Design and Evaluation of a Near-Sensor Magneto-Electric FET-Based Event Detector

Mehrdad Morsali, *Student Member, IEEE*, Sepehr Tabrizchi, *Graduate Student Member, IEEE*, Andrew Marshall, *Fellow, IEEE*, Arman Roohi, *Senior Member, IEEE*, Durga Misra, *Fellow, IEEE*, and Shaahin Angizi, *Senior Member, IEEE*

***Abstract*— As a recently developed post-CMOS FET, magneto-electric FETs (MEFETs) offer high-speed and low-power design characteristics for logic and memory applications. In this article, a near-sensor processing (NSP) platform leveraging the MEFETs is presented that enables event detection for edge vision sensors at a low cost by eliminating the need for power-hungry analog-to-digital circuits (ADCs). Besides, an efficient background comparison method is presented with adjustable precision that offers the output quality efficiency tradeoff, depending on the application's needs. Our device-to-architecture evaluations show that the proposed hardware–software codesign reduces the energy consumption and execution time on average by a factor of ∼15× and ∼2.4× compared to the SOT-MRAM counterpart running employing the same method.**

***Index Terms*— Magneto-electric FET (MEFET), near-sensor computing, nonvolatile memory.**

I. Introduction

THE CMOS imagers with on-chip feature extraction and compression have been developed extensively in the last decade with the primary goal of optimizing computing resources and reducing overall power consumption [1]. In recent years, emerging nonvolatile memories (NVMs) have attracted significantly increased attention for replacing conventional volatile memory technologies in main memory (i.e., DRAM) or cache (i.e., SRAM) due to their unique features such as nonvolatility, robustness, long endurance, extremely low standby power, suitability for intermittent computing, and integration density [2], [3], [4], [5]. In embedded applications and low-power IoT systems where the on-chip cache is required, a robust NVM can potentially enhance memory capacity and performance. Among conventional NVM technologies, recent experiments and fabrication of spintronics demonstrate the capability of fast magnetization switching (subnanosecond) at magnetic tunnel junction (MTJ) devices utilizing spin-transfer torque (STT) or spin-orbit torque (SOT) switching mechanisms [6], [7], [8]. Notwithstanding their long retention time (10 years) and low write energy (fJ/bit), these technologies suffer from low ON/OFF ratios (typically less than 10), resulting in reliability issues because of the current-driven switching scheme [7]. On the other hand, ReRAM, as another popular NVM technology [9], offers a higher ON/OFF ratio resulting in a larger sense margin but suffers from a slow and power-hungry write operation with a lower endurance than MTJs [4]. Based on the antiferromagnetic magneto-electric (ME) phenomena, magneto-electric FET (MEFET) has been recently introduced and experimentally analyzed [2], [4], [10], [11]. This promising spintronic device offers superior performance and enhanced temperature stability. The unique characteristics of the MEFET device which distinguish it from conventional spintronic devices are its much faster switching speed and very large ON/OFF ratio. As the switching of the ferromagnet or movement of a ferromagnetic (FM) domain wall is not required, a very small switching time (<20 ps) at a low energy cost (<20 aJ) is achievable with a coherent rotation, as the domain switching mechanism [12].

In this work, we propose a near-sensor event-driven architecture using MEFETs, allowing for a tradeoff between power consumption and accuracy at the IoT edge devices. To the best of our knowledge, this work is the first that utilizes MEFET NVMs to store the static background, which leads to a notable reduction in standby power consumption. The main contributions of this article are as follows: 1) we develop a Verilog-A MEFET model by capturing the switching dynamics of the ME layer for more accurate benchmarking at the circuit level; 2) we develop a near-sensor processing (NSP) platform with a MEFET array based on a set of innovative microarchitectural and circuit-level schemes optimized for event detection with energy efficiency and speedup; 3) we complete the design with a flexible hardware-aware event-detection method to detect an event through background variations; and 4) we introduce a

Manuscript received 18 May 2023; revised 11 July 2023; accepted 13 July 2023. Date of publication 27 July 2023; date of current version 23 August 2023. This work was supported by the National Science Foundation under Grant 2228028, Grant 2216772, and Grant 2216773. The review of this article was arranged by Editor L. Liao. *(Corresponding author: Shaahin Angizi.)*

Mehrdad Morsali, Durga Misra, and Shaahin Angizi are with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: mm2772@njit.edu; dmisra@njit.edu; shaahin.angizi@njit.edu).

Sepehr Tabrizchi and Arman Roohi are with the School of Computing, University of Nebraska–Lincoln, Lincoln, NE 68588 USA (e-mail: aroohi@unl.edu; stabrizchi2@huskers.unl.edu).

Andrew Marshall is with the Department of Electrical and Computer Engineering, The University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: andrew.marshall@utdallas.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TED.2023.3296389.

Digital Object Identifier 10.1109/TED.2023.3296389

0018-9383 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.

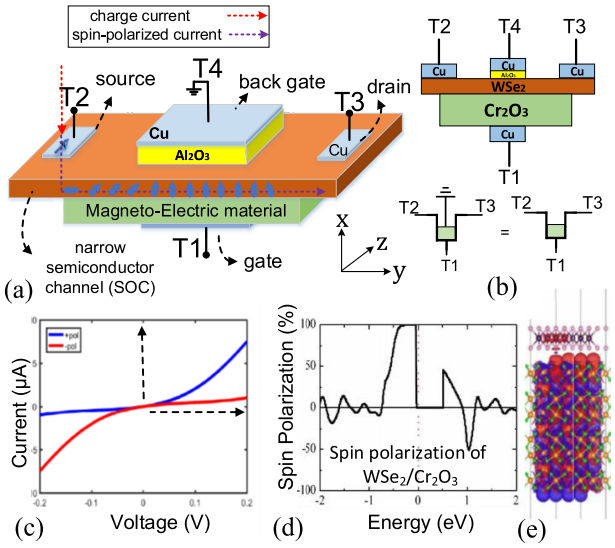

Fig. 1. (a) MEFET device structure, (b) MEFET 2-D view and the circuit scheme, (c) sample source-to-drain current versus voltage at T1, (d) induced spin polarization in WSe_2, and (e) interaction with chromia based on [2].

comprehensive bottom-up evaluation framework to assess the overall performance of the system.

