

Received 16 May 2024; accepted 15 June 2024. Date of publication 19 June 2024; date of current version 29 July 2024.

The review of this article was arranged by Editor P.-W. Li.

Digital Object Identifier 10.1109/JEDS.2024.3416441

# Efficient Implementation of Mahalanobis Distance on Ferroelectric FinFET Crossbar for Outlier Detection

MUSAIB RAFIQ<sup>1b</sup> (Graduate Student Member, IEEE), YOGESH SINGH CHAUHAN<sup>1b</sup> (Fellow, IEEE),  
AND SHUBHAM SAHAY<sup>1b</sup> (Senior Member, IEEE)

Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

CORRESPONDING AUTHOR: S. SAHAY (e-mail: ssahay@iitk.ac.in)

This work was supported in part by the Prime Minister's Research Fellowship (PMRF), Semiconductor Research Corporation SRP Task under Grant 3056.001; in part by the Ministry of Education's Scheme for Transformational and Advanced Research in Sciences (STARS) Project under Grant MoE-STARS/STARS-2/2023-0023; and in part by the Swarnajayanti Fellowship under Grant DST/SJF/ETA/02/17-18.

**ABSTRACT** The developments in the nascent field of artificial-intelligence-of-things (AIoT) relies heavily on the availability of high-quality multi-dimensional data. A huge amount of data is being collected in this era of big data, predominantly for AI/ML algorithms and emerging applications. Considering such voluminous quantities, the collected data may contain a substantial number of outliers which must be detected before utilizing them for data mining or computations. Therefore, outlier detection techniques such as Mahalanobis distance computation have gained significant popularity recently. Mahalanobis distance, the multivariate equivalent of the Euclidean distance, is used to detect the outliers in the correlated data accurately and finds widespread application in fault identification, data clustering, single-class classification, information security, data mining, etc. However, traditional CMOS-based approaches to compute Mahalanobis distance are bulky and consume a huge amount of energy. Therefore, there is an urgent need for a compact and energy-efficient implementation of an outlier detection technique which may be deployed on AIoT primitives, including wireless sensor nodes for in-situ outlier detection and generation of high-quality data. To this end, in this paper, for the first time, we have proposed an efficient Ferroelectric FinFET-based implementation for detecting outliers in correlated multivariate data using Mahalanobis distance. The proposed implementation utilizes two crossbar arrays of ferroelectric FinFETs to calculate the Mahalanobis distance and detect outliers in the popular Wisconsin breast cancer dataset using a novel inverter-based threshold circuit. Our implementation exhibits an accuracy of 94.1% which is comparable to the software implementations while consuming a significantly low energy (27.2 pJ).

**INDEX TERMS** Crossbar array, deep learning, ferroelectric FETs, Mahalanobis distance, outlier detection.

## I. INTRODUCTION

The recent developments in the field of machine learning and artificial intelligence platforms have led to a transition from the compute-centric paradigm to the data-centric paradigm in this era of artificial-intelligence-of-things (AIoT) and big data [1], [2], [3], [4], [5], [6]. Most machine learning and deep learning techniques rely heavily on data availability for training and testing of the models [7], [8]. Since data is collected in huge volumes, it may contain some entries that are erroneous or significantly different from most entries. These entries are popularly known as outliers and degrade

the accuracy and reliability of the models. Therefore, their detection becomes crucial prior to processing or utilizing the data. Moreover, the outliers often provide a useful insight into the data and should be analyzed carefully before discarding from the dataset [9].

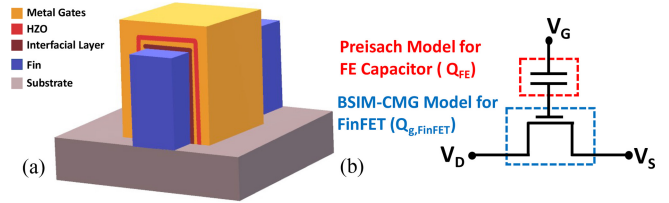
Detecting outliers in one-dimensional data exploiting techniques such as three standard deviations, box plots, etc., is not a complicated task. However, in the case of multi-dimensional data, which is more prevalent in ML/DL models, detecting outliers is a complex task. In multi-dimensional data, even the probability of the occurrence of an outlier is

higher. Although Euclidean distance can be utilized to detect outliers in the multi-dimensional data, it fails miserably when the variables in the data are related to each other such that change in the value of one variable causes a change in the values of other variable(s). In such complex cases where data is correlated, and the data dimensions are not equally weighted, one of the most promising methods to detect outliers is by computing the Mahalanobis distance [10]. This approach relies on finding the distance between the data points and the center of the distribution with the help of a covariance matrix of the variables. Apart from outlier detection, Mahalanobis distance has been used for a wide range of applications including fault identification, single-class classification, clustering, information security, data mining, etc., [11], [12], [13]

Moreover, the computing paradigm is gradually shifting towards non-von-Neumann architectures, owing to their high area and energy-efficiency as compared to traditional digital computing primitives with memory wall. The rapid advancements in the field of emerging non-volatile memories-based cross-point array platforms has enabled this transition [14], [15], [16], [17]. Various implementations of these neuro-morphic platforms based on Resistive RAMs (RRAMs), Magnetoresistive memories (MRAMs), and Ferroelectric FETs (FeFETs) have been shown [18], [19], [20]. Among the emerging non-volatile memories, FeFETs have attracted significant attention recently due to the discovery of ferroelectricity in CMOS-compatible hafnium oxide ( $HfO_2$ ) [21]. Doped  $HfO_2$  based FeFETs have been successfully integrated into the 28-nm high-k metal gate (HKMG) [22] and 22-nm fully depleted silicon on insulator (FDSOI) nodes [23]. In addition to their CMOS compatibility, FeFETs offer high scalability, multi-level storage capability, field-driven programming, and low-voltage switching [22], [23], [24], [25]. Consequently, FeFETs have found their use as embedded nonvolatile memory, ultra-efficient noise immune nanoscale circuits and systems, programmable delay element, and in many non-von-Neumann computing platforms [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36].

