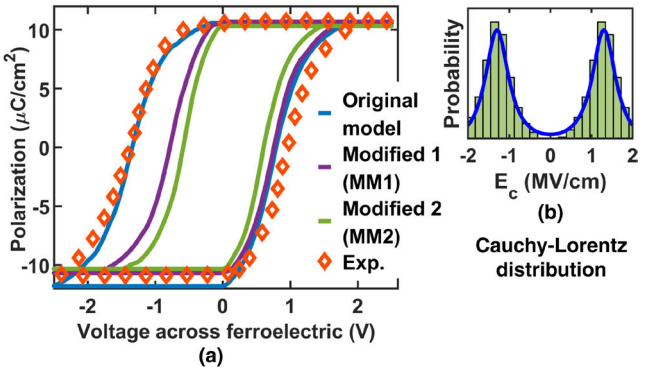

FIGURE 1. (a) Device models used in circuit simulation. (b) Sample distribution of coercive electric field.

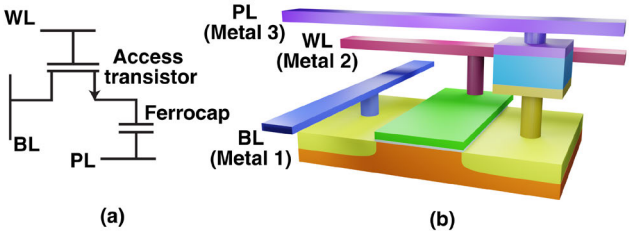

FIGURE 2. (a) Schematic of single memory cell. (b) Layout of memory cell.

parameters to achieve the target polarization switching voltage of −1.5 V, which provides us with the modified model 2 (MM2). The parameters of the coercive field distribution used in three different models are shown in Table 1.

III. ARRAY-LEVEL MEMORY CIRCUIT SIMULATIONS

The structure of a single memory cell of 1T1C FERAM is similar to that of DRAM [Fig. 2(a)]. A conventional DRAM cell consists of a single access transistor and a storage capacitor that together store one bit of information. The word-line (WL) controls the access transistor, which connects the capacitor to the bitline (BL) during read and write operations. To write data, a voltage corresponding to the desired logic state is applied to the BL while the WL is activated, charging or discharging the capacitor. During a read operation, the stored charge is shared with the BL, creating a small voltage deviation that is detected by sense amplifiers to determine the stored bit. Since the read process is destructive, the original charge must be restored immediately after sensing. Moreover, due to charge leakage from the storage capacitor, DRAM cells require periodic refresh cycles to preserve data integrity. In contrast, a FERAM cell also employs an access transistor but replaces the conventional capacitor with a ferroelectric capacitor. The WL activates the access transistor, while the BL and plate line (PL) are used to read or write the polarization state of the ferroelectric capacitor. Owing to the nonvolatile nature of ferroelectric polarization, FERAM does not require periodic refresh operations. The state-of-the-art DRAM circuits are designed with proprietary 28-nm HKMG technology PDK, which are not publicly available. Hence, here we have used the closest available 28-nm TSMC Si-CMOS PDK for our peripheral circuitry. We have drawn the layout, as shown in Fig. 2(b), and extracted the parasitic resistance and capacitance of the memory cells and interconnects.

A. READ/WRITE SCHEME

For writing positive polarization state (“1”) to the memory cell, we apply a positive electric field across the ferroelectric capacitor by raising the BL to write voltage and keeping PL grounded. Alternatively, for writing negative polarization state (“0”), we raise the PL to write voltage and keep the BL grounded. In both cases, the word line is overdriven by 0.3 V above the write voltage to ensure the access transistor has switched on completely. Now, to write a random word of “10110...” to any particular row of the memory block in a single write cycle, we have followed the write scheme, as shown in Fig. 3 [42]. During the second half of the write duration, PL is always raised, and depending on whether we are writing “1” or “0” to that cell, we either make the corresponding BL high or keep it grounded for the full duration. As a result, during the first half of the time period, there is positive electric field across the ferroelectric capacitor while writing “1” and positive polarization is written at that time. Alternatively, while writing “0,” in the second half period, there is negative electric field across the ferroelectric