II. Magneto-Electric Field Effect Transistor

A. Device Structure and Operation Principle

The magneto-electric spin field effect transistor (MEFET) shares structural similarities with the CMOS FET device. Fig. 1(a) shows the basic single-source version of MEFET, which is a four-terminal device with gate (at T1), source (T2), drain (T3), and back gate (T4) terminals [2], [13]. This device comprises a narrow semiconductor channel positioned between two dielectrics: the ME material, such as chromia (Cr_2O_3), and the insulator, for example, alumina (Al_2O_3). The narrow semiconductor channel can be constructed from various suitable materials such as PbS, graphene, InP, WSe_2, and others. Two electrodes are attached to the stacked structure. One is located at the bottom gate (T1) via the ME layer, and the other is located at the top (T4) via the alumina layer that forms the back gate. As demonstrated in Fig. 1(b), the channel is tungsten diselenide (WSe_2), providing a high on–off ratio with high hole mobility [2], [13]. As for source and drain material, both conductors and FM polarizers can be employed (at the T2/T3 terminal). The MEFET circuit schematic used in this work is shown in Fig. 1(b). As T4 is grounded, the simplified three-terminal design is used hereafter.

The MEFET operates as a transistor by first biasing the semiconductor channel via T1 and T4 terminals (similar to gate biasing in CMOS) and then applying the current from T2 to T3 (similar to source–drain biasing in CMOS). Applying an extremely low voltage of around ±100 mV [2], [14] across the gate (T1) and back gate terminals (T4: ground) corresponds to the charging of the ME capacitor. Hence, a vertical electric field is created across the gate depending on whether T1 is positively or negatively charged. When an electrical field is applied, it switches paraelectric polarization and anti-FM (AFM) order in the ME insulator layer. As a result, chromia spin vectors are first reoriented. Through exchange interactions and spin-orbit coupling (SOC), the ME layer's high boundary polarization can then polarize carriers' spins in the semiconductor channel, yielding preferred conduction, for example, much lower resistance, along only one axis. Impossible through conventional gate dielectrics, the MEFET's surface magnetization on the channel induces a directionality in the conductance as shown in Fig. 1(c). In the end, the channel spin vector changes to either the "up" or "down" direction. Based on nonequilibrium Green's function (NEGF) transport simulations, we achieved the current versus voltage dependent on the direction of ME polarization [2], [15] in a 20-nm-width 2-D ribbon with a band mass of $0.1m_e$. At 300 K, a conservative exchange splitting the value of 0.1 eV, T3–T2 = 0.1 V, is taken into account. Extremely high spin polarization induced by the chromia layer in WSe_2 channel can be observed in Fig. 1(d), nearly 100% at the top of the valence band for hole conduction. Fig. 1(e) shows the induced polarization interaction with chromia. Once the semiconductor channel is biased, the charge current is injected from the source, resulting in the generation of a spin-polarized current at T3. For sensing the MEFET, as a resistive device with a higher ON/OFF ratio and a higher tunneling magnetoresistance (TMR) compared to previous spin-based paradigms, a simplified and energy-efficient read circuitry is expected. Experimentally, the ON/OFF current ratio for WSe_2 can be extended up to 10^6 similar to that of FeFET [3]. An experimental demonstration revealed the deterministic state switching in the prototype ME antiferromagnet Cr_2O_3 without any magnetic field, operating at CMOS-compatible temperatures of 300–400 K [11].

B. Device Modeling

As shown in Fig. 2, the Verilog-A model consists of three modules to address two key aspects: 1) ME dynamics to control and induce the polarization through proximity effect in the semiconductor channel and 2) source-to-drain spin injection and detection feature [10]. Module 1 (input) is designed to capture the electrical charging of the ME capacitor. The relevant capacitance of the ME layer is modeled through a resistor–capacitor circuit network by $C_{\mathrm{ME}} = ((\varepsilon_{\mathrm{ME}} A)/t_{\mathrm{ME}})$. Here, $\varepsilon_{\mathrm{ME}}$ represents the dielectric constant of the ME layer, which has a thickness of t_{ME}, while A denotes the cross-sectional area. R_{in} represents the load resistance of the driving level at the input. When a voltage difference is applied to the gate electrodes, the capacitor undergoes charging. By considering the gate–source voltage (V_g) as the input and the threshold voltage for chromia state inversion (V_{th} = 0.050 V [16]), the model can initialize the memory and assign the resulting voltage across the drain and source terminals. Table I displays the experimental switching parameters employed for the chromia layer and SOC channel in the presented model. Module 2 captures the switching dynamics associated with the boundary magnetization between the ME layer and the narrow semiconductor channel and the electrical transit time through the channel. The transition from the source to the drain incorporates a delay factor that affects the overall performance.

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.

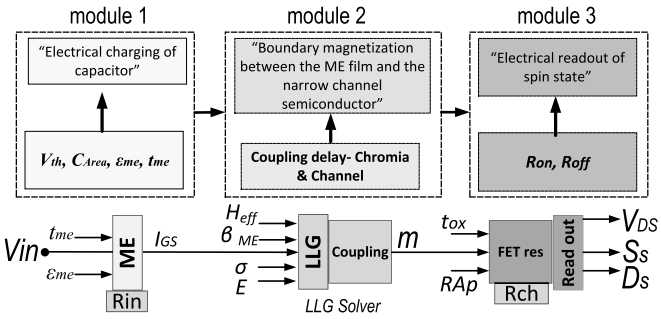

Fig. 2. Verilog-A modules developed for MEFET modeling.