As we are heading towards a new computing generation defined by compact and energy-efficient architectures, detecting outliers efficiently within this framework becomes imperative. Although hardware demonstration of Euclidean distance on memristive crossbar arrays has already been shown [37], [38], Euclidean distance fails to detect outliers in correlated multi-dimensional data. Detecting outliers in multidimensional data is crucial in this era of AIoT, particularly while performing AI/ML workloads that rely predominantly on multi-dimensional datasets.

Considering the efficacy of Mahalanobis distance while detecting outliers in multi-dimensional data in this work, for the first time, a dedicated hardware implementation of Mahalanobis distance for outlier detection is demonstrated on the Wisconsin Breast cancer dataset. To achieve this, two ferroelectric FinFET-based crossbars are used to calculate the Mahalanobis distance. The calculated Mahalanobis distance



**FIGURE 1. (a) 3D view of a ferroelectric (Fe)-FinFET. (b) The equivalent circuit model for Fe-FinFET with the series combination of the ferroelectric capacitor (modeled using multi-domain Preisach model) and an underlying FinFET (modeled using BSIM-CMG model).**

is then applied as an input to the detect and fire circuit made of a simple CMOS inverter whose threshold value is tuned to the corresponding cut-off value calculated from the Chi-square distribution with  $n$  degrees of freedom. The proposed implementation of detecting outliers is extremely energy-efficient and shows an accuracy of 94.1%, which is comparable to the software implementations.

The rest of the manuscript is organized as follows: device simulation methodology used for FeFETs is discussed in Section II. The methodology used for calculating Mahalanobis Distance and the performance metrics are discussed in Sections III and IV, respectively. Finally, the conclusions are drawn in Section V.

## II. DEVICE SIMULATION METHODOLOGY

The 3D view of the Fe-FinFET used in this study is shown in Fig. 1. A zirconium-doped hafnium oxide (HZO) is used in the gate stack as the ferroelectric layer. To accurately capture the Fe-FinFET characteristics, we have utilized an experimentally calibrated compact model of the FE capacitor based on the multi-domain Preisach model [39], [40] and connected it to the industry standard FinFET (BSIM CMG) compact model, which is calibrated to the measured characteristics of a commercially fabricated minimum channel length n-FinFET in 14 nm technology [39], [41]. The inclusion of an internal metal gate within the gate stack of Fe-FinFET, as depicted in Fig. 1, simplifies the process for model development for Fe-FinFETs by considering two different circuit entities: Fe-Cap ( $C_{FE}$ ) and the underlying conventional FinFET [39]. The equivalent circuit used for modeling is also shown in Fig. 1(b). For capturing the switching characteristics, an auxiliary voltage  $V_{aux}$  to which the ferroelectric dipoles respond after relaxation is calculated first [39]. The value of  $V_{aux}$  is computed from (1), which is modeled using a simple R-C delay network:

$$V_{aux} = V_{in} - \tau v \frac{d}{dt} V_{aux} \quad (1)$$

In (1),  $\tau v$  represents the relaxation time, and the applied input voltage is represented by  $V_{in}$ . Furthermore, the polarization is calculated by (2):

$$P_{aux} = m \cdot P_s \cdot \tanh(w(V_{aux} \pm V_c)) + P_{off} \quad (2)$$

$$w = \frac{t_{fe}}{2V_c} \cdot \ln\left(\frac{P_s + P_r}{P_s - P_r}\right) \quad (3)$$

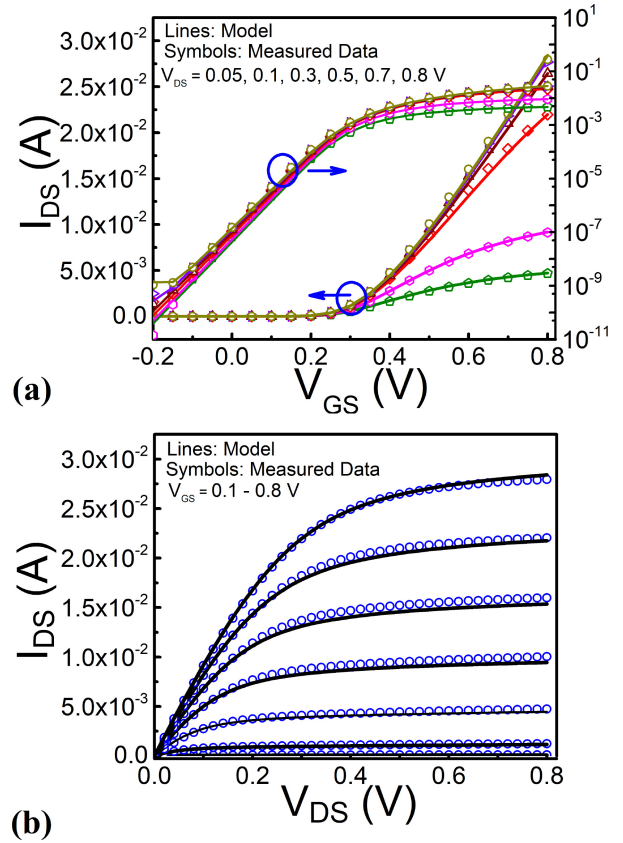
$P_r$  represents the remnant polarization,  $t_{fe}$  represents the thickness of the FE layer,  $m$  is the slope of the curve,  $P_s$  is saturation polarization, and  $P_{off}$  represents the offset polarization. The upward (+) and downward trajectory (−) of the loop dictates the sign of the tanh function. Moreover, the values of  $m$  and  $P_{off}$  are calculated by using the polarization history. ‘ $m$ ’ and  $P_{off}$  determine the P-V characteristics of the minor loops and for the major (saturation) loop,  $m = 1$  and  $P_{off} = 0$ . Turning point tracing is used for polarization history. Another RC delay network is used to determine the turning points.  $m$  and  $P_{off}$  are given by (4) and (5), respectively

