Exploiting Ferroelectric FETs for Low-Power Non-Volatile Logic-in-Memory Circuits

Xunzhao Yin[1], Ahmedullah Aziz[2], Joseph Nahas[1], Suman Datta[3],
Sumeet Gupta[2], Michael Niemier[1], Xiaobo Sharon Hu[1]
[1]Department of CSE, University of Notre Dame, Notre Dame, IN 46656, USA
[2]Department of ECE, Pennsylvania State University, University Park, PA 16802, USA
[3]Department of EE, University of Notre Dame, Notre Dame, IN 46556, USA
Email: xyin1@nd.edu, afa5191@psu.edu, {jnahas, sdatta}@nd.edu
skgupta@psu.edu, {mniemier, shu}@nd.edu

ABSTRACT

Numerous research efforts are targeting new devices that could continue performance scaling trends associated with Moore's Law and/or accomplish computational tasks with less energy. One such device is the ferroelectric FET (FeFET), which offers the potential to be scaled beyond the end of the silicon roadmap as predicted by ITRS. Furthermore, the I_{ds} vs. V_{gs} characteristics of FeFETs may allow a device to function as both a switch and a non-volatile storage element. We exploit this FeFET property to enable fine-grained logic-in-memory (LiM). We consider three different circuit design styles for FeFET-based LiM: complementary (differential), dynamic current mode, and dynamic logic. Our designs are compared with existing approaches for LiM (i.e., based on magnetic tunnel junctions (MTJs), CMOS, etc.) that afford the same circuit-level functionality. Assuming similar feature sizes, *non-volatile* FeFET-based LiM circuits are more efficient than functional equivalents based on MTJs when considering metrics such as propagation delay (2.9X, 6.8X) and dyanmic power (3.7X, 2.3X) (for 45 nm, 22 nm technology respectively). Compared to CMOS functional equivalents, FeFET designs still exhibit modest improvements in the aforementioned metrics while also offering non-volatility and reduced device count.

1. INTRODUCTION

Both industry and academia researchers are looking for new devices to continue performance scaling trends associated with Moore's Law. One obvious target is a new device that (i) operates at lower voltages and (ii) mitigates leakage currents that have throttled clock rates, and introduced phenomena such as dark silicon [29, 11]. However, even if improved switches evolve, and traditional core scaling continues unabated, one must also address how to provide each core with its requisite data. Any energy spent in moving/acquiring data is energy that cannot be spent to *process* said data. Recent work [6] suggests that in order for future microprocessors to match traditional Moore's Law performance scaling trends, 576 terabits of data must be moved from registers/memory to logic *every second*. If each operand moves a distance of 1 mm (i.e., over 10% of the die), 58W of a 65W power budget would be allocated to just data transfer. However, if this distance can be reduced by 10X, 90% of a 65W power budget could be devoted to computation. Co-locating processor elements and memory would obviously have a significant, positive impact.

ICCAD'16, November 07–10, 2016, Austin, TX, USA
 ISBN 978-1-4503-4466-1/16/11$15.00
DOI: http://dx.doi.org/10.1145/2966986.2967037

"Near data processing" [5]/processing-in-memory (PIM) prototypes have been heavily pursued since the 1990s (e.g., [16, 10]). While projections suggested that many application classes could benefit from PIM systems, commoditized products did not materialize – a trend due no small part to the economics of manufacturing logic in a DRAM process (or vice versa) [5]. More recently, 3D integration is paving a new path toward realizable PIM systems. As examples, studies suggest that systems such as Micron's hybrid memory cube [24] could reduce execution time and system energy by 15X and 18X respectively for MapReduce [25] while the N3XT project suggests 1000X improvements in energy efficiency for abundant data applications [4].

In contrast to "coarse-grained" efforts (i.e., with logic and memory on separate dies), "fine-grained" logic-in-memory (LiM) structures could also bring data closer to processing elements. Frequently based on emerging technologies, these LiM structures may be broadly captured by two categories: (i) leveraging local, non-volatile (NV) storage to preserve system state, and (ii) integrating NV storage elements with the logic itself. An example of the former includes [17] where ferroelectric capacitors save register state at periodic intervals (e.g., for continued operation given unreliable power supplies). As an example of the latter, [22] integrated magnetic tunnel junctions (MTJs) with MOS transistors to build dynamic current mode logic circuits. The MTJs were used to store data words that might be repeatedly used over the course of a given computation, e.g., for a sum-of-absolute differences (SAD) calculation commonly used in compression, motion detection, etc. (More recently, SAD has also been identified by the SRC/DARPA SONIC center as a *nanofunction*, i.e., a computational kernel that, if realized efficiently, can have significant application-level impact [2].)

Our work is directly in line with the last example. We consider fine-grained local storage elements based on ferroelectric field effect transistors (FeFETs) that are directly integrated with logic switches. Both logic and storage elements are based on the same underlying transistor structure, eliminating the need to integrate MTJs [22], other RRAM [1, 12] elements, etc. with CMOS transistors.

An FeFET is made by integrating a ferroelectric material layer in the gate stack of a MOSFET. It has faced many fabrication challenges since its early introduction (e.g., [34]). However, progress in ferroelectric materials and the potential of offering steep subthreshold slopes (SS) have led to the renaissance of FeFET– e.g., it has recently been shown that FeFETs exhibit negative capacitance which enables step-up voltage conversion of the applied gate bias to the surface potential leading to a switching slope (SS) that is steeper than 60 mV/decade [27]. What is even more interesting is that an FeFET device can be tuned such that it either has or does not have hysteresis when considering a plot of I_{ds} vs. V_{gs} [14]. This opens the door to local, NV storage elements (assuming device hysteresis centering at $V_{gs} = 0$).