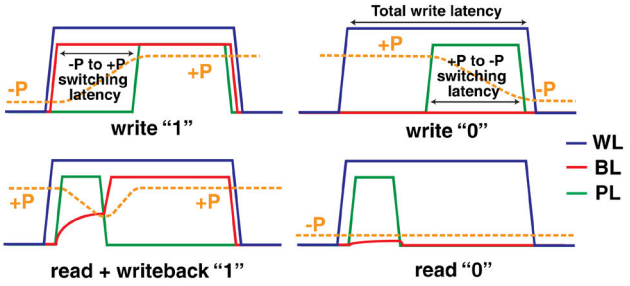

FIGURE 3. **Write/read scheme of memory cells.**

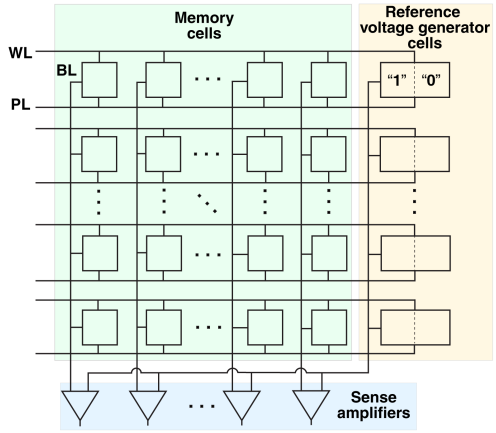

FIGURE 4. **Word line parallel to PL circuit architecture used in simulations.**

capacitor and negative polarization is written at that time. Here, the latency for full switching is considered as the time it takes for the polarization to switch to 90% of its final value from the opposite state after applying the voltage pulse. So, the total write latency of the memory array = switching latency of $-P \rightarrow +P$ + switching latency of $+P \rightarrow -P$, as shown in Fig. 3.

Although there is a recent very promising demonstration of nondestructive read scheme for ferroelectric capacitive memories [43], the technology is still under development. Hence, we follow the traditional destructive process to read the saved data (Fig. 3). In this process, we first precharge the bit lines to zero voltage. Then, we turn on the access transistor and apply a voltage pulse to PL with an intention to switch the saved polarization state. Now, if the saved polarization state was positive, the polarization state begins to flip and it generates a higher voltage on the BL. Alternatively, if the saved polarization state was negative, it remains as it was and a smaller voltage is generated at the BL. We consider the time required to generate a sense margin of 100 mV as the read latency. Finally, with the help of sense amplifiers, we differentiate between the generated BL voltages and read the data. Now, since the reading process is destructive, we need to write back the data if the stored polarization state was positive. So, the overall delay for the read/writeback latency is the summation of the BL precharge delay, the read latency, and the latency to set the cell state to the positive polarization state (write back).

B. CIRCUIT ARCHITECTURE

The circuit architecture for memory array used in the simulation is shown in Fig. 4. In the memory array, the word and PLs are shared among the memory cells in a row. The BL runs perpendicular to the word line and connects to the sense amplifiers. Each row contains a pair of memory cells used as a reference voltage generator, one of which is tied to the positive polarization state (“1”) and another to the negative polarization state (“0”). During the read operation, they will generate a reference voltage, which is at the midlevel of the BL voltages generated by the saved positive and negative polarization states [42]. Without delving into further circuit-level details, we provide here a higher level overview; a comprehensive circuit-level analysis of FERAM can be found in [44]. We consider four cases of circuit simulation scheme, as described in Table 2. To match with the DRAM specifications for page size in JEDEC guideline [45], we first considered 8192 memory cells in every row of the memory bank in Case 1, but due to this large number of memory cells in a row, the parasitic capacitances of the word and PLs become too large, which results in an excessively large read/write latency. To decrease the read/write latency, in Case 2, we divide the memory bank into eight subbanks running in parallel with a minimal increase (∼1.53%, mostly due to additional reference cells and WL, PL drivers) in layout area compared to Case 1. This novel approach keeps the page size the same as Case 1 since the subbanks are used in parallel while reducing the parasitics associated with the word and PLs in each subbank. We can see the drop in latency, as shown Table 2. As will be shown in the next section, though the latency has decreased in Case 2, the energy consumption at the system level is still high compared to DRAM. To reduce the energy consumption, we look at the two modified ferroelectric capacitor models MM1 and MM2 with lower polarization switching voltages in Cases 3 and 4, respectively. Besides, since the polarization switching voltage is 1.5 V in case 4, we have used the access transistor model with a 1.8-V nominal voltage from the TSMC 28-nm process node, whereas in the first three cases, we have used an access transistor model with 2.5 − V nominal voltage. The results from the circuit-level simulations are summarized in Table 2. In Cases 2–4, almost the same polarization dynamics are followed while switching from $-P$ to $+P$ state with gradually lower read/write voltages as evident from Fig. 1(a). As a result, in these cases, a gradual increase in write “1” latency and writeback to “1” latency is observed. Consequently, there is a gradual increment in both “total write latency” and “read + writeback latency,” as shown in Table 2. On the other hand, the read/write energy consumption per bit in Cases 2–4 gradually decreases due to lower read/write voltages.