TABLE I
COMPACT MODEL PARAMETERS OF THE MEFET, USED IN THE VERILOG-A MODEL ADAPTED FROM [10], [18]

Parameter	Value	Description of Parameter and Units
ϵ_{ME}	12	Dielectric constant of chromia [19]
$\epsilon_{Al_2O_3}$	10	Dielectric constant of Alumina
t_{ME}	10	thickness of magnetoelectric layer, nm
$W_{ME} \times L_{ME}$	900	area of magnetoelectric layer, nm^2
t_{ox}	2	Oxide barrier thickness, nm
V_{th}	0.05	Threshold of Chromia state inversion, V
V_g	0.1	Voltage applied across ME layer, V
R_{on}	1.05	ON Resistance, $k\Omega$
R_{off}	63.4	OFF Resistance, $M\Omega$

The switching time of the MEFET device is determined by the dynamics of ME. The spin dynamics (m) is modeled by the widely used Landau–Lifshitz–Gilbert (LLG) equation and takes into account thermal fluctuation, electron/spin transport, and the voltage-controlled ME effect [2], [17]

$$\frac{dm}{dt} = -|\gamma| m \times H_{\text{eff}} + \alpha\left(m \times \frac{dm}{dt}\right) + \sigma\beta_{\text{ME}}(m \times E) \quad (1)$$

where γ is the gyromagnetic ratio and H_{eff} is the effective magnetic field. β_{ME} is the ME susceptibility, which is determined by temperature as discussed in [2] and [4]. σ is a scaling factor to help tune the calculated ME-induced momentum magnitude with our experimental data. Module 3 determines the appropriate channel resistance (R_{ch}) and computes associated electrical parameters, including the output voltage at the drain. The 2-D narrow channel's R_{ch} is evaluated in series to input resistance R_{in} to establish the switching boundary conditions. Additionally, the spin states at the source and drain terminals ("Ss" and "Ds") are verified using two spin-state terminals. In Fig. 2, the "Ss" terminal is set to "+1 V" for the "up" spin and "−1 V" for the "down" spin. The drain terminal ("Ds" in Fig. 2) is used to detect the spin current. Our model incorporates a fixed delay of 200 ps to account for the processional delay across the FM layer, estimated based on reliable coupling delay data [12]. It is noteworthy that we scaled up and compared the I–V curve and R_{ON}–R_{OFF} ratio in Fig. 1(c) with the corresponding values reported in Table I.

III. NSP PLATFORM

We propose an always-on event detector architecture based on an NSP scheme. The proposed NSP architecture mainly consists of two sets of arrays as depicted in Fig. 3(a): 1) a standard 128 × 128 pixel array to capture frames and 2) a 128 × 128 MEFET background array. Peripheral circuits include a row controller (Ctrl), a command decoder connected

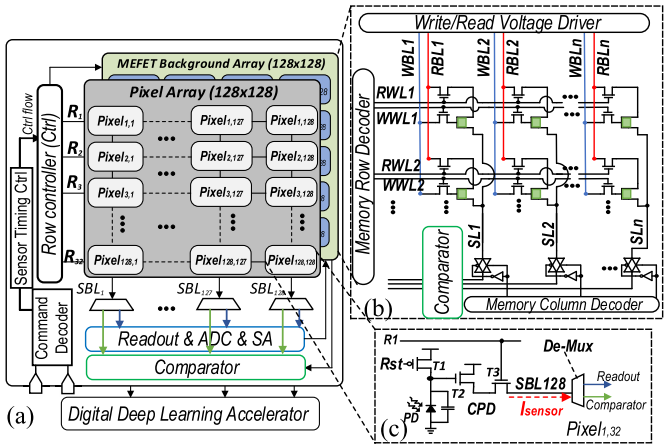

Fig. 3. (a) Proposed NSP platform, (b) MEFET background array, and (c) 3T pixel structure.

to a sensor timing Ctrl, and readout/analog-to-digital circuit (ADC)/sense amplifier (SA)/comparator circuitry. The key idea behind developing the NSP is to leverage the nonvolatile MEFET array as the storage element with a high TMR on the sensor to save some algorithmically selected pixels from input frames as a static background. The NSP then works in event detection mode by comparing the selected pixels with new input in an ADC-less fashion to detect a moving object. To reduce the overall power consumption, the NSP only updates (sends) the modified pixels on the MEFET array. Once the object is detected, the NSP can switch to the sensing mode to capture the frame and send it into any digital deep-learning accelerator to detect it.

A. Pixel Array

Illustrated in Fig. 3, the NSP platform consists of a pixel array where each pixel consists of a typical three-transistor/1-photodiode (PD) [Fig. 3(c)]. In the *sensing mode*, when Rst = "high," the PD connected to T1 turns into inverse polarization, and the readout component will be driven by the source bitline (SBL) and captures a V_1 = VDD voltage. When T1 is deactivated, the PD generates a photo-current corresponding to the external light intensity. This photo-current results in a voltage drop (V_{PD}) at the gate of T2. The new voltage (V_2) will also drive the readout component through SBL. The voltage values before and following the exposure to image light are captured by the capacitors. The difference between these voltages is sensed, amplified, and then converted to digital data through an ADC. It is worth pointing out that each ADC sample when the voltage drops, then it subtracts the pixel reset voltage and converts the output signal. Accordingly, the ADC can skip to the next row of the array. In the proposed *event detection mode*, however, the pixels are disconnected from the readout/ADC module by demultiplexers [Fig. 3(a)], where the generated current (I_{sensor}) passing through T2 and T3 [Fig. 3(c)] will drive a near-sensor comparator module for an efficient operation.