In this paper, we propose and study tradeoffs between three different circuit design styles for FeFET-based LiM: (i) complementary (differential), (ii) dynamic current mode, and (iii) dynamic logic. We present designs for a ternary content addressable memory (TCAM/CAM) and a full adder (FA). We also compare our designs to other fine-grained LiM designs at the circuit level. Namely, assuming identical circuit functionality and minimum feature sizes, we evaluate our FeFET-based structures against approaches that leverage CMOS technology and MTJs. Notably, FeFET-based approaches are more efficient than MTJ-based designs when considering metrics such as propagation delay (2.9X, 6.8X) and dynamic power (3.7X, 2.3X) (45, 22/28 nm technology respectively). Compared to CMOS functional equivalents, FeFET designs still exhibit modest improvements in the aforementioned metrics while also offering non-volatility and reduced device count.

The rest of this paper is organized as follows. Sec. 2 presents relevant background and reviews related work for fine-grained LiM structures. In Sec. 3, Sec. 4, and Sec. 5, we describe design efforts with complementary logic, dynamic current mode and dynamic logic styles, respectively. In Sec. 6, we present evaluation results, while Sec. 7 concludes.

2. BACKGROUND

Here, we discuss FeFET devices, simulation models and related work.

2.1 The FeFET device

Devices with hysteretic characteristics offer intriguing possibilities for efficient LiM implementation by virtue of an inherent memory present in the device. Tunability of the hysteresis, both in terms of the width of the hysteretic window and its position on the voltage axis, adds circuit design flexibility and enables circuit/application-driven device optimization. Properly designed FeFETs belong to such a class of devices and is the design target in this work. The hysteresis in the FeFET characteristics stem from the polarization retention in the ferroelectric (FE) layer present in

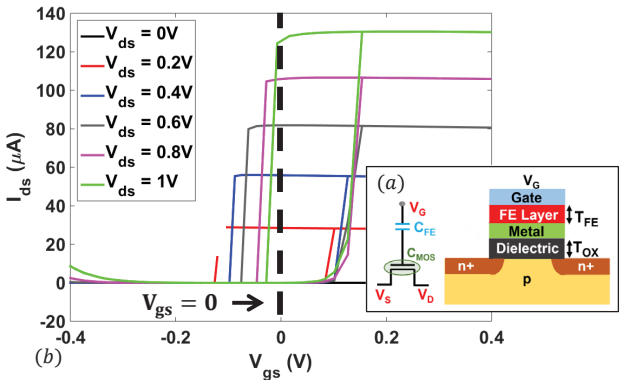

Figure 1. **(a)** FeFET device structure and an equivalent circuit representation showing ferroelectric capacitance (CFE) and the capacitance of the underlying MOSFET (CMOS). **(b)** FeFET I-V curves with tunable hysteresis.

the gate stack (Fig. 1(a)). Since the ferroelectric capacitance (CFE) interacts with the capacitance of the underlying MOSFET (CMOS) in the FeFET structure, the polarizability of the FE layer strongly depends on the relative capacitance CFE/CMOS [27]. For large CFE/CMOS, the FE stabilizes in the negative capacitance region and therefore, is not able to retain the polarization. Negative CFE leads to a voltage step-up action in the FeFET, which yields steep switching behavior (SS $<$ 60mV/decade at room temperature) and higher ON-OFF current ratios compared to the standard MOSFETs [27]. (In the literature, this type of FeFETs are typically referred to as negative capacitance FETs (NCFETs). In this paper, we mainly use FeFETs as we are interested in devices with hysteresis.)

Depending on the non-linearity of CMOS, FeFETs may exhibit non-hysteretic or a mild-hysteretic behavior in this mode of operation. In order to retain the polarization in the FE layer, CFE/CMOS needs to be sufficiently decreased so as to introduce a large hysteresis spanning the positive and negative gate voltages. This can be achieved by proper selection of the FE material and/or device design with *increased FE thickness* (to lower CFE). Moreover, theoretical analyses have revealed that the hysteresis loop in the transfer characteristics of FeFETs can be altered by changing the drain voltage [14] as electrostatic coupling of the channel to the drain changes CFE and CMOS. Thus, the position as well as the width of the hysteresis of FeFETs can be tailored to suit the needs of the circuit. In this work, we utilize this hysteretic feature of FeFETs to design NV, energy efficient LiM circuits.

2.2 FeFET simulation models

For our analysis, we employ a SPICE model for FeFETs based on time-dependent Landau Khalatnikov (LK) equations [28]. The model is calibrated to our experimental data on Hafnium Zirconium Oxide (HZO). The LK equations are self-consistently solved in SPICE with the transistor equations based on 45 nm predictive technology models [30] to obtain the FeFET characteristics. (See Fig. 1(b) for a representative, simulation-based example of a FeFET with tunable hysteresis assuming the FE material parameters are as in Table 1. α, β and γ are the coefficients in the LK equations; the data presented is representative of

HZO material.) Finally, note that in this paper, we restrict our simulations/design space exploration to planar transistor structures (i.e., the 22 nm node above). Integrating FE material with a FinFET would obviously be more challenging as it would likely require the use of cut masks, etc. As such, our simulations do not consider an underlying FinFET device structure at this juncture.

Table 1. FE material parameters

α	-7e9 m/F
β	3.3e10 $m^5/F/coul^2$
γ	-0.2e10 $m^9/F/coul^4$
FE layer thickness	5.7 nm

2.3 Related work: Fine-grained LiM

Here, we briefly review related work with respect to fine-grained LiM. Regarding **CAM designs**, a 4-MOSFET/2-MTJ cell circuit was proposed and fabricated for a NV TCAM [21]. This design requires 60%/86% less area than 12T/16T SRAM-based functional equivalents – essentially due to the fact that the MTJs were placed on top of MOSFETs. A novel logic style based on FE capacitors, i.e., complementary FE capacitors (CFCs), was also considered in this context [15]. The circuits in [15] enable NV storage and logic via CFC capacitive coupling, which reduces device counts for storage and logic blocks. Hybrid circuits based on CMOS, MTJs and FE tunnel junctions (FTJs) have been considered for **fine-grained LiM** due to non-volatility, fast access capability and high, scalable write endurance [23, 9, 31]. NV LiM full adders were designed based on CMOS/MTJ and CMOS/FTJ technology [9, 31, 18, 19], and lower delay, dynamic power, and static power were reported when compared to CMOS functional equivalents.