TABLE 2. Circuit simulation schemes and results.

	Case 1	Case 2	Case 3	Case 4
Ferroelectric model	Original	Original	MM1	MM2
Nominal voltage of access transistor	2.5V	2.5V	2.5V	1.8V
Read/write scheme	2.5V/2.8V	2.5V/2.8V	2.0V/2.3V	1.5V/1.8V
Memory block size	1024 rows × 8192 memory cells	1024 rows × 1024 memory cells × 8 sub-banks	1024 rows × 1024 memory cells × 8 sub-banks	1024 rows × 1024 memory cells × 8 sub-banks
Total write latency	65.78 ns	20.84 ns	21.1 ns	23.3 ns
Read+writeback latency	30.71 ns	18.73 ns	20.72 ns	23.57 ns
Write energy per bit	0.65 pJ	0.65 pJ	0.45 pJ	0.27 pJ
Read+writeback energy per bit	0.51 pJ	0.47 pJ	0.29 pJ	0.21 pJ

TABLE 3. System specification.

CPU clock frequency = 3.2 GHz Single core x86, 4-wide instruction issue CPU L3 cache size 1MB, 8 way set associative Cache line = 64 byte	
Memory clock frequency = 1.2 GHz Memory organization = 1 channel, 1 rank, 4 bank groups, 4 banks Memory access protocol = DDR4-2400	
System 1	**System 2**
DRAM = 16MBx8	FERAM = 16MBx8

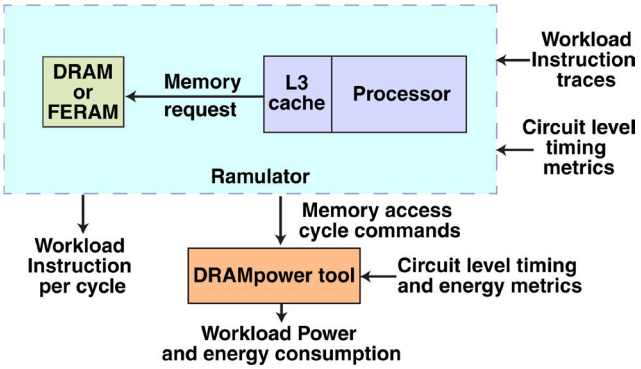

FIGURE 5. System-level benchmarking workflow.

IV. WORKLOAD-DEPENDENT DTCO ANALYSIS

A. EXPERIMENTAL SETUP

After performing circuit-level simulations, we consider two representative systems shown in Table 3 to compare the performance at the system level. Keeping all other specifications the same, the DRAM of System 1 is replaced by FERAM in System 2. Each 16-MB chip comprises four bank groups with each containing four memory banks. In systems incorporating either DRAM or Case 1 of FERAM, each memory bank has 1024 rows, with each row containing 8192 memory cells. For Cases 2–4 with FERAM, every bank is subdivided into eight subbanks, each comprising 1024 rows and 1024 memory cells per row. Notably, in Cases 2–4, there is no need for separate addressing of individual subbank due to their parallel access. We consider a sufficiently long CPU instruction trace produced by merging multiple traces from SPEC 2006 [46] to evaluate the system-level performance. To evaluate the execution time of the instruction trace, we use Ramulator, which is a standard DRAM simulator [47]. It emulates the behavior of a processor and can generate memory requests to be sent to the DRAM when a CPU instruction trace is given. We can obtain the number of instructions executed per CPU cycle (IPC) as an output from Ramulator. A faster system is indicated by higher IPC. The Ramulator also has a memory controller unit, which can generate internal memory access cycle commands. After extracting these memory access cycle commands from Ramulator, we use them in DRAMpower tool [48] to estimate the energy and power consumed by the memory to execute the CPU instruction. The simulation flow for system-level evaluation is shown in Fig. 5.