B. MEFET Array

The MEFET array depicted in Fig. 3(b) is a 128 × 128 memory array to store the background data. We used the

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.

TABLE II
BIAS CONFIGURATION OF MEFET BACKGROUND ARRAY

	Operation	WWL	RWL	WBL	RBL	SL
Accessed	Read	0	1	-	I_{read}	-
Unaccessed		0	0	-	-	-
Accessed	Write	1	0	V_{write}/-V_{write}	-	-
Unaccessed	('1'/'0')	0	0	-	-	-
All	Hold	0	0	-	-	-

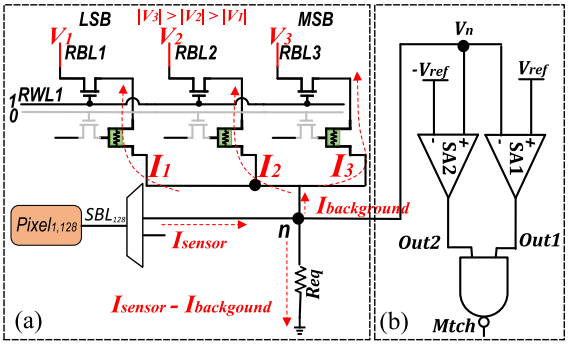

Fig. 4. (a) Near-sensor comparator on top of the KCL calculation. (b) Circuit to generate match output.

2T-1MEFET RAM bit-cell [20] that consists of one MEFET as the main storage element and two access transistors for separate read and write paths. Five controlling signals, namely read word line (RWL), read bitline (RBL), write word line (WWL), write bitline (WBL), and source line (SL) are utilized to control each cell. The read/write access transistor is governed by RWL and WWL, allowing for selective read and write operations on the cells within a specific row. The bias conditions listed in Table II are applied during these operations. In addition to the typical memory components, the MEFET background array contains two additional novel peripheral circuits: A write/read voltage driver and a comparator unit. The write/read voltage driver is responsible for providing the writing voltage during updating the background data and also generating calibrated negative reading voltages (V1–V3) during event detection mode.

C. Near-Sensor Comparison

During the event-detection mode, the NSP platform continuously compares the new pixels' sensed data and the background data to find mismatches. When the number of mismatches exceeds a predefined threshold, which means a new object is detected, the architecture switches to the sensing mode. Our main contribution here is to avoid using power-hungry ADC circuits and to compare in a hybrid analog–digital mode without ADCs as shown in Fig. 4(a). Each pixel value can be stored with a configurable precision of 2-bit or 3-bit in the MEFET background array storing elements from the LSB to the MSB. We consider the stored data of MEFET as binary "1" when MEFET is in a low-resistance state (LRS) and "0" when the MEFET is in a high-resistance state (HRS). As shown in Fig. 4(a), assuming the background array is updated by 3-bit per pixel precision, three negative read voltages (V_1, V_2, and V_3) are calibrated based on the current characteristics of the MEFET device to translate the bit position of each MEFET data. As a result, when all MEFETs stored data is "1," the generated currents by MEFETs are doubled from the LSB to MSB. Leveraging the HRS of MEFET (63 MΩ), when all MEFETs stored data is "0," the drained current by MEFETs would be almost zero.

TABLE III
CURRENT LEVELS FOR THE 3-BIT PRECISION BACKGROUND ARRAY

$I_{sensor}(\mu A)$	Background data	$I_{background}(\mu A)$
0 to 9	"000"	0
9 to 26	"001"	17
26 to 43	"010"	34
43 to 60	"011"	51
60 to 77	"100"	69
77 to 94	"101"	86
94 to 111	"110"	103
111 to 120	"111"	120

The NSP platform performs a comparison on top of the Kirchhoff's current law (KCL) calculation in the node-n, as depicted in Fig. 4(a), with a current coming from the pixel (I_{sensor}) and a weighted current drained by our background MEFET storage cells depending on the pixel's value ($I_{\text{background}}$). Considering the 3-bit precision for the background array, the analog current generated from the new pixel (I_{sensor}) can be divided into $2^3 = 8$ levels, that will be compared with the weighted current summation of three MEFET bit-cells holding the previous pixel value as tabulated in Table III. At node-n, if I_{sensor} is the counterpart of the data stored in the MEFETs, I_{sensor} will be entirely drained by the MEFETs, for example, with $I_{\text{sensor}} = 0-9\ \mu$A when the background stored data is "000" draining $I_{\text{background}} = 0$. If I_{sensor} is not peer to $I_{\text{background}}$, the difference of two currents ($I_{\text{sensor}} - I_{\text{background}}$) passes through a resistance (R_{eq}) resulting in a negative or positive voltage at node-n. Now, if the pixel current is greater than the MEFET's weighted current, V_n will be positive and if the pixel current is smaller than the MEFET's weighted current, the resultant V_n will be negative. The NSP platform is equipped with two Sas to detect the mismatch. As shown in Fig. 4(b), V_n is fed to two CMOS Sas [Fig. 3(d)]. Each time, V_n is small enough ($-V_{\text{ref}} < V_n < +V_{\text{ref}}$), the output of both SA which are inputs of the CMOS NAND gate becomes "1," resulting in "0" at the output of NAND [Mtch in Fig. 4(b)].