That said, the devices used to construct NV LiM cells described above have a number of limitations. First, all these devices (CFCs, FTJs and MTJs) are two-terminal devices, which means that additional access transistors are needed to assist the read/write operations. Second, both MTJs and FTJs are used as variable resistors with rather small high/low resistance ratios (between 120%-250%). This implies that (a) the current flowing through the devices is weak with respect to drive capability and still contributes to leakage power, and (b) the devices can only be used in current mode logic (CML)-style circuits (by storing complementary bits and performing a logic function via sensing the differential currents flowing through them.). Third, for CFC-based LiM, a restore operation is required after a logic operation to recover the remnant-polarization charge in the FE capacitor, resulting in extra delay and dynamic power dissipation. In contrast, FeFETs can serve as both a switch and a NV storage element due to the three terminal structure and high I_{on}/I_{off} ratio. As such, these FeFETs could play an unique/flexible role for fine-grained LiM.

3. COMPLEMENTARY LOGIC

We first consider FeFET-based LiM circuit designs based on the "complementary logic" style employed for MTJs [20] and CFCs [15]. Following terminology from [15], "complementary logic" here means that complementary bits are stored in a LiM structure. Note that due to the two terminal nature of MTJs and CFCs, extra writing transistors are required

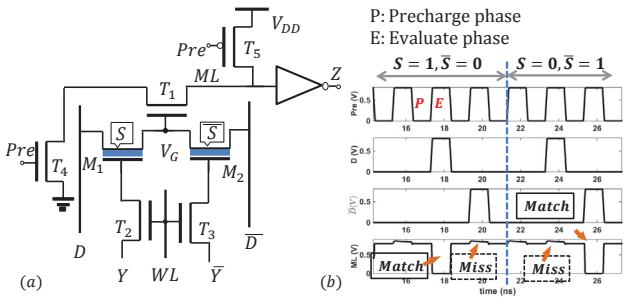

Figure 2. **(a)** FeFET-based core cell for NAND-type CAM and **(b)** its simulation waveforms.

and complementary bits must be stored for proper read and write operations in LiMs based on these devices.

We first consider a NV CAM circuit as in [15, 20]. The CFC-based CAM cell in [15] adopts the complementary logic style and uses 4 MOSFETs (precharge circuitry is not included) in addition to a CFC element (two FE capacitors). Direct drop-in replacement of the CFC pair with two FeFETs does not work since the writing schemes are different. A CFC senses the voltage difference between the two terminals to set the polarization, while a FeFET uses the gate-source bias to set the polarization. We introduce a new FeFET-based complementary logic style CAM cell as depicted in Fig 2(a). The FeFET CAM cell design connects two FeFETs in series, and includes 3 additional transistors (instead of 4 as in [15]). Furthermore, it exhibits better performance due to the fact that the CFCs need one extra operation to recover remnant-polarization charge after each search phase. However, with the FeFET approach, a search can be performed without an extra recovery operation.

The circuit in Fig 2(a) works as follows. The S and $\overline{S}$ bits stored in the FeFETs are written via the access transistors T_2, T_3, controlled by wordline WL as well as the external inputs Y and $\overline{Y}$. The inputs Y (resp., $\overline{Y}$) is set to have either a positive (resp., negative) or negative (resp., positive) gates-source voltage for the FeFETs to change their state to ON or OFF respectively–thus achieving the NV bit storage based on the hysteretic behavior. The circuit compares the stored bits S and $\overline{S}$ with the corresponding search data D and $\overline{D}$ using the two FeFETs and the pass transistor T_1 as comparison transistors. Fig 2(b) shows the waveforms obtained from SPICE simulation. The matchline ML is initially pre-charged. In the case of a match, e.g. $S = 1$ and $D = 1$, the FeFET M_1 is in the ON state and passes the logic '1' on search line D to intermediate node V_G. Given a logic '1', the pass transistor T_1 is turned on and ML is discharged (indicating a match). Similar switching events occur with the complementary FeFET M_2 when $S = 0$ and $D = 0$. In other cases where $S \neq D$ (i.e., a mismatch, labelled as "miss" in Fig. 2(b)), logic '0' is passed through an ON state FeFET. The pass transistor is turned off, and the matchline ML remains at a high level. Given the NAND-like nature of the matchline structure, multiple FeFET-based CAM cells can be serially connected (via pass transistors) to form a word, hence realizing the desired CAM functionality.