$$m_{\uparrow,\downarrow} = \frac{P_{aux,u} - P_{aux,l}}{P_s \cdot (\tanh(w(V_{aux,u} \pm V_c)) - \tanh(w(V_{aux,l} \pm V_c)))} \quad (4)$$

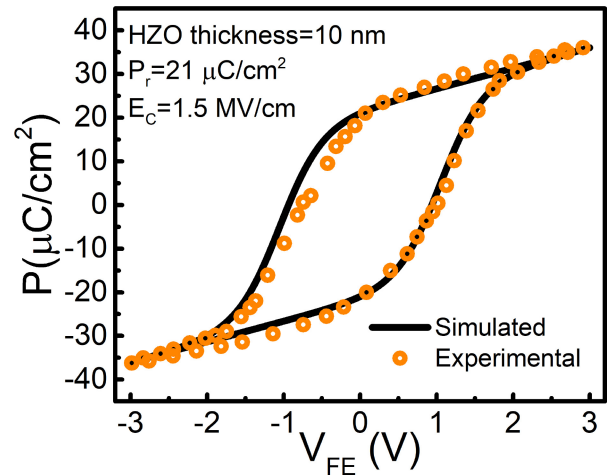
$$P_{off\uparrow,\downarrow} = P_{aux,u/l} - m_{\uparrow,\downarrow} \cdot P_s \cdot \tanh(w(V_{aux,u/l} \pm V_c)) \quad (5)$$

Arrows represent the trajectory direction and, consequently, the sign of the tanh function. Furthermore, following the principle of charge conservation, the FE capacitor ( $Q_{FE}$ ) charge must be equal to the gate charge on the underlying FinFET ( $Q_{g,FinFET}$ ). This charge balance condition is solved simultaneously with the principle of voltage division ( $V_g = V_{FE} + V_{FinFET}$ ) to obtain the characteristics of Fe-FinFET [39]. Furthermore, a similar approach for modeling FeFETs has been used in prior works [42]. The simulations were performed using the commercial SPICE simulator Cadence Spectre. Moreover, for baseline FinFET calibration, experimental data is extracted from a commercially fabricated minimum channel length n-FinFET in 14 nm technology having  $W_{eff}/FIN$  of 85 nm and a fin thickness of 7 nm with fin height of 39 nm while a 10 nm thick ferroelectric layer is used in the gate stack of Fe-FinFET. To extract the on-wafer characteristics, a Cascade Summit 11K probe station, along with a Keysight B-1500A parameter analyzer, was used. The experimentally extracted DC  $I_{DS}-V_{DS}$  and  $I_{DS}-V_{GS}$  characteristics are fitted by tuning the model parameters of BSIM CMG as shown in Fig. 2. In addition, model parameters of the FE-cap model were also fine-tuned to reproduce the experimental [42] polarization-voltage characteristics shown in Fig. 3. To further validate the accuracy of the compact model for Fe-FinFET, we have reproduced the experimental characteristics of the Fe-FinFET reported in [43] with our compact model by simple fine-tuning of the parameters of the underlying industry-standard model for FinFETs (BSIM-CMG). Fig. 4 shows the fitting of the read characteristics after the program and erase operation. As can be observed from Fig. 4, the simulation results show excellent agreement with the experimental data of [43] validating the efficacy of our compact model.

The capability to program more than 32 polarization states (which correspond to 32 intermediate threshold voltage levels) between the extreme states has already been demonstrated experimentally in [44]. Fig. 5(a) shows the incremental pulse-amplitude waveform scheme, which can be utilized to achieve intermediate  $V_{th}$  programming in Fe-FinFETs. The transfer(read) characteristics of the

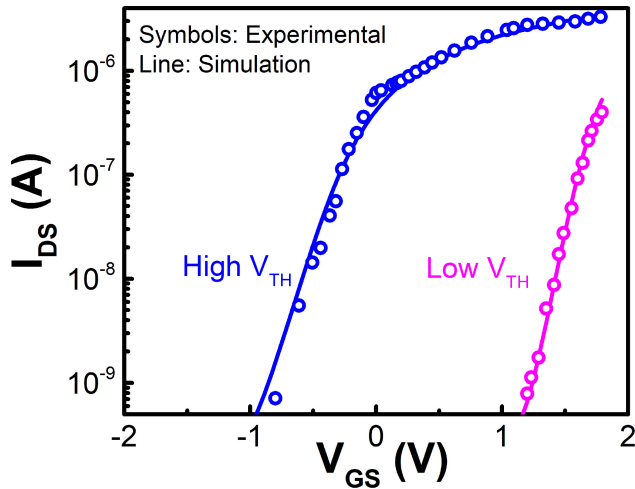


**FIGURE 2.** FinFET model calibration with experimentally measured data for 14 nm FinFETs: (a) transfer characteristics for different drain voltages and (b) output characteristics for different gate voltages.