D. NSP Event Detection Methodology

The NSP supports various configurations to enable accuracy-energy-efficiency tradeoffs at runtime. Different design configurations are determined by the notion of box_size $\in \{3, 5, 7\}$ and precision $\in \{2, 3\}$, where box_size represents the height and width of defined groups, and precision denotes the selected pixel data's bit-width. This data will be compared with the prior pixel value stored in the MEFET background array. As shown in Fig. 5, each $n \times n$ pixel box (here, e.g., 3×3) includes only one ON pixel, $(n-1)$ disconnected pixels, and $(n^2 - n)$ OFF pixels. Each column in the NSP's pixel array is enabled via a distinct VDD, and each row is enabled by a common row selector (R) signal as depicted in Fig. 3(a) to provide a sharper output. From the NSP's controller perspective, for the ON pixels, R and the column signals are both enabled. For the OFF pixels, R is disabled, but the column is enabled, and for the disconnected pixels the column is disabled regardless of the R value. The R signal is valued using $nx - 1$, where $n \in \{3, 5, 7\}$ and x is the row index $\in \{1, 2, \ldots, \lfloor 120/n \rfloor\}$. As a result, all the columns without central pixels are disconnected from the power supply

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.

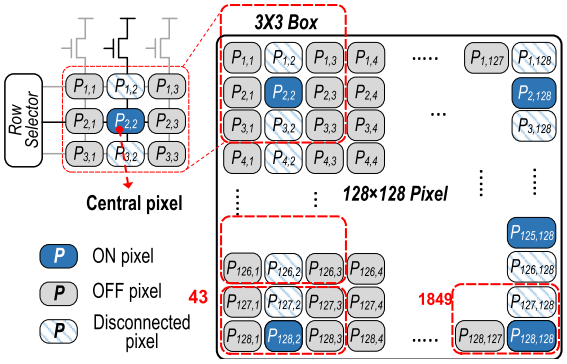

Fig. 5. Boxing pixels with the size of 3 × 3 with the possible situations.

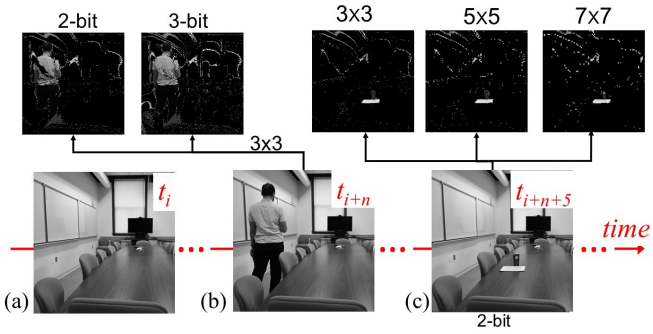

Fig. 6. Detecting object timeframes using the NSP platform. (a) → (b) detects a person leveraging different precision (2 and 3 bits) and (a) → (c) calculates differences in the images based on different box sizes.

(OFF), while the rest of the pixels in the columns containing the central pixel is disconnected using R signal. For instance, by setting box_size to 3 as in Fig. 5, all the pixels are grouped into a 3 × 3 shape, where the central pixel (i.e., $P_{2,2}$) is ON, the other two pixels ($P_{1,2}$ and $P_{3,2}$) in the same column are disconnected from ADCs because of the R values, and the rest ($P_{1,1}$, $P_{2,1}$, $P_{3,1}$, $P_{1,3}$, $P_{2,3}$, and $P_{3,3}$) are OFF.

Fig. 6 qualitatively compares different scenarios, including various box sizes and precisions. First, the NSP platform captures Fig. 6(a) and stores it as a static background within the MEFET background array. Then, an event has occurred in Fig. 6(b), and its results in precisions are shown on top. Fig. 6(c) illustrates the results for varied box sizes. After comparing the new input (t_{i+n+5}) with the background at t_i, the NSP detects that the mug and some papers are left on the desk. A smaller box size (e.g., 3 × 3) provides sharper output.

Algorithm 1 shows all the steps, including the event detection and sensing modes provided by the proposed architecture. The algorithm takes the size of the box, precision, and two thresholds, that is, $\text{threshold}_{\text{pixels}}$ and time_τ. The former is used for minimum changes, whereas the latter is leveraged to update the background. First, every row containing a central pixel, line-9, is activated, and the parallel comparison is performed in line-11 between all the central values and the previous value of the same pixel. The **parallel_comp** function takes the precision, which determines the required number of compared bits. In line-12, if the number of changes is greater than or equal to threshold $_{\text{pixels}}$, the row index is held in the turn_on_list. After checking all rows, the length of the turn_on array is checked. In the case of nonequality to zero, the mode is changed to the sensing mode, and the time counter is increased by one. This variable indicates how many times the

Algorithm 1 Proposed NSP Event Detection Algorithm

1: **Input$_1$:** $box\_size \in \{3, 5, 7\}$ & $precision \in \{2, 3\}$-bit
2: **Input$_2$:** $\text{threshold}_{\text{pixels}}$, $time_\tau$
3: **Output:** sensor_mode status
4: turn_on_list = []
5: **procedure** EVENT-DETECTION
6: **if** time $\geqslant time_\tau$: ▷ Merge steady objects with the background.
7: **update** (background)
8: **for** $i = \lfloor \frac{box\_size}{2} \rfloor + 1$ to 128 with step= box_size
9: **activate** (row_i)
10: pixel_values ← **parallel_read** ($\text{column}_{i,j}$) ▷ $j \in \{\lfloor \frac{box\_size}{2} \rfloor + 1, \ldots, 128\}$, with step= box_size
11: num_changes ← **parallel_comp** ($precision$, pixel_values, old_values)
12: **if** num_changes $\geqslant \text{threshold}_{\text{pixels}}$:
13: turn_on_list.**push** (i) ▷ i is row index.
14: **if** (length (turn_on_list) !=0)
15: time + = 1 ▷ Use it to update the background.
16: **enable** SENSOR MODE
17: **else**:
18: time = 0
19: **end procedure**
20: **procedure** SENSOR MODE
21: **while** (length (turn_on_list) !=0) **do**
22: row = turn_on_list.**pop**
23: **transfer** ($row - \lfloor box\_size \rfloor$ **to** $row + \lfloor box\_size \rfloor$)
24: **end while**
25: **end procedure**

NSP platform is switched to the sensing mode continuously. If this variable reaches time_τ, the NSP platform updates the background with the new values (line-7).