Serial connections in a CAM based on the cell design in Fig 2(a) require that the stored bits be complementary,

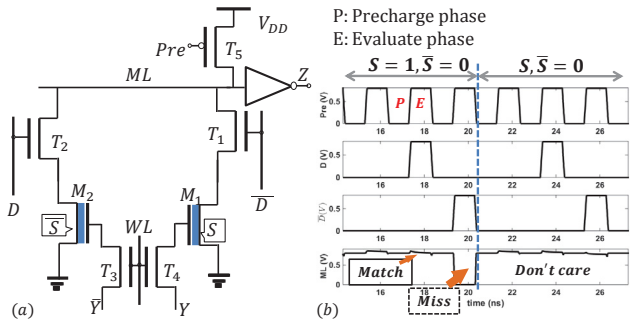

Figure 3. **(a)** FeFET-based core cell for NOR-type TCAM and **(b)** its simulation waveforms.

which limits the CAM functionality to only binary decisions (i.e., match or mismatch). For many applications (e.g., routers for packet forwarding, etc.), it is desirable to also allow an additional '*don't care*' state, which is achieved with a ternary CAM (TCAM). A FeFET-based TCAM can be designed by connecting two FeFETs in parallel and connecting both to the matchline *ML* via two transistors. Fig. 3(a) depicts our FeFET TCAM cell design, where the two FeFETs can both store a '0' in addition to storing complementary bits. The two pairs of comparison transistors, M_1/T_1 and M_2/T_2 serve as pull-down branches from the match line *ML* to ground, and perform an XNOR function, $\overline{S \oplus D}$, at the *ML*. The simulation waveforms of our FeFET-based TCAM circuit are given in Fig. 3(b). In the case of a match, neither pull-down branch conducts, and *ML* stays high. In the case of a mismatch, at least one of the pull-down branches is turned on, and *ML* is discharged. In addition to the above two cases, the '*don't care*' case can be realized if both FeFETs store logic '0'. This case causes a match regardless of the input search data.

Our proposed FeFET TCAM cell bears some similarity with the MTJ-based TCAM introduced in [20]. However, it is important to note that FeFETs cannot be readily used as drop-in replacements for the MTJs in the MTJ TCAM because MTJs and FeFETs have different write schemes as well as search schemes. The MTJ approach senses a voltage difference to set the polarization by two access transistors, while the FeFET approach uses a gate-source bias to set the polarization via two access transistors. Furthermore, an MTJ TCAM cell requires 4 transistor for the search operation and only obtains a poor output swing. Three additional transistors are required to achieve full swing. Thus, 7 transistors in all are required. The FeFET TCAM design – which adopts the search scheme from a CMOS TCAM – requires only 2 transistors to perform a search. Comparing the FeFET TCAM design in Fig 3(a) with the one in [20], one can find that the FeFET TCAM leads to reduced transistor count and incurs no extra area overhead. (More detailed architectural and application level assessments of the FeFET-based CAM/TCAM will be considered in our future work.)

4. DYNAMIC CURRENT MODE LOGIC

In this section, we discuss another circuit design style that is amenable to fine-grained LiM structures, specifically dynamic current mode logic (DyCML). DyCML has been used by many emerging NV devices to realize LiM, e.g., the MTJ [23, 9, 18, 19] and FTJ [31] based work reviewed in Sec. 2. In

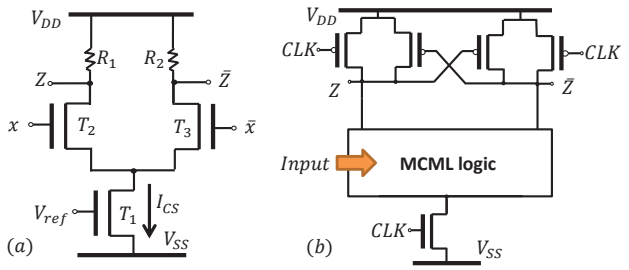

Figure 4. **(a)** Schematic of MOSFET CML inverter; **(b)** general DyCML circuit structure.

this section, we first briefly review the concepts of current mode logic (CML) and dynamic current mode logic (DyCML), and then discuss how to employ FeFETs to realize DyCML for fine-grained, NV LiM.

4.1 Dynamic current mode logic concept

The CML design style aims at reducing dynamic power and delay by limiting the output voltage swing. Unlike "full swing" logic circuits where different transistors switch between ON/OFF states to create a charge or discharge path from the output to the supply rails, with CML all transistors are (at least partially) turned on with a reduced output swing. To determine output, the currents flowing through two branches of the circuit are compared. An MOS CML (MCML) inverter example is shown in Fig 4(a). As can be seen, the inverter operates on both x and $\overline{x}$ with the differential pair circuit T_2 and T_3, and outputs both z and $\overline{z}$. Transistor T_1 serves as a current source. Despite low dynamic power/delay, given present-day concerns about static power, and the fact that static power dissipation can be significant in CML, MCML is not widely employed.

To address the static power concern, DyCML has been proposed [3]. Per the schematic in Fig. 4(b), DyCML circuits employ an MCML block for logic functionality, a 4-transistor latch to precharge and preserve the output before and after logic evaluation respectively, and a dynamic current source that is controlled by the CLK signal to mitigate static power dissipation. This logic family not only utilizes the CML scheme to reduce dynamic power and enhance circuit performance, but simultaneously applies dynamic operation to eliminate the static power dissipation associated with MCML circuits.

DyCML design style has been exploited by many emerging NV devices to realize NV LiM such as MTJs [23, 18, 19, 8, 9] and FTJs [31]. Besides the lower power consumption and higher performance advantages, DyCML's property of using the complementary signals is also desirable for these two-terminal devices. However, since the writing mechanisms of MTJs and FTJs are different from that of FeFETs, drop-in replacements of these devices with FeFETs are not readily achievable.