B. METHODOLOGY

We follow the DDR4 protocol [45] and map the circuit-level timing and energy metrics to the system-level interface for all the four cases with FERAM (Tables 4 and 5). The values for DRAM-based system are taken from [49]. Here, N_{RCD} is the delay between the activate command and internal read/write command in the memory chip. While reading, charge sharing to bit lines happens within N_{RCD} clock cycles, and after that, internal read command is issued [50]. The internal read command starts to send the data toward output I/O buffers. In DRAM, a separate precharge command (N_{RP} duration of precharge command) is issued after every read/write command to precharge the BLs to $(V_{\mathrm{DD}}/2)$ to minimize charge leakage. Whereas FERAM being nonvolatile, we do not need this precharge operation, but in FERAM, we need to precharge the BLs to zero voltage before any read operation. We take into account the delay and energy consumption due to this precharging within the read cycle in the case of FERAM. To ignore the precharge command issued after read/write operation and make its energy consumption zero, we put $N_{\mathrm{RP}} = 0$. Hence, in our framework for FERAM, $N_{\mathrm{RCD}} = f_{\mathrm{CLK}} \times$ (BL precharge delay + read

TABLE 4. Parameters for Ramulator.

	DRAM	Case 1	Case 2	Case 3	Case 4
N_{RCD}	16	19	10	10	10
N_{RAS}	32	39	30	30	30
N_{WR}	18	79	26	26	28
N_{RFC}	192	1	1	1	1
N_{REFI}	9360	infinity	infinity	infinity	infinity

* The values represent number of clock cycles for memory clock frequency 1.2GHz

TABLE 5. Parameters for DRAMpower tool.

	DRAM	Case 1	Case 2	Case 3	Case 4
IDD4R (mA)	145	88	98	84	82
IDD4W (mA)	123	63	109	99	83
IDD51 (mA)	518	0	0	0	0
IDD61 (mA)	24	0	0	0	0
VDD1 (V)	1.2	2.5	2.5	2.0	1.5
VDD2 (V)	2.5	2.8	2.8	2.3	1.8

latency), where f_{CLK} is the memory clock frequency. The total read + writeback command duration is $N_{\text{RAS}} = \text{Max}(N_{\text{RCD}} + f_{\text{CLK}} \times \text{writeback latency}, N_{\text{RCD}} + N_{\text{CL}} + N_{\text{BL}})$. Here, N_{CL} is the latency to arrive the data at output buffer and N_{BL} is the duration of data burst. We keep the values of N_{CL} and N_{BL} the same as those in the DRAM technology.

While writing, an internal write command is issued after N_{RCD} clock cycles of activate command and data starts to be written [50]. So, total write cycle duration is $N_{\text{RCD}} + N_{\text{WR}}$, where N_{WR} is the write recovery time in clock cycles. We calculate the average read current IDD4R and write current IDD4W to be used by DRAMpower tool using the following equations:

$$\text{Write power per bit} = f_{\text{CLK}} \cdot \frac{\text{Write energy per bit}}{N_{\text{RCD}} + N_{\text{WR}}} \tag{2}$$

$$\text{Read + writeback power per bit} = f_{\text{CLK}} \cdot \frac{\text{Read + writeback energy per bit}}{N_{\text{RAS}}} \tag{3}$$

$$\text{IDD4W} = \text{Number of bits written in a cycle} \times \left(\frac{\text{Write power per bit} - \text{WL power per bit}}{\text{VDD1}} + \frac{\text{WL power per bit}}{\text{VDD2}} \right) + \text{IDD3N} \tag{4}$$

$$\text{IDD4R} = \text{Number of bits read in a cycle} \times \left(\frac{\text{Read + writeback power per bit} - \text{WL power per bit}}{\text{VDD1}} + \frac{\text{WL power per bit}}{\text{VDD2}} \right) + \text{IDD3N}. \tag{5}$$