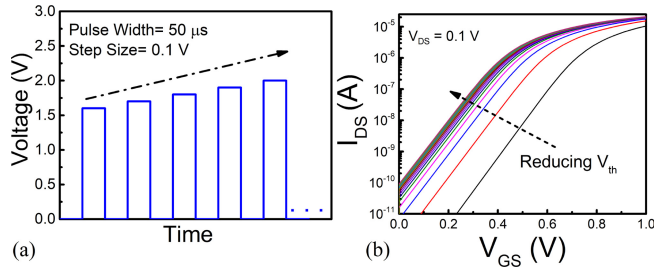


**FIGURE 3.** Polarization-voltage characteristics of an HZO FE capacitor model, with parameters tuned to reproduce experimental characteristics [42].

intermediate states corresponding to waveforms are extracted after every write pulse and are shown in Fig. 5(b). In this work, we have utilized this experimentally calibrated compact model for emulating Fe-FinFETs and realize five bit-weights in cross-point architecture for outlier detection.



**FIGURE 4.** Read (transfer) characteristics after the program (+3 V) and erase (-3 V) operations in Fe-FinFET. Model parameters were tuned to reproduce the experimental data from [43], showing the efficacy of the compact model.



**FIGURE 5.** (a) Incremental pulse amplitude programming of Fe-FinFETs (increasing amplitude from 1.6 V to 3 V) (b) Read (transfer) characteristics of Fe-FinFETs after incremental pulse amplitude programming.

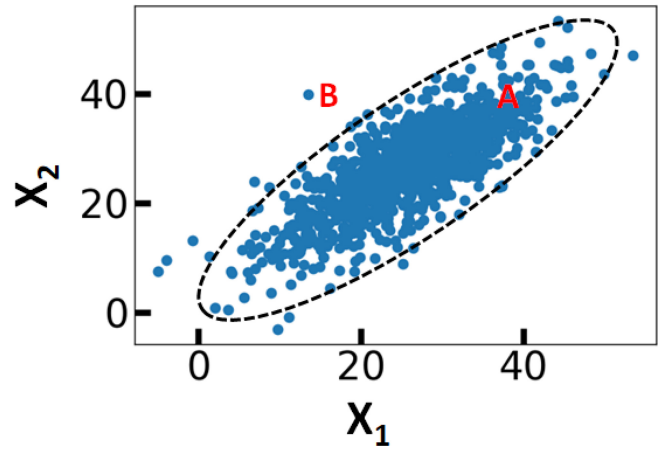
### III. MAHALANOBIS DISTANCE COMPUTATION

#### A. MAHALANOBIS DISTANCE

Applications such as the  $k$ -nearest neighbor,  $k$ -means clustering, outlier detection, etc., rely on the calculation of distance for multi-variate data. Euclidean distance generally used to find out the shortest distance between the two points as:

$$ED^2 = (x - \mu)^T \cdot (x - \mu) \quad (6)$$

where  $x$  is the vector of observation, and  $\mu$  is the mean value. Euclidean distance has been used extensively in many machine-learning techniques, like clustering, nearest neighbor classification, etc., as a distance metric. For brain-inspired competitive learning and machine learning applications like  $k$ -means clustering, energy-efficient memristive Euclidean distance engines have also been demonstrated [37], [38]. However, in case of outlier detection, Euclidean distance works only when the data dimensions are not correlated. To illustrate this, scatter plot of two variables that are positively correlated with each other is shown in Fig. 6. If Euclidean distance is used as the metric for classifying outliers, it will mark data present on both points A and B as non-outliers or outliers since the Euclidean distance of both points A and B from the centroid of the scatter plot is same. However,



**FIGURE 6.** Scatter plot of two variables that are correlated. The dashed boundary is marked by using the Mahalanobis distance.

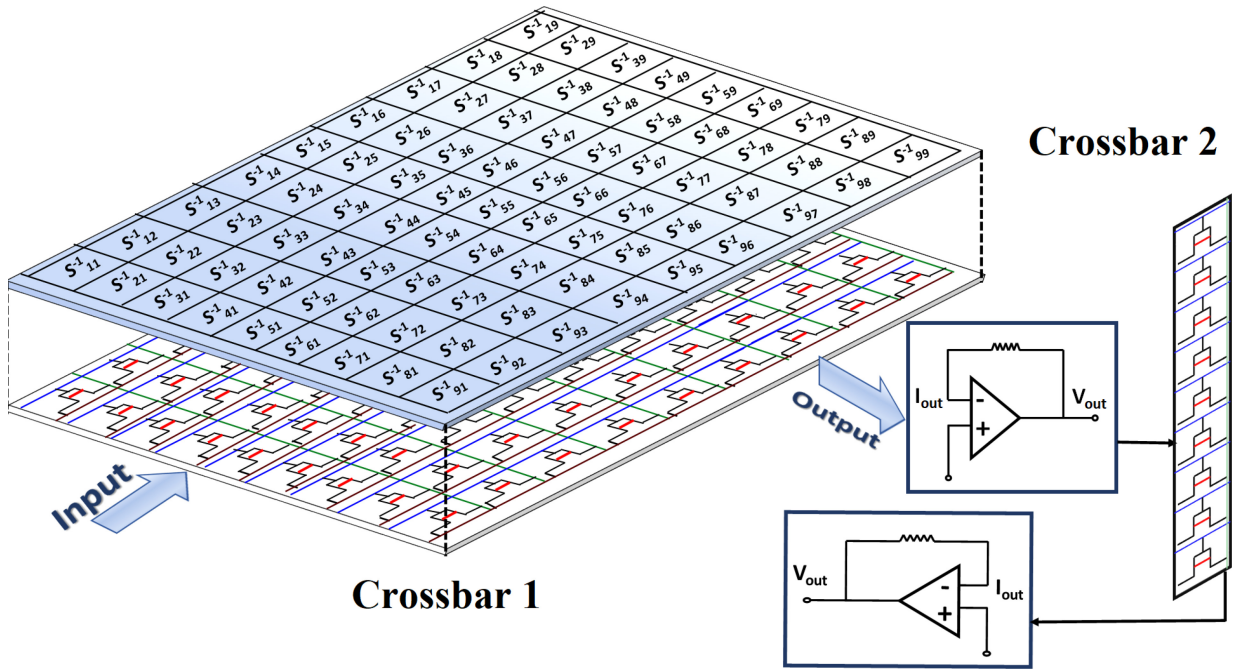
as can be observed from Fig. 6, only data present at point A is a non-outlier, and point B is an outlier as it falls outside the error ellipse. In most of the real-world datasets used in the ML/DL applications the data dimensions are correlated. In such cases, the Mahalanobis distance has been found to be a promising alternative to the Euclidean distance while detecting outliers in multivariate data. In fact, the dashed boundary of the error ellipse in Fig. 6, which accurately captures the distribution of data, is marked using the Mahalanobis distance. Additionally, Mahalanobis distance has been effectively utilized in applications such as classification problems, clustering, information security, etc. The Mahalanobis distance is defined as the distance between a point and distribution and requires the calculation of a covariance matrix which makes it computationally intensive. The expression for computing Mahalanobis distance is given as:

$$MD^2 = (x - \mu)^T \cdot S^{-1} \cdot (x - \mu) \quad (7)$$

where  $MD^2$  represents the squared Mahalanobis distance,  $x$  is the vector of the observation,  $\mu$  is the vector of mean values of the independent variables, and  $S^{-1}$  represents the inverse covariance matrix of independent variables. The covariance matrix removes the redundant information from the correlated variables. The general form of a covariance matrix of  $n$ -dimensional data is given below

$$S = \begin{bmatrix} \text{Var}(x_1) & \text{Cov}(x_1, x_2) & \dots & \text{Cov}(x_1, x_n) \\ \text{Cov}(x_2, x_1) & \text{Var}(x_2) & \dots & \text{Cov}(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(x_n, x_1) & \text{Cov}(x_n, x_2) & \dots & \text{Var}(x_n) \end{bmatrix} \quad (8)$$

where  $\text{Var}(x_m)$  represents the variance of variable  $x_m$  and  $\text{Cov}(x_m, x_n)$  represents the covariance between  $x_m$  and  $x_n$ . Mahalanobis distance reduces to Euclidean distance if  $S$  is an identity matrix, which implies all dimensions are statistically independent of each other.



**FIGURE 7.** A two-Fe-FinFET based crossbar array implementation for calculating Mahalanobis distance. The inverse covariance matrix of data is mapped onto crossbar 1. The column currents of crossbar 1 are applied to the TIAs for conversion to voltage values which act as inputs to crossbar 2. The single-column current of crossbar 2 is applied to TIA, whose output represents the squared Mahalanobis distance.

### B. CROSSBAR ARRAY-BASED IMPLEMENTATION OF MAHALANOBIS DISTANCE

In this work, we have proposed a two-step approach for hardware implementation of Mahalanobis distance based on Fe-FinFET crossbar arrays since the calculation of Mahalanobis distance requires two dot-product operations (equation (7)). We have utilized two crossbar arrays: the output of the dot product operation performed on the first crossbar is fed as input to the second crossbar for the final Mahalanobis distance calculation, as shown in Fig. 7. In our proof-of-concept implementation, we have used the Wisconsin Breast Cancer [45] Dataset, which has a dimensionality of 9. Therefore, the size of the inverse-covariance matrix is  $9 \times 9$  and the corresponding size of Fe-FinFET crossbar 1 is  $9 \times 18$  (since a differential weight mapping scheme is utilized), and the size of crossbar 2 is  $9 \times 2$ . The main function of crossbar 1 is to perform one of the two dot product operations given in equation (7). In order to perform this, first, the inverse covariance matrix of the dataset needs to be computed. The inverse covariance matrix also has a size equal to  $n \times n$ . The values of the elements from the computed inverse covariance matrix are then correspondingly mapped onto crossbar 1 as the different polarization states, which yield different conductance states when measured at a fixed gate and drain voltage during the read operation, as shown in Fig. 7. This matrix element-to-conductance state mapping is done utilizing a differential weight-mapping scheme [46] (for brevity, only positive weight mappings are shown in Fig. 7). A linear mapping technique was used while programming the conductance-state ( $G$ ) of the Fe-FinFETs according to

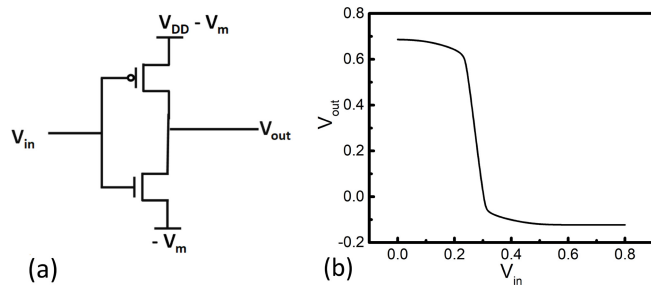
the weight ( $W$ ) as:

$$G_{\pm} = G_{\text{avg}} \pm \frac{K_g W}{2} \quad (9)$$

where

$$G_{\text{avg}} = \frac{G_{\text{min}} + G_{\text{max}}}{2} \text{ and } K_g = \frac{G_{\text{max}} - G_{\text{min}}}{\max |W|}$$