4.2 FeFET-based CML LiM

FeFETs, when considered as switches, are also suitable to use for DyCML in order to build NV LiM, and could offer additional improvements with respect to performance and energy efficiency. Here proposes a new FeFET-based DyCML LiM circuit structure as illustrated in Fig. 5. It consists

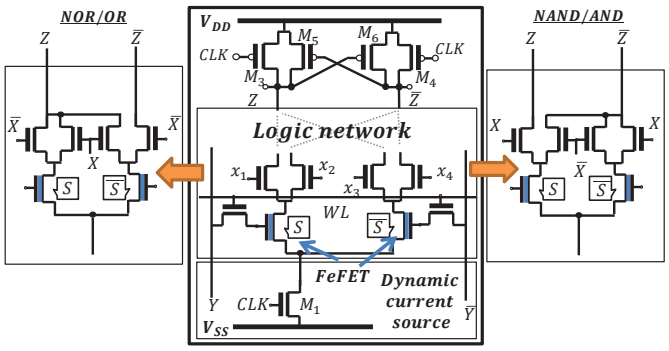

Figure 5. General circuit structure of FeFET-based DyCML LiM circuits. (The inverter for generating $\overline{CLK}$ and the access transistors in the two logic cases are not drawn.)

of four basic parts: **(i)** a clock-controlled pull-up network (top center part) where M_3 and M_4 facilitate pre-charging, and M_5 and M_6 perform latching operations to maintain the circuit output post-evaluation; **(ii)** a logic network that implements the desired logic functionality; **(iii)** a NV storage based on two FeFETs plus two access transistors; and **(iv)** a dynamic current source (lower center part). Depending on the complementary bits stored in the two FeFETs and the implementation of the logic network, the pull-up network generates the corresponding complementary outputs. Our proposed FeFET-based LiM design enables flexible wiring connections within the logic network and employs the FeFETs to perform various logic functions such as NOR/OR and NAND/AND, etc. The insets in Fig. 5 show the NAND/AND and NOR/OR LiM design.

Though our FeFET-based DyCML LiM structure appears similar to the MTJ-based LiM structure in [8, 9], the writing mechanisms of the two structures are very different, which leads to differences in the control transistors to the storage devices. In [8, 9] MTJ-based LiM requires two inverters whose outputs are connected via the two serial MTJs to implement the writing operation, while FeFET-based LiM requires two access transistors to pass a bias voltage to the FeFETs.

Based on the FeFET-based DyCML LiM circuit structure presented above, we have designed a full adder (FA) whose schematic is illustrated in Fig. 6. Fig. 7 shows the simulation waveforms of the FeFET-LiM FA, which demonstrate the correct functionality. This LiM FA requires 4 FeFETs, 28 MOSFETs (24 for DyCML, 4 to facilitate FeFET writes). Note that the 28 MOSFETs may be readily replaced by FeFETs without hysteresis, i.e., NCFETs. The steep SS offered by NCFETs should lead to even lower power consumption. (This will be explored and quantified in future work.) Unlike its CMOS equivalent, the FeFET-based DyCML LiM FA employs FeFETs in the pull-down network as both switches and memory, and obtains non-volatility at the expense of four additional access transistors (which can be reduced to 2 if the FeFETs storing the same bits share an access transistor). Compared to the MTJ-based LiM FA in [9], our proposed design has less devices but much higher performance. More detailed comparisons will be given in Section 6. (In the context of application-level utility, we again refer to the case study for SAD in [22] that begins to

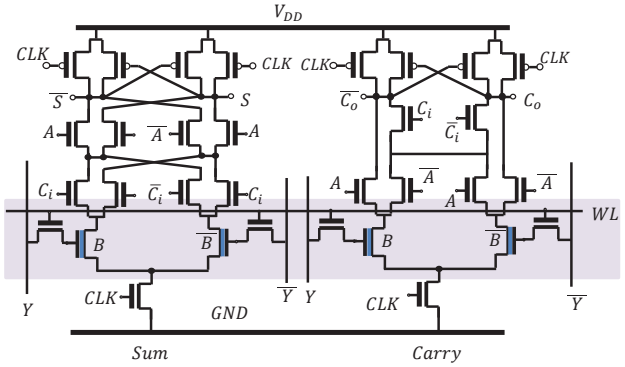

Figure 6. Schematic of FeFET-based DyCML LiM 1-bit FA.

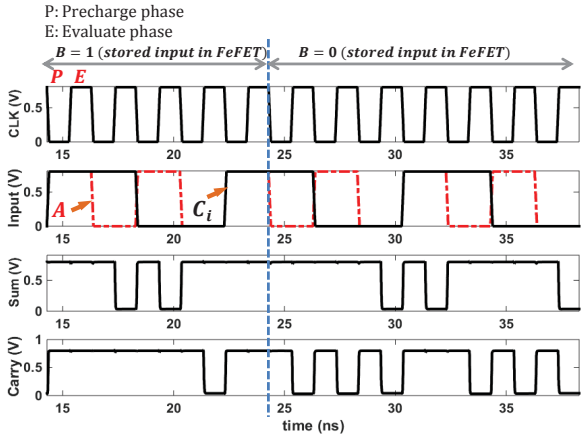

Figure 7. Simulation waveforms of FeFET-based DyCML LiM 1-bit FA.

consider the benefits of non-volatility for specific problems of interest.)

5. DYNAMIC LOGIC

The FeFET-based circuits discussed in Secs. 3 and 4 build off of similar work that employ emerging technologies to realize fine-grained LiM. As will be seen in Sec. 6, these FeFET circuit designs are more efficient than MTJ-based designs (for example) when considering metrics such as propagation delay, dynamic power and device count. Furthermore, delay and dynamic power are similar to CMOS-based designs, while also offering the ability to (non-volatility) store data directly at the gate. However, additional improvements are also possible by leveraging a third logic design style that is uniquely suited to the FeFET. Namely, in this section we consider dynamic logic (DL) style FeFET-based LiM design. FeFETs-based DL LiM is possible because (i) the high I_{on}/I_{off} ratio of FeFETs allow them to be used as switches rather than variable resistors and (ii) unlike MTJs (for example) FeFETs possess a third terminal. FeFET-based DL LiM offers additional improvements with respect to static power and device count.