Here, IDD3N is the measured background current that is consumed by the memory chip when clock enable signal of DRAM chip is high and any memory bank is in the active mode. Since it is difficult to estimate this current without measuring the actual hardware memory chip, in our simulations, we kept IDD3N = 37 mA the same as the DRAM technology in all cases [51]. The IDD51 and IDD61 currents are related to refresh currents. For FERAM, we put them to zero because FERAM being nonvolatile does not require refresh operations. For the same reason, in FERAM cases, we put N_{RFC} to a very low value, which is the duration of refresh command, and N_{REFI} to a very high value, which is the interval between successive refresh commands. Following the above novel procedure, we can map the circuit-level metrics to system-level simulators, which can be applicable for other emerging memory technologies also.

C. RESULTS AND DISCUSSION

The results obtained from the FERAM-based system-level simulation are normalized with respect to that of DRAM

TABLE 6. Summery of system-level results with FERAM compared to DRAM-based system.

	Case 1	Case 2	Case 3	Case 4
Instruction per cycle	6.6% lower	14.79% higher	14.79% higher	14.23% higher
Energy consumption	57.14% higher	36.14% higher	5.93% higher	21.28% lower
Power consumption	46.77% higher	56.58% higher	21.59% higher	10.08% lower
Energy-delay product	68.24% higher	18.83% higher	7.72% lower	31.09% lower

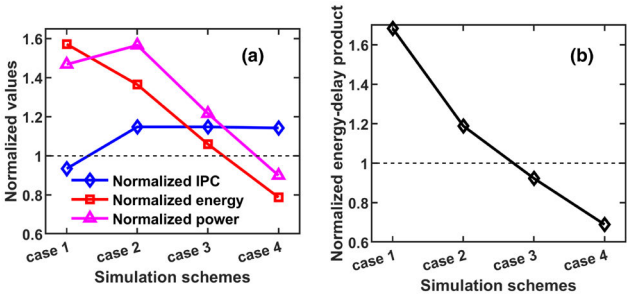

FIGURE 6. System-level simulation results with FERAM normalized with respect to DRAM. (a) IPC, energy, and power. (b) Energy-delay product.

and presented in Fig. 6 and Table 6. For Case 1, we see the performance of the FERAM-based system has a lower IPC and a higher energy/power consumption than the DRAM system. The reason behind a lower IPC is the larger parasitics in the circuit causing slower read/write speed in FERAM memory arrays. Besides, due to higher parasitics of the interconnects, raising the worldline, PL, and BL voltages to V_{DD} requires more energy. When we divide our memory block into eight subbanks (cases 2–4), it reduces the parasitics and read/write latency of the memory array, thereby increasing IPC. Since in Cases 2–4 the mapped values of read/write latency are almost the same (Table 4), the IPC in all these cases is also similar. We observed around 14% increase in IPC in all these cases. On the other hand, the energy consumption has decreased as we move from Cases 2–4. In Case 2, the latency of the read/write cycle has decreased significantly. As a result, though the energy consumption is less, we observe an increase in power consumption. In Cases 3 and 4, the reduction in energy and power is mainly due to the reduction in the polarization switching voltage; hence, lower read/write voltages. Finally, in Case 4, when we use our modified ferroelectric capacitor model (MM2) with a polarization switching voltage of 1.5 V, we have observed a 21.3% reduction in energy consumption, 10.1% decrease in power consumption, and 31.1% decrease in energy-delay product compared to DRAM-based system. We believe that these results substantiate our main claims in a quantitative way. So, it shows that the novel DTCO methodology enables clear power and performance gains of using 1T1C FERAM as a main memory.