IV. Experimental Results

At the device level, we developed a Verilog-A compact model for the MEFET-RAM based on Section II to cosimulate with other peripheral CMOS circuits displayed in Fig. 3(a) in SPICE. To have an extensive comparison with existing MRAM counterparts, we used STT-MRAM and SOT-MRAM cells by jointly applying NEGF and LLG without and with spin Hall effect equations [7], [17]. At the circuit level, we use the 45-nm North Carolina State University (NCSU) Product Development Kit (PDK) library [21] to fully design and verify the NSP background and pixel arrays in HSPICE and to extract performance parameters such as delay and energy consumption. We use the Synopsys Design Compiler to develop the NSP's controller using a standard industry-level 45-nm technology. At the architecture level, we extensively modified NVSIM [22] as a memory performance evaluation tool to take memory configuration and circuit data for MRAM counterparts and a newly developed high-level MEFET library and report the array-level read/write energy and latency. At the application level, we developed an HW/SW simulator equipped with the proposed NSP event detection method that takes the architecture-level data for the MEFET background array and pixel array to estimate the system performance.

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.

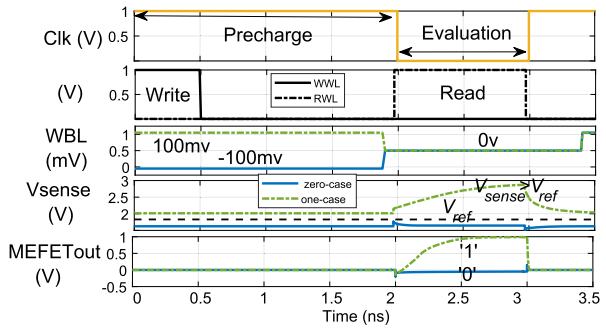

Fig. 7. Transient result of the read/write of MEFET array.

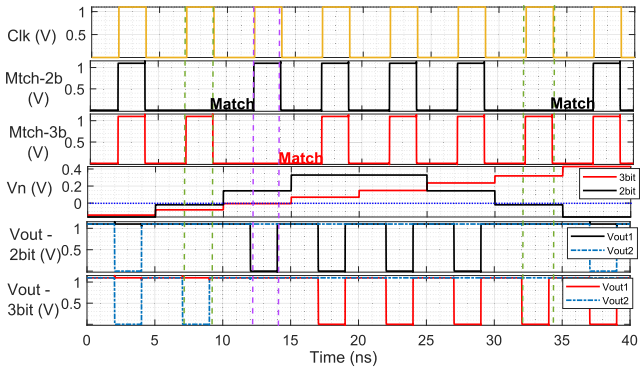

Fig. 8. Transient simulation result of the NSP's comparison unit.

A. Functionality Verification

1) MEFET Array: The transient simulation results of a single MEFET-RAM cell in the 128 × 128 background array are depicted in Fig. 7. We performed one write-and-read operation for each case (write zero case and write one case), as indicated by the solid blue and the dotted green lines. A 3-ns period clock is considered to synchronize the write/read operations. However, our experiment shows that a reliable read operation can be accomplished in <1-ns period. For the write operation, in the precharge phase when Clk = 1, the write voltage (±100 mv) is applied to WBL. This will change the MEFET resistance to $R_{\text{low}} = 1.05$ KΩ or $R_{\text{high}} = 63.4$ MΩ as discussed in Section II. For the read operation, the platform first grounds WWL and WBL signals and feed RBL by the tiny read current, $I_{\text{read}} = 900$ nA. By activating RWL, V_{sense} can be generated with respect to the resistance state of the MEFET bit-cell. As shown in the fourth waveform, the comparison between V_{sense} and a predefined V_{ref} is readily used to generate "0" and "1" at the output (MEFETout).

2) Near-Sensor Comparison Unit: To clarify the functionality of the proposed near-sensor comparator, the transient simulation waveforms are shown in Fig. 8. Here, the waveforms of V_n, the output of our two SAs, both in 2-bit and 3-bit configurations, and the output of the NAND gate (Mtch signal) are indicated. In our case study, a constant sensing current is applied to the node-n, and in each clock cycle, $2^3(/2^2)$ possible background bit combinations for 3(/2)-bit precision are considered. It can be observed when the data stored in the MEFETs is the counterpart to the current coming from the pixel, the V_ns' value is very close to zero, resulting in output "1" for two SAs at the same time. Thus, the Mtch signal is set to "1." For the 3-bit case, it occurs at the third clock cycle marked in Fig. 8 by the purple dashed line. For the 2-bit case, it occurs two times and is marked by the green dashed lines. We performed a Monte-Carlo simulation with 10 000 trials to investigate both device- and circuit-level sensing path variations with a 10% variation on the MEFET's TMR and a 5% variation on the access transistor. The results for 2-bit and 3-bit comparison scenarios are shown in Fig. 9(a) and (b). We observe that in the 2-bit scenario, with four combinations of stored background data, resulting in four voltage levels at the V_n node, at the worst case ∼80-mV voltage margin between two consecutive data is achieved. Thus, the SAs can easily and precisely detect match and mismatch using these voltages. In the 3-bit case, however, as the number of voltage stages is doubled, the margins between two consecutive levels are smaller. Yet, no overlapping among stages is observed and at the worst case, ∼45-mV voltage margin is obtained. Note that the very large resistance of the MEFET ($R_{\text{high}} = 63.4$ MΩ) completely cuts the reading current path that in turn enhances the voltage margins and generates a perfect "0." The difference in the ranges displayed in Figs. 8 and 9 is attributed to the utilization of distinct constant sensing currents in each case, thereby leading to different ranges of V_n. In fact, V_n is a 0-centered continuous voltage.