Dynamic logic gates find utility when improved performance and reduced area are demanded (e.g., in ARM Cortex A8 processors [33]). A dynamic logic circuit consists of a pull-up network which is simply a PMOS transistor with a clocked gate, a NMOS pull-down network that is similar in composition to the ones implemented in CMOS, and a clocked

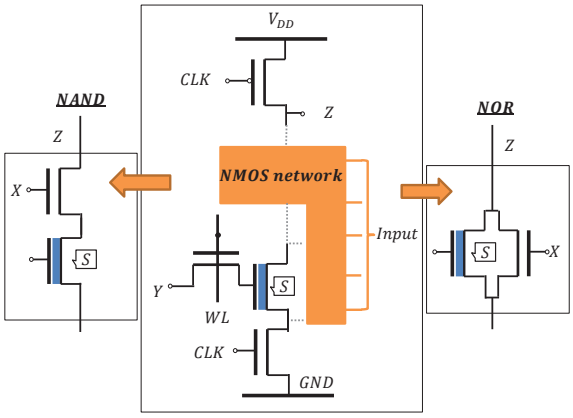

Figure 8. General structure of FeFET-based dynamic logic circuits.

NMOS device that connects the pull-down network and ground [32]. By applying a clock signal, dynamic logic circuits use a sequence of precharge and conditional evaluation phases to realize complex logic functions. Transistor count can essentially be halved, logic delay improves, and no static power dissipation. However, most emerging NV devices (e.g., MTJs and FTJs) cannot leverage the advantages offered by DL since they behave as variable resistors thus always conduct a considerable amount of current even when in a high resistance state. On the contrary, FeFETs can be employed to cut off conducting paths, which makes DL circuits more appealing.

We propose a generic FeFET-based DL LiM circuit structure as shown in Fig. 8. A NAND and a NOR gate are also shown as representative examples. (Note that we assume one of the two inputs is stored locally.) Using conventional DL as context, FeFETs (along with an associated access transistor) can be distributed in the pull-down network (i.e., with other N-type devices) and can serve as both a logic switch and an NV storage element. Specifically, the S bit stored in the FeFET is written via the access transistor, which is controlled by wordline WL as well as the external input Y. The input Y is set to have either a positive or negative gate-source voltage for the FeFET to change its state to ON or OFF respectively thus achieving the NV bit storage based on device hysteresis, albeit at the expense of an access transistor.

Fig. 9 shows the schematic of an FeFET-based DL 1-bit FA. It is similar to a conventional DL FA, but the transistors associated with input B are replaced by the FeFET-based NV memory elements. As the memory elements store the same bit, the access transistor can be shared by the three FeFETs, which reduces transistor count. Fig. 10 shows the simulation waveforms of the FeFET-based DL FA. All the input combinations (with different stored bits) have been tested. Correct functionality of the circuit is observed. (Following the dynamic stage, the circuit employs two inverting static gates at the outputs; this domino gate enables full output swing/cascadability.) Due to reduced transistor count and the DL style employed, as will be seen in Sec. 6, this FeFET-based NV LiM FA achieves better dynamic power efficiency as well as delay than other NV LiM FAs.

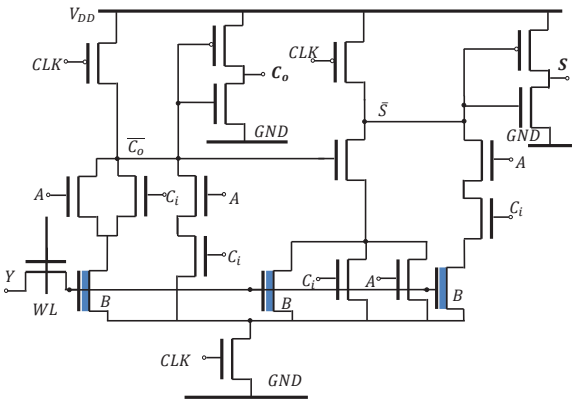

Figure 9. Schematic of FeFET-based dynamic logic 1-bit FA.

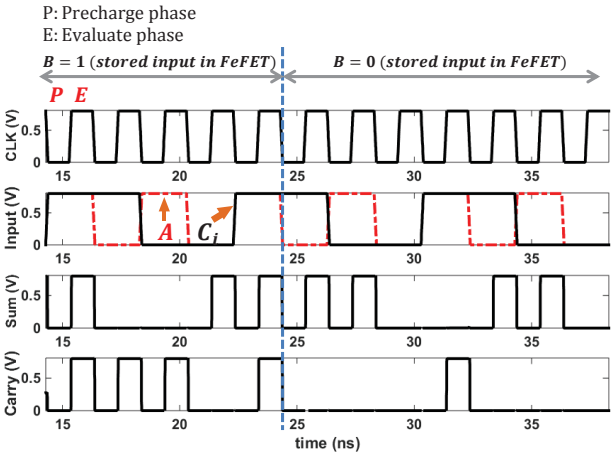

Figure 10. Simulation waveforms of FeFET-based DL LiM 1-bit FA.

6. EVALUATION

In this section, we compare and contrast the different FeFET-based LiM circuits described in the previous sections with LiM circuits based on CMOS and other emerging technologies at similar feature sizes. We specifically consider performance and power of the LiM FA. The metrics considered include supply voltage (V_{DD}), propagation delay (T_d), dynamic power (P_{DYN}), static power dissipation (P_{static}), and whether a design is non-volatile or volatile (NV or V). The simulation results for FeFET-based designs are obtained using HSPICE based on the FeFET device model discussed in Sec. 2.1 and 45nm/22nm ASU predictive technology models (PTMs) [30]. Data for other implementations are directly drawn from the relevant papers.