V. CONCLUSION

In this article, we have devised a structured route from device to system level to analyze the performance of a memory system with FERAM replacing the DRAM. We have demonstrated that if the 2.5 V polarization switching voltage of the ferroelectric capacitor can be lowered to 1.5 V, we can reduce the energy consumption of the memory system by 21.3% compared to a DRAM one. On the other hand, the speed of the instruction execution in FERAM-based system can be increased by more than 14% by reducing the parasitics through dividing the individual memory block into eight parallelly accessible subbanks with minimal increase in layout area and complexity of peripheral circuitry. The proposed framework can also be employed for estimating the anticipated advantage from other emerging memory device technology-based systems.

REFERENCES

[1] X. Yin et al., "Exploiting ferroelectric FETs for low-power non-volatile logic-in-memory circuits," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2016, pp. 1–8.

[2] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder, and U. Böttger, "Ferroelectricity in hafnium oxide thin films," *Appl. Phys. Lett.*, vol. 99, no. 10, Sep. 2011.

[3] S. Mueller, S. R. Summerfelt, J. Müller, U. Schroeder, and T. Mikolajick, "Ten-nanometer ferroelectric Si:HfO_2 films for next-generation FRAM capacitors," *IEEE Electron Device Lett.*, vol. 33, no. 9, pp. 1300–1302, Sep. 2012.

[4] S. Riedel, P. Polakowski, and J. Müller, "A thermally robust and thickness independent ferroelectric phase in laminated hafnium zirconium oxide," *AIP Adv.*, vol. 6, no. 9, Sep. 2016, Art. no. 095123.

[5] G. R. Fox, R. Bailey, W. B. Kraus, F. Chu, S. Sun, and T. Davenport, "The current status of FeRAM," in *Ferroelectric Random Access Memories: Fundamentals and Applications*. Cham, Switzerland: Springer, 2004, pp. 139–148.

[6] S.-C. Chang et al., "Anti-ferroelectric $Hf_xZr_{1-x}O_2$ capacitors for high-density 3-D embedded-DRAM," in *IEDM Tech. Dig.*, Dec. 2020, pp. 28.1.1–28.1.4.

[7] R. Khosla and S. K. Sharma, "Integration of ferroelectric materials: An ultimate solution for next-generation computing and storage devices," *ACS Appl. Electron. Mater.*, vol. 3, no. 7, pp. 2862–2897, Jul. 2021.

[8] Y. Peng et al., "$HfZrO_x$ hybrid DRAM/FRAM arrays featuring excellent endurance and low latency," *IEEE Electron Device Lett.*, vol. 43, no. 12, pp. 2101–2104, Dec. 2022.

[9] K. Florent et al., "Vertical ferroelectric HfO_2 FET based on 3-D NAND architecture: Towards dense low-power memory," in *IEDM Tech. Dig.*, Dec. 2018, pp. 2.5.1–2.5.4.

[10] A. K. Saha, M. Si, K. Ni, S. Datta, P. D. Ye, and S. K. Gupta, "Ferroelectric thickness dependent domain interactions in FEFETs for memory and logic: A phase-field model based analysis," in *IEDM Tech. Dig.*, Dec. 2020, pp. 4.3.1–4.3.4.

[11] S. S. Cheema et al., "Enhanced ferroelectricity in ultrathin films grown directly on silicon," *Nature*, vol. 580, no. 7804, pp. 478–482, Apr. 2020.

[12] N. Feng et al., "A dynamic compact model for ferroelectric capacitance," *IEEE Electron Device Lett.*, vol. 43, no. 3, pp. 390–393, Mar. 2022.

[13] C.-T. Tung, G. Pahwa, S. Salahuddin, and C. Hu, "A compact model of polycrystalline ferroelectric capacitor," *IEEE Trans. Electron Devices*, vol. 68, no. 10, pp. 5311–5314, Oct. 2021.

[14] W. Tang et al., "FeFET-based logic-in-memory supporting SA-free write-back and fully dynamic access with reduced bitline charging activity and recycled bitline charge," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 70, no. 6, pp. 2398–2411, Jun. 2023.

[15] C. Wang et al., "FeFET-based synaptic cross-bar arrays for deep neural networks: Impact of ferroelectric thickness on device-circuit non-idealities and system accuracy," in *Proc. Device Res. Conf. (DRC)*, Jun. 2023, pp. 1–2.