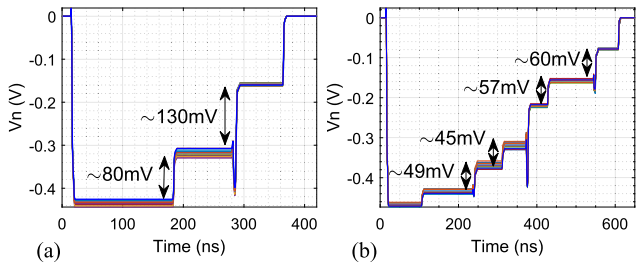

Fig. 9. Monte-Carlo simulation of V_n for (a) 2-bit and (b) 3-bit scenarios.

B. Energy Consumption and Execution Time

Fig. 10(a) reports the energy consumption breakdown for event detection mode (background updating and mismatch detection) across three platforms with STT-MRAM, SOT-MRAM, and MEFET as the background array. To have a fair comparison, STT-MRAM and SOT-MRAM platforms also leverage the NSP event detection method; however, a conventional ADC-based CMOS comparator is leveraged to detect the mismatch as opposed to the proposed ADC-less method. We assume that in the event detection mode, 5% of the time is allocated to updating the background and the rest of the time to detecting mismatches between a digitized pixel value and the prestored background in three under-test counterparts. The total power consumption for the central pixel comparison in the NSP platform can be estimated by $P_{\text{total}} = P_{\text{pixel}} + P_{\text{MEFET}} + P_{\text{compare}}$, where P_{pixel} represents the pixel sensing power. P_{MEFET} is the MEFET's read power during the current comparison, and P_{compare} denotes the power consumed by the near-sensor CMOS SAs and NAND gate. We observe that: 1) The average energy consumption in the STT-MRAM and SOT-MRAM platforms is 5477 and 5020 pJ, respectively, while the average energy consumption of the MEFET platform is 351 pJ. Therefore, the NSP platform with MEFET achieves on average ∼15.5× and 14.2× energy efficiency compared with the STT-MRAM and

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.

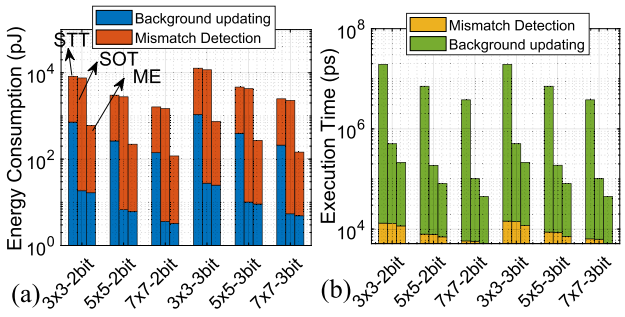

Fig. 10. Breakdown of (a) energy consumption and (b) execution time for three under-test platforms. In each bar group from left to right: STT-MRAM-, SOT-MRAM-, and MEFET-RAM-based designs.

SOT-MRAM-based designs, respectively. This comes mainly from eliminating the energy-hungry ADCs in the proposed near-sensor comparator and the higher write energy imposed by the STT-MRAM and SOT-MRAM-based technologies to update the background. 2) Averaged across six configurations, background updating accounts for 3% of the whole event detection operation. 3) The higher the precision is (here from 2- to 3-bit), the higher the energy budget is required for the edge device to perform such a near-sensor computation. 4) Within a particular precision, the larger box size brings higher energy efficiency to the system. Fig. 10(b) plots the execution time for three under-test platforms in six possible combinations of box sizes and precision. We observe 1) as the reading time of all three under-test technologies is almost the same, and execution times in the mismatch detection are almost the same. However, the proposed near-sensor comparator method has slightly reduced the required time for mismatch detection as opposed to the conventional ADC-comparator method and 2) a remarkable reduction in execution time can be observed in the background updating. The average execution time of the STT-MRAM platform is 10 110 ns, and for the SOT-MRAM-based platform, this number is 265 ns. However, the execution time of the MEFET platform is 114 ns. Therefore, relying on the MEFET switching at this stage has resulted in on average ∼88× and 2.3× improvements over the STT-MRAM and the SOT-MRAM platforms. Note that, the FeFET [3] also offers a high ON/OFF current ratio along with high off-state resistance. However, the writing voltage of the FeFET is much higher (∼3 V [3]) than the MEFET that is crucial for such small power-restricted edge sensors.

V. Conclusion

In this work, a reconfigurable NSP platform leveraging the MEFETs was proposed with an efficient background comparison method that enables event detection for edge sensors at a low cost by eliminating the need for power-hungry ADCs. Our evaluation results showed the proposed design can reduce energy consumption and execution time by a factor of ∼15× and ∼2.4× compared to the SOT-MRAM counterpart running the same method.

References

[1] F. D. V. R. Oliveira, H. L. Haas, J. G. R. C. Gomes, and A. Petraglia, "CMOS imager with focal-plane analog image compression combining DPCM and VQ," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 5, pp. 1331–1344, May 2013, doi: 10.1109/TCSI.2012.2226505.

[2] P. A. Dowben et al., "Towards a strong spin–orbit coupling magneto-electric transistor," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 4, no. 1, pp. 1–9, Jun. 2018, doi: 10.1109/JXCDC.2018.2809640.

[3] S. Das and J. Appenzeller, "FeTRAM. An organic ferroelectric material based novel random access memory cell," *Nano Lett.*, vol. 11, no. 9, pp. 4003–4007, Sep. 2011, doi: 10.1021/nl2023993.