Table 2 summarizes the data for the different FA designs. Specifically, we examine the FeFET-based DyCML and DL LiM FA designs (rows 2 and 7, and rows 3 and 8, respectively), a conventional CMOS DyCML FA (rows 4 and 9), an MTJ-based LiM FA (rows 5 and 10) and an FTJ-based LiM FA (row 6). Data in rows 2-6 are based on similar technology nodes (40 or 45 nm) while data in rows 7-9 are based on the same technology node (22 nm). In Table 2, we also include data (row 10) from a recently published MTJ-based FA at the 28 nm technology node [8]. This gives a sense of scaling trends as both MTJ and FeFETs are scaled down.

(Note that we do not report 28 nm FeFET data since there is no corresponding PTM.) For the FeFET-based FAs, Table 2 shows the static power values when the FAs are powered on instead of standby power which can be extremely low due to non-volatility.

Let us first examine the data associated with DyCML style FAs (rows 2, 4-6, 7, and 9, 10). Considering comparison with CMOS DyCML FA, from Table 2, one can see that a FeFET-based DyCML FA has similar delay and dynamic power to a conventional CMOS DyCML FA, which is expected since they have similar topologies. However, the FeFET-based DyCML FA exhibits non-volatility with minimal area/transistor overhead. When comparing FeFET NV FA to other NV approaches, improvements over other approaches in the published literature are observed. Notably, comparing FeFET DyCML FA at the 45 nm technology node to MTJ/FTJ designs at the 40 nm technology node, the propagation delay of the FeFET approach is 2.9X/16.6X better, while dynamic power is 3.7X/3.2X better. When devices are scaled, the FeFET DyCML FA is 6.8X better in terms of propagation delay, and 2.3X better in terms of dynamic power than an MTJ-based approach. The device count of FeFET DyCML FA is also smaller than the other approaches. (Static power dissipation is unavailable for the MTJ/FTJ approaches, so no comparison is made across this metric.)

Comparing the FeFET-based DL FA with FeFET-based DyCML FA, one can see from Table 2 that FeFET-based DL FA has much lower transistor count while still achieves comparable performance, power consumption as well as NV functionality. Since CMOS DL FA and FeFET-based DL FA also have similar topologies, they are expected to have similar delay and dynamic power except for non-volatility and thus we skip the presentation of the data for CMOS DL FA.

There are several reasons for the improvements associated with the FeFET-based circuits. First, FeFETs conduct higher currents ($\sim 100\mu A$) while MTJs and FTJs have smaller sensing currents ($\sim 10\mu A$), which leads to FeFETs's improved performance over MTJs and FTJs. Furthermore, FeFETs have a high I_{on}/I_{off} ratio ($\sim 10^6$), while MTJs and FTJs simply serve as a tunable resistor (with just 120% magneto-resistance ratio and 220% tunnel electro-resistance ratios, respectively). This feature enables FeFET-based LiM circuits to have stronger driving capability as well as less dynamic power consumption. Additionally, the FeFET LiM structures require fewer transistors than other NV LiM circuits due to FeFETs being a three-terminal device. Consequently, FeFETs's ON/OFF states are controlled by changing the gate bias via a single access transistor, while MTJ/FTJ-based designs need to monitor tunnel resistance by reversing either the current direction or applied voltage, which requires two or more write transistors. Finally, in the context of dynamic logic, FeFETs can serve as both a storage element *and a switch*, while MTJs (for example) cannot.

7. CONCLUDING DISCUSSION

We exploited the unique hysteretic behavior of FeFET devices to design three different NV LiM circuit structures: complementary, dynamic current mode, and dynamic logic. Simulation results suggest that, for a full-adder benchmark/case study, our FeFET DyCML structure slightly out-performs a (volatile) CMOS functional equivalent in terms of propagation delay and dynamic power. When considering the DL design style, additional improvements in static power and device count are also observed. When comparing our designs to other NV alternative LiM structures based on emerging technologies, our designs consistently perform better across all metrics. Furthermore, all of our projections are likely pessimistic in that we assume MOSFETs in lieu of FeFETs for non-hysteretic devices. From a fabrication perspective, non-hysteretic FeFETs should be compatible with FeFET structures assumed here, and should also offer improved switching slopes, which should lead to additional improvements with respect to delay and power. This will be explored in future work. Additionally, we will consider more application-specific case studies (i.e., for CAM/TCAM designs [13], nanofunctions such as SAD, and circuitry required for critical path operations per [33]).

Acknowledgment

This work was supported in part by the Center for Low Energy Systems Technology (LEAST), one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA.

8. REFERENCES

[1] Logic-in-rram: Design of low-power circuits based on oxide memories. http://lsi.epfl.ch/page-79126-en.html. Accessed: 2016-04-20.

[2] Systems on nanoscale information fabrics (SONIC). https://www.sonic-center.org/research/nano.php. Accessed: 2016-04-20.

[3] M. Allam et al. Dynamic current mode logic (DyCML): a new low-power high-performance logic style. *IEEE JSSC*, 36(3), March 2001.

[4] M. M. S. Aly, et al. Energy-efficient abundant-data computing: The n3xt 1,000x. *Computer*, 48(12):24–33, Dec 2015.

[5] R. Balasubramonian, et al. Near-data processing: Insights from a micro-46 workshop. *IEEE Micro*, 34(4):36–42, July 2014.

[6] S. Borkar et al. The future of microprocessors. *Commun. ACM*, 54(5):67–77, May 2011.

[7] S. Crolles, France. Design Rule Manual for CMOS 40nm, 2012.

[8] E. Deng, et al. High-frequency low-power magnetic full-adder based on magnetic tunnel junction with spin-hall assistance. *Magnetics, IEEE Transactions on*, 51(11):1–4, 2015.

[9] E. Deng, et al. Low power magnetic full-adder based on Spin Transfer Torque MRAM. *IEEE Transaction on Magnetics*, 49(9), September 2013.

[10] J. Draper, et al. The architecture of the diva processing-in-memory chip. In *Proceedings of the 16th International Conference on Supercomputing*, ICS '02, pages 14–25, New York, NY, USA, 2002. ACM.