[16] T. Francois et al., "Demonstration of BEOL-compatible ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ scaled FeRAM co-integrated with 130nm CMOS for embedded NVM applications," in *IEDM Tech. Dig.*, Dec. 2019, pp. 15.7.1–15.7.4.

[17] Y. S. Kim et al., "Critical thickness of ultrathin ferroelectric $BaTiO_3$ films," *Appl. Phys. Lett.*, vol. 86, no. 10, Mar. 2005, Art. no. 102907.

[18] R. Moazzami, C. Hu, and W. H. Shepherd, "Electrical characteristics of ferroelectric PZT thin films for DRAM applications," *IEEE Trans. Electron Devices*, vol. 39, no. 9, pp. 2044–2049, Sep. 1992.

[19] E. Parsonnet et al., "Toward intrinsic ferroelectric switching in multiferroic $BiFeO_3$," *Phys. Rev. Lett.*, vol. 125, no. 6, Aug. 2020, Art. no. 067601.

[20] V. Garcia and M. Bibes, "Ferroelectric tunnel junctions for information storage and processing," *Nature Commun.*, vol. 5, no. 1, p. 4289, Jul. 2014.

[21] J. Okuno et al., "SoC compatible 1T1C FeRAM memory array based on ferroelectric $Hf_{0.5}Zr_{0.5}O_2$," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2020, pp. 1–2.

[22] Y.-C. Luo, J. Hur, Z. Wang, W. Shim, A. I. Khan, and S. Yu, "A technology path for scaling embedded FeRAM to 28 nm and beyond with 2T1C structure," *IEEE Trans. Electron Devices*, vol. 69, no. 1, pp. 109–114, Jan. 2022.

[23] Y. Long et al., "A ferroelectric FET-based processing-in-memory architecture for DNN acceleration," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 5, pp. 113–122, 2019.

[24] J. Okuno et al., "High-endurance and low-voltage operation of 1T1C FeRAM arrays for nonvolatile memory application," in *Proc. IEEE Int. Memory Workshop (IMW)*, Dresden, Germany, May 2021, pp. 1–3.

[25] K. Tahara et al., "Strategy toward HZO BEOL-FeRAM with low-voltage operation (= 1.2 V), low process temperature, and high endurance by thickness scaling," in *Proc. Symp. VLSI Technol.*, Jun. 2021, pp. 1–2.

[26] T. Mikolajick, U. Schroeder, and S. Slesazeck, "The past, the present, and the future of ferroelectric memories," *IEEE Trans. Electron Devices*, vol. 67, no. 4, pp. 1434–1443, Apr. 2020.

[27] Y. Huai et al., "Spin-transfer torque MRAM (STT-MRAM): Challenges and prospects," *AAPPS Bull.*, vol. 18, no. 6, pp. 33–40, 2008.

[28] H.-S. P. Wong et al., "Phase change memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, Dec. 2010.

[29] M. Mao, Y. Cao, S. Yu, and C. Chakrabarti, "Optimizing latency, energy, and reliability of 1T1R ReRAM through cross-layer techniques," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 3, pp. 352–363, Sep. 2016.

[30] K. Toprasertpong et al., "Low operating voltage, improved breakdown tolerance, and high endurance in $Hf_{0.5}Zr_{0.5}O_2$ ferroelectric capacitors achieved by thickness scaling down to 4 nm for embedded ferroelectric memory," *ACS Appl. Mater. Interface*, vol. 14, no. 45, pp. 51137–51148, 2022.

[31] Y. Liu et al., "A 130-nm ferroelectric nonvolatile system-on-chip with direct peripheral restore architecture for transient computing system," *IEEE J. Solid-State Circuits*, vol. 54, no. 3, pp. 885–895, Mar. 2019.

[32] A. Keshavarzi, K. Ni, W. Van Den Hoek, S. Datta, and A. Raychowdhury, "FerroElectronics for edge intelligence," *IEEE Micro*, vol. 40, no. 6, pp. 33–48, Nov. 2020.

[33] S.-C. Chang et al., "FeRAM using anti-ferroelectric capacitors for high-speed and high-density embedded memory," in *IEDM Tech. Dig.*, Dec. 2021, pp. 33.2.1–33.2.4.