$G_{\text{max}}$  and  $G_{\text{min}}$  represent the maximum and minimum conductance of the Fe-FinFET corresponding to the extreme polarization states, respectively, and  $\max |W|$  represents the maximum magnitude of the weight. After the mapping is done successfully and the Fe-FinFET crossbar is programmed, the inputs are applied for the computation of the first dot-product. Here inputs are encoded as the scaled drain voltages (in the range of  $[0, 50 \text{ mV}]$  for 100 ns to ensure non-destructive read operation) corresponding to the term  $x - \mu$ , and applied to the rows of the crossbar array (while gate voltages are kept at 0.5 V). After the application of the inputs, the multiply and accumulate operation is inherently performed in-situ in an energy-efficient way due to the physical laws (Ohm's law and Kirchoff's law). The column currents obtained from this crossbar 1 of Fe-FinFETs represent the first dot-product. These output current values are converted to the voltage domain using a trans-impedance amplifier (TIA) having a gain of  $10 \text{ k}\Omega$ , as shown in Fig. 7. The output voltage coming from crossbar 1 acts as the input for the next crossbar. Crossbar 2 is used for performing the final dot product operation. The gain of the trans-impedance amplifier has been adjusted to yield voltages in the range of read voltage for the Fe-FinFETs of crossbar 2. The weights



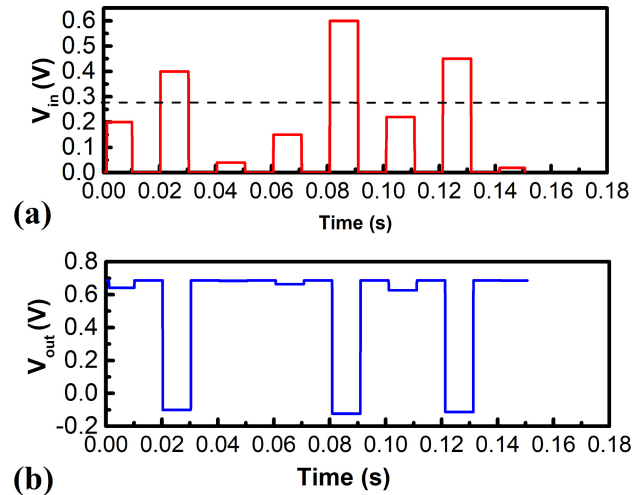
**FIGURE 8.** (a) Schematic of an inverter (b) voltage transfer characteristics. The switching threshold voltage of the inverter was tuned according to the critical value by applying a voltage  $V_{DD} - V_m$  and  $-V_m$  ( $V_m = 0.123$  V) at the source terminals of PMOS and NMOS FETs, respectively.

of crossbar 2 are mapped corresponding to the vector  $x - \mu$  utilizing the differential weight mapping scheme owing to its capability to represent negative weights and resiliency to noise. It may be noted here that while the input drain voltages of Fe-FinFETs in crossbar1 are mapped according to  $x - \mu$ , the weights (conductance state measured at  $V_{GS} = 0.5$  V,  $V_{DS} = 0.05$  V) of the crossbar 2 are mapped according to  $x - \mu$ . The output current represents the squared Mahalanobis distance value between the input vector  $x$  and the distribution, which is then applied to another TIA, feeding it to a CMOS-inverter-based outlier detection circuitry.

**IV. PERFORMANCE METRICS**

**A. OUTLIER DETECTION**

For detecting the outliers based on the calculated squared Mahalanobis distance, we have proposed a novel CMOS-inverter-based firing circuit as shown in Fig. 8(a). The switching threshold voltage of the inverter is tuned to the critical value computed from the Chi-Square distribution at a 0.001 significance level and with ‘ $n$ ’ degrees of freedom. The critical value was found to be 27.87 for this input dataset. The tripping point (inverter switching threshold voltage) of the inverter was set to 274 mV, corresponding to the critical value of 27.87. The inverter switching threshold voltage (tripping point) was tuned according to the critical value by applying a voltage  $V_{DD} - V_m$  and  $-V_m$  ( $V_m = 0.123$  V) at the source terminals of the P-FinFET and N-FinFET, respectively, as shown in Fig. 8(a). The voltage transfer characteristic of the inverter with a tuned switching threshold voltage is shown in Fig. 8(b). The output voltages from crossbar 2, which represent the squared Mahalanobis distance values, are then applied to the input of the inverter. Fig. 9(a) shows the input voltage waveform to the inverter with a dashed line representing the cutoff value. The classification accuracy drops only when the Mahalanobis distance corresponding to the input lies at/near the threshold boundary (critical value), and even a slight miscalculation due to the inherent non-linear behavior/spatial variation in the Fe-FinFETs results in an incorrect classification. Furthermore, a negative spike is observed at the output of the outlier detection circuit (Fig. 9(b)) only when the Mahalanobis distance is greater than the critical value (the tripping point of the inverter 274 mV corresponding to the critical value of 27.87), which signifies an outlier. However, it may be noted



**FIGURE 9.** (a) Input and (b) output voltage waveforms of the outlier detection circuit. Negative spikes in (b) represent an outlier.

that not all the negative spikes observed in Fig. 9(b) correspond to the correct classification of inputs as an outlier. A negative spike can also occur when a non-outlier is marked as an outlier which is the case when the original squared Mahalanobis distance is less than the critical value, but the value obtained from the proposed implementation comes out to be greater than the critical value. Also, an outlier can be misclassified as a non-outlier when the original squared Mahalanobis distance is greater than the critical value, but the value from the proposed implementation comes out to be less than the critical value, resulting in no spike generation.

**B. ACCURACY**

To analyze the efficacy of our proposed implementation, we randomly selected inputs from the Wisconsin breast cancer dataset and fed them as the input to crossbar1. We compared the Mahalanobis distance computed at the output of the TIA of crossbar 2 against the results obtained via software simulations. The classification accuracy is measured as the ratio of the number of inputs correctly marked as outlier/non-outlier to the total number of inputs tested. The inputs are misclassified in the proposed implementation either when an outlier is marked as a non-outlier or vice-versa. Although the output obtained from the proposed implementation deviates somewhat from the software results owing to the limited computational precision, non-linear characteristics, and spatial variation of the Fe-FinFETs utilized in this work, the proposed implementation exhibits a high accuracy of 94.1% while detecting outliers. Moreover, the Mahalanobis distance values computed by the Fe-FinFET-based implementation show a mean relative error of 12.76%. It may be noted that even an appreciable deviation in the calculated value may not degrade the accuracy of the implementation for outlier detection significantly as long as the threshold of the firing circuit is chosen appropriately.

To analyze the impact of spatial variations on the proposed FeFET-based implementation of Mahalanobis distance, we

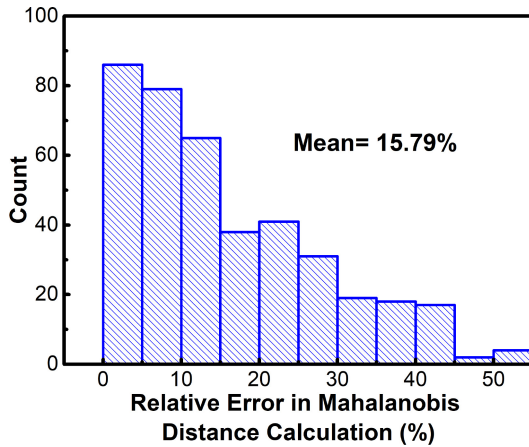


FIGURE 10. Impact of  $V_{TH}$  variation on Mahalanobis Distance Calculation.

have performed Monte-Carlo simulations considering the device-to-device variability for the 14 nm technology node. For performing the Monte-Carlo simulations, we have introduced a stochastic distribution of the threshold voltage,  $V_{TH}$  of the underlying FinFETs by changing the model parameter “DELVTRAND” of the industry-standard BSIM-CMG model with the best-case variation ( $\sigma V_{TH}$ ) of 15 mV [47], [48], [49]. The Monte-Carlo simulations were performed for 20 different inputs sampled from the dataset such that the squared Mahalanobis distance values corresponding to these inputs spans the entire range with 20 iterations for each input. The results of Monte-Carlo simulations, shown in Fig. 10, indicate that the proposed implementation exhibits a mean relative error of 15.79% (3.03% degradation from the baseline value (12.76%)) in the presence of spatial (device-to-device) variations. Moreover, the presence of spatial variations does not degrade the accuracy of outlier detection, and an accuracy close to 94% (the variation-free baseline value) is still achieved. However, for the inputs where the corresponding squared Mahalanobis distance value is very close to the threshold boundary, variability in the devices can play a significant role in degrading the accuracy for outlier detection, as even a slight variation will lead to a false detection.

The retention of the intermediate polarization-states of the Fe-FinFET is crucial for the proposed implementation since the computation of dot product I involves the utilization of the stored polarization-states in crossbar I. Therefore, any loss of the programmed polarization-states in crossbar I will significantly affect the classification accuracy. However, FeFETs have shown excellent retention characteristics in MFIS configuration [30], [50], [51]. Also, with advanced techniques such as utilizing dual ferroelectric layer gate stacks [52] and Al:HfO<sub>2</sub> thin films [53], FeFETs in MFMS configurations have also shown good retention characteristics. Therefore, Fe-FinFET cell retention should not degrade the classification accuracy significantly.

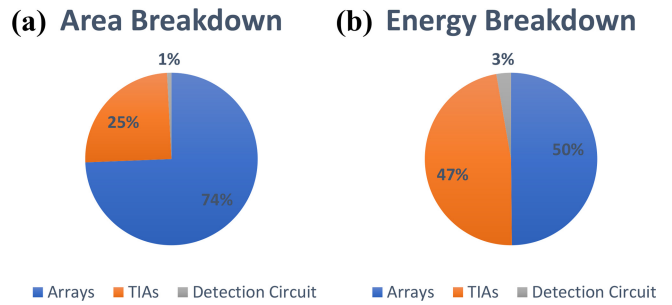


FIGURE 11. (a) Area and (b) energy breakdown of the implementation.