[4] P. A. Dowben, D. E. Nikonov, A. Marshall, and C. Binek, "Magneto-electric antiferromagnetic spin–orbit logic devices," *Appl. Phys. Lett.*, vol. 116, no. 8, Feb. 2020, Art. no. 080502, doi: 10.1063/1.5141371.

[5] C. Ma, Y. Wang, Z. Shen, R. Chen, Z. Wang, and Z. Shao, "MNFTL: An efficient flash translation layer for MLC NAND flash memory," *ACM Trans. Design Autom. Electron. Syst.*, vol. 25, no. 6, pp. 1–19, Nov. 2020, doi: 10.1145/3398037.

[6] Y. Pan et al., "A multilevel cell STT-MRAM-based computing in-memory accelerator for binary convolutional neural network," *IEEE Trans. Magn.*, vol. 54, no. 11, pp. 1–5, Nov. 2018, doi: 10.1109/TMAG.2018.2848625.

[7] X. Fong et al., "Spin-transfer torque devices for logic and memory: Prospects and perspectives," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 35, no. 1, pp. 1–22, Jan. 2016, doi: 10.1109/TCAD.2015.2481793.

[8] W. Zhao, E. Belhaire, C. Chappert, and P. Mazoyer, "Spin transfer torque (STT)-MRAM-based runtime reconfiguration FPGA circuit," *ACM Trans. Embedded Comput. Syst.*, vol. 9, no. 2, pp. 1–16, 2009, doi: 10.1145/1596543.1596548.

[9] H. Akinaga and H. Shima, "Resistive random access memory (ReRAM) based on metal oxides," *Proc. IEEE*, vol. 98, no. 12, pp. 2237–2251, Dec. 2010, doi: 10.1109/JPROC.2010.2070830.

[10] N. Sharma, C. Binek, A. Marshall, J. Bird, P. Dowben, and D. Nikonov, "Compact modeling and design of magneto-electric transistor devices and circuits," in *Proc. IEEE Int. Syst. Chip Conf. (SOCC)*, Sep. 2018, pp. 146–151, doi: 10.1109/SOCC.2018.8618494.

[11] A. Mahmood et al., "Voltage controlled Néel vector rotation in zero magnetic field," *Nature Commun.*, vol. 12, no. 1, pp. 1–8, Mar. 2021, doi: 10.1038/s41467-021-21872-3.

[12] D. E. Nikonov and I. A. Young, "Benchmarking of beyond-CMOS exploratory devices for logic integrated circuits," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 1, pp. 3–11, 2015, doi: 10.1109/JXCDC.2015.2418033.

[13] H.-J. Chuang et al., "Low-resistance 2D/2D ohmic contacts: A universal approach to high-performance WSe_2, MoS_2, and $MoSe_2$ transistors," *Nano Lett.*, vol. 16, no. 3, pp. 1896–1902, Mar. 2016, doi: 10.1021/acs.nanolett.5b05066.

[14] N. Sharma, J. P. Bird, C. Binek, P. A. Dowben, D. Nikonov, and A. Marshall, "Evolving magneto-electric device technologies," *Semicond. Sci. Technol.*, vol. 35, no. 7, Jul. 2020, Art. no. 073001, doi: 10.1088/1361-6641/ab8438.

[15] M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov, "Modeling of nanoscale devices," *Proc. IEEE*, vol. 96, no. 9, pp. 1511–1550, Sep. 2008, doi: 10.1109/JPROC.2008.927355.

[16] N. Sharma, A. Marshall, J. Bird, and P. Dowben, "Verilog—A based compact modeling of the magneto-electric FET device," in *Proc. 5th Berkeley Symp. Energy Efficient Electron. Syst. Steep Transistors Workshop (E3S)*, Oct. 2017, pp. 1–3, doi: 10.1109/E3S.2017.8246186.

[17] X. Fong, S. K. Gupta, N. N. Mojumder, S. H. Choday, C. Augustine, and K. Roy, "KNACK: A hybrid spin-charge mixed-mode simulator for evaluating different genres of spin-transfer torque MRAM bit-cells," in *Proc. Int. Conf. Simulation Semiconductor Processes Devices*, Sep. 2011, pp. 51–54, doi: 10.1109/SISPAD.2011.6035047.

[18] N. Sharma, A. Marshall, J. Bird, and P. Dowben, "Novel ring oscillator design using ME-MTJ based devices," in *Proc. 5th Berkeley Symp. Energy Efficient Electron. Syst. Steep Transistors Workshop (E3S)*, Oct. 2017, pp. 1–3, doi: 10.1109/E3S.2017.8246187.

[19] A. Iyama and T. Kimura, "Magnetoelectric hysteresis loops in Cr_2O_3 at room temperature," *Phys. Rev. B, Condens. Matter*, vol. 87, no. 18, May 2013, Art. no. 180408, doi: 10.1103/PhysRevB.87.180408.

[20] S. Angizi, N. Khoshavi, A. Marshall, P. Dowben, and D. Fan, "MeF-RAM: A new non-volatile cache memory based on magneto-electric FET," *ACM Trans. Design Autom. Electron. Syst.*, vol. 27, no. 2, pp. 1–18, Mar. 2022, doi: 10.1145/3484222.

[21] (2011). *NCSU EDA Freepdk45*. [Online]. Available: http://www.eda.ncsu.edu/wiki/FreePDK45

[22] X. Dong, C. Xu, Y. Xie, and N. P. Jouppi, "NVSim: A circuit-level performance, energy, and area model for emerging nonvolatile memory," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 31, no. 7, pp. 994–1007, Jul. 2012, doi: 10.1109/TCAD.2012.2185930.

Authorized licensed use limited to: New Jersey Institute of Technology. Downloaded on November 29,2023 at 01:09:07 UTC from IEEE Xplore. Restrictions apply.