[34] H. Toyoshima et al., "FeRAM device and circuit technologies fully compatible with advanced CMOS," in *Proc. IEEE Custom Integr. Circuits Conf.*, May 2001, pp. 171–178.

[35] S. Baek, J. Choi, D. Lee, and S. H. Noh, "Energy-efficient and high-performance software architecture for storage class memory," *ACM Trans. Embedded Comput. Syst.*, vol. 12, no. 3, pp. 1–22, Mar. 2013.

[36] S. A. Chamazcoti et al., "Exploring Pareto-optimal hybrid main memory configurations using different emerging memories," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 70, no. 2, pp. 733–746, Feb. 2023.

[37] M. K. Qureshi, V. Srinivasan, and J. A. Rivers, "Scalable high performance main memory system using phase-change memory technology," in *Proc. 36th Annu. Int. Symp. Comput. Archit.*, Jun. 2009, pp. 24–33.

[38] O. Zilberberg, S. Weiss, and S. Toledo, "Phase-change memory: An architectural perspective," *ACM Comput. Surveys*, vol. 45, no. 3, pp. 1–33, Jun. 2013.

[39] Y. Xiang et al., "Compact modeling of multidomain ferroelectric FETs: Charge trapping, channel percolation, and nucleation-growth domain dynamics," *IEEE Trans. Electron Devices*, vol. 68, no. 4, pp. 2107–2115, Apr. 2021.

[40] M. N. K. Alam et al., "HfZrO ferroelectric characterization and parameterization of response to arbitrary excitation waveform," in *Proc. IEEE SOI-3D-Subthreshold Microelectron. Technol. Unified Conf. (S3S)*, Oct. 2019, pp. 1–3.

[41] S. J. Kim et al., "Low-voltage operation and high endurance of 5-nm ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ capacitors," *Appl. Phys. Lett.*, vol. 113, no. 18, 2018, Art. no. 182903.

[42] A. Sheikholeslami and P. G. Gulak, "A survey of circuit innovations in ferroelectric random-access memories," *Proc. IEEE*, vol. 88, no. 5, pp. 667–689, May 2000.

[43] S. Mukherjee et al., "Capacitive memory window with non-destructive read in ferroelectric capacitors," *IEEE Electron Device Lett.*, vol. 44, no. 7, pp. 1092–1095, Jul. 2023.

[44] M. Adnaan, S.-C. Chang, H. Li, Y.-C. Liao, I. A. Young, and A. Naeemi, "Design considerations for sub-1-V 1T1C FeRAM memory circuits," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 10, pp. 107–115, 2024, doi: 10.1109/JXCDC.2024.3488578.

[45] *JEDEC Standard*. Accessed: Dec. 20, 2023. [Online]. Available: https://www.jedec.org/standards-documents/docs/jesd79-4a

[46] (2006). *Standard Performance Evaluation Corp. SPEC CPU 2006*. [Online]. Available: https://www.spec.org/cpu2006

[47] Y. Kim, W. Yang, and O. Mutlu, "Ramulator: A fast and extensible DRAM simulator," *IEEE Comput. Archit. Lett.*, vol. 15, no. 1, pp. 45–49, Jan. 2016.

[48] K. Chandrasekar, C. Weis, Y. Li, B. Akesson, N. Wehn, and K. Goossens. (2012). *DRAMPower: Open-Source DRAM Power & Energy Estimation Tool*. [Online]. Available: http://www.drampower.info

[49] *Micron DRAM Datasheet*. Accessed: Dec. 20, 2023. [Online]. Available: https://www.micron.com/products/dram/ddr4-sdram

[50] D. Lee, Y. Kim, V. Seshadri, J. Liu, L. Subramanian, and O. Mutlu, "Tiered-latency DRAM: A low latency and low cost DRAM architecture," in *Proc. IEEE 19th Int. Symp. High Perform. Comput. Archit. (HPCA)*, Feb. 2013, pp. 615–626.

[51] *Micron Technical Note*. Accessed: Dec. 20, 2023. [Online]. Available: https://www.micron.com/-/media/client/global/documents/products/technical-note/dram/tn4007_ddr4_power_calculation