### C. ENERGY

Since we have performed the dot product operation in-situ on a cross-point array of Fe-FinFETs utilizing the physical laws, it is highly energy-efficient. Our comprehensive analysis indicates that the worst-case total energy (including the contribution of the Fe-FinFET arrays and the trans-impedance amplifier and inverter-based outlier detector in the peripheral circuitry) required to classify a given input as an outlier/non-outlier after the application of inputs is only 27.2 pJ. The energy- and area breakdown of the proposed implementation are also shown in Fig. 11, which clearly indicates that the Fe-FinFET arrays and peripheral circuitry nearly take up an equal amount of energy. Since a dedicated CMOS/emerging technology-based hardware implementation of Mahalanobis distance algorithm is still elusive, for efficient benchmarking of our proposed Fe-FinFET-based implementation, we have analyzed the performance of CMOS-based logic gates using FinFETs (at the same technology node [54]) and carbon nanotube (CNT) FETs [55] for realizing the exact Mahalanobis distance algorithm in the digital domain and compared the performance metrics in Table 1. The proposed implementation outperforms the FinFET- and CNTFET-based implementations in terms of area footprint. Furthermore, it also exhibits a higher energy-efficiency as compared to the FinFET-based CMOS implementation. Moreover, it may also be noted that the energy estimates for the CMOS-based digital implementations in Table 1 do not include the data transfer (access) energy, which dominates their energy landscape and remains the bottleneck for energy-efficient circuit design. Since the proposed Fe-FinFET-based implementation also alleviates the need for frequent data transfer, it exhibits an inherent energy benefit. Although we have included the line parasitics in our simulations and energy estimates, they do not contribute significantly due to the small array size in our implementation. However, the contribution of the parasitics may not be non-negligible as the array sizes increase to address complex high-dimensional datasets.

### V. CONCLUSION

In this work, we have proposed an energy-efficient hardware implementation of Mahalanobis distance utilizing Fe-FinFET crossbars and a novel CMOS-inverter-based firing circuit

**TABLE 1. Benchmarking of our implementation.**

	CMOS (14 nm)	CNTFET	Our Work
Energy (pJ)	54.6*	11.21*	27.2
Area ( $\mu\text{m}^2$ )	33.163	19.48	2.3716

\* Excluding access energy

for detecting outliers in data. Although the Mahalanobis distance computed utilizing the proposed implementation shows deviation from the ideal software-calculated values, it shows an excellent accuracy of 94.1% in detecting outliers in the Wisconsin breast cancer dataset. We believe that our proposal will provide an incentive for experimental demonstration of the Mahalanobis distance computing engine, which will not only be useful for outlier detection but also pave the way for efficient implementation of AI/ML tasks, including cluster analysis, image processing, fault identification, etc., in this era of AIoT. Moreover, the proposed hardware implementation can also be extended to other outlier detection problems [56].

## REFERENCES

- [1] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014, doi: [10.1109/ACCESS.2014.2332453](https://doi.org/10.1109/ACCESS.2014.2332453).
- [2] Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, "Health-CPS: Healthcare cyber-physical system assisted by cloud and big data," *IEEE Syst. J.*, vol. 11, no. 1, pp. 88–95, Mar. 2017, doi: [10.1109/JSYST.2015.2460747](https://doi.org/10.1109/JSYST.2015.2460747).
- [3] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison," *IEEE Access*, vol. 6, pp. 3585–3593, 2018, doi: [10.1109/ACCESS.2018.2793265](https://doi.org/10.1109/ACCESS.2018.2793265).
- [4] R. C. Carlos, C. E. Kahn, and S. Halabi, "Data science: Big data, machine learning, and artificial intelligence," *J. Amer. Coll. Radiol.*, vol. 15, no. 3, pp. 497–498, 2018, doi: [10.1016/j.jacr.2018.01.029](https://doi.org/10.1016/j.jacr.2018.01.029).
- [5] H. Yan, J. Wan, C. Zhang, S. Tang, Q. Hua, and Z. Wang, "Industrial big data Analytics for prediction of remaining useful life based on deep learning," *IEEE Access*, vol. 6, pp. 17190–17197, 2018, doi: [10.1109/ACCESS.2018.2809681](https://doi.org/10.1109/ACCESS.2018.2809681).
- [6] D. E. O'Leary, "Artificial intelligence and big data," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 96–99, Mar./Apr. 2013, doi: [10.1109/MIS.2013.39](https://doi.org/10.1109/MIS.2013.39).
- [7] A. Paullada, I. D. Raji, E. M. Bender, E. Denton, and A. Hanna, "Data and its (dis)contents: A survey of dataset development and use in machine learning research," *Patterns*, vol. 2, no. 11, 2021, Art. no. 100336, doi: [j.patter.2021.100336](https://doi.org/10.1016/j.patter.2021.100336).
- [8] W. Liang et al., "Advances, challenges and opportunities in creating data for trustworthy AI," *Nat. Mach. Intell.*, vol. 4, no. 8, pp. 669–677, Aug. 2022, doi: [10.1038/s42256-022-00516-1](https://doi.org/10.1038/s42256-022-00516-1).
- [9] C. C. Aggarwal, *An Introduction to Outlier Analysis*. Cham, Switzerland: Springer Int. Publ., 2017, pp. 1–34.
- [10] P. C. Mahalanobis, "On the generalized distance in statistics," *Proc. Nat. Inst. Sci.*, vol. 2, pp. 49–55, Apr. 1936.
- [11] S. Zeng, X. Wang, X. Duan, S. Zeng, Z. Xiao, and D. Feng, "Kernelized Mahalanobis distance for fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 10, pp. 3103–3117, Oct. 2021, doi: [10.1109/TFUZZ.2020.3012765](https://doi.org/10.1109/TFUZZ.2020.3012765).
- [12] S. Kumar, T. W. S. Chow, and M. Pecht, "Approach to fault identification for electronic products using Mahalanobis distance," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2055–2064, Aug. 2010, doi: [10.1109/TIM.2009.2032884](https://doi.org/10.1109/TIM.2009.2032884).
- [13] F. Guo, W. Susilo, and Y. Mu, "Distance-based encryption: How to embed fuzziness in biometric-based encryption," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 247–257, 2016, doi: [10.1109/TIFS.2015.2489179](https://doi.org/10.1109/TIFS.2015.2489179).
- [14] 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, no. 2, pp. 113–122, Dec. 2019, doi: [