[11] H. Esmaeilzadeh, et al. Dark silicon and the end of multicore scaling. In *ISCA, 2011 38th Annual International Symposium on*, pages 365–376, June 2011.

Table 2. Performance and Power of Full adders

Row	*Device*	*Technology Node*	*Transistor Count*	V_{DD} (V)	T_d (ps)	P_{DYN} (μW)	P_{static} (nW)	*NV or V*
2	FeFET Fig.6	45nm*	28 MOS+ 4 FeFETs	0.8	30.1	0.53	106.6	NV
3	FeFET Fig.9	45nm*	17 MOS+ 3 FeFETs	0.8	32.9	0.45	60.0	NV
4	CMOS [26]	45nm*	28	0.8	36.3	0.55	97.3	V
5	MTJ [9]	40nm†	34 MOS+ 4 MTJs	1.2	87.4	1.98	N.A.	NV
6	FTJ [31]	40nm†	30 MOS+ 4 FTJs	1.2	500	1.70	N.A.	NV
7	FeFET Fig.6	22nm‡	28 MOS+ 4 FeFETs	0.8	20.2	0.27	41.2	NV
8	FeFET Fig.9	22nm‡	17 MOS+ 3 FeFETs	0.8	18.1	0.21	34.0	NV
9	CMOS [26]	22nm‡	28	0.8	27.0	0.27	36.9	V
10	MTJ [8]	28nm$^\$$	52 MOS+ 4 MTJs	1	137.7	0.62	N.A.	NV

Notes: (i) In SPICE simulations, a 500MHz signal is applied to CLK; **(ii)** $\star$: based on ASU 45nm PTM [30]; **(iii)** †: based on STMicroelectronics 40nm design kit [7]; **(iv)** ‡: based on ASU 22nm PTM [30]; **(v)** \$: based on STMicroelectronics 28nm design kit.

[12] M. Imani, et al. Approximate computing using multiple-access single-charge associative memory. 2016.

[13] M. Imani, et al. Resistive configurable associative memory for approximate computing. In *2016 DATE*, pages 1327–1332. IEEE, 2016.

[14] A. I. Khan. *Negative Capacitance for Ultra-low Power Computing*. PhD thesis, University of California at Berkeley, 2015.

[15] H. Kimura, et al. Complementary ferroelectric-capacitor logic for low-power logic-in-memory VLSI. *IEEE JSSC*, 39(6), June 2004.

[16] C. E. Kozyrakis, et al. Scalable processors in the billion-transistor era: Iram. *Computer*, 30(9):75–78, Sep 1997.

[17] K. Ma, et al. Architecture exploration for ambient energy harvesting nonvolatile processors. In *2015 HPCA*, pages 526–537, Feb 2015.

[18] S. Matsunaga, et al. Fabrication of a non-volatile full adder based on logic-in-memory architecture using magnetic tunnel junctions. *Applied Physics Express*, 1.9(9):091301, August 2008.

[19] S. Matsunaga, et al. MTJ-based nonvolatile logic-in-memory circuit, future prospects and issues. In *DATE*, pages 978–3–9810801–5–5, 2009.

[20] S. Matsunaga, et al. Design of a nine-transistor/two-magnetic-tunnel-junction-cell-based low-energy nonvolatile ternary content-addressable memory. *Japanese J. of Applied Physics*, 51(2S):02BM06, 2012.

[21] S. Matsunaga, et al. A 3.14 um 2 4t-2mtj-cell fully parallel tcam based on nonvolatile logic-in-memory architecture. In *VLSIC, 2012 Symposium on*, pages 44–45. IEEE, 2012.

[22] A. Mochizuki, et al. Tmr-based logic-in-memory circuit for low-power vlsi*. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, E88-A(6):1408–1415, 2005.

[23] A. Mochizuki, et al. TMR-based logic-in-memory circuit for low-power VLSI. *ITC-CSCC*, E88-A(6):1408–1415, June 2005.

[24] J. T. Pawlowski. Hybrid memory cube (hmc). In *Proceedings of HotChips 23*, 2011.

[25] S. H. Pugsley, et al. Ndc: Analyzing the impact of 3d-stacked memory+logic devices on mapreduce workloads. In *2014 ISPASS*, pages 190–200, March 2014.

[26] F. Ren et al. True energy-performance analysis of the MTJ-based logic-in-memory architecture (1-bit full adder). *IEEE TED*, 57(5), May 2010.

[27] S. Salahuddin et al. Use of negative capacitance to provide voltage amplification for low power nanoscale devices. *Nano Letters*, 8(2):405–410, 2008.

[28] T. Song. Landau-khalatnikov simulations for ferroelectric switching in ferroelectric random access memory application. *Journal of the Korean Physical Society*, 46(1):5–9, 2005.

[29] T. N. Theis et al. In quest of the next switch: Prospects for greatly reduced power dissipation in a successor to the silicon field-effect transistor. *Proceedings of the IEEE*, 98(12):2005–2014, Dec 2010.

[30] R. Vattikonda, et al. Modeling and minimization of PMOS NBTI effect for robust nanometer design [Online]http://ptm.asu.edu/. In *DAC*, pages 1047–1052, 2006.

[31] Z. Wang, et al. A physics-based compact model of ferroelectric tunnel junction for memory and logic design. *Journal of Physics D*, 47(045001), December 2013.

[32] L. Wanhammar. *DSP integrated circuits*. Academic press, 1999.

[33] D. Williamson. Arm cortex-a8: A high-performance processor for low-power applications. *Unique Chips and Systems*, page 79, 2007.

[34] S.-Y. Wu. A new ferroelectric memory device, metal-ferroelectric-semiconductor transistor. *IEEE TED*, 21(8):499–504, 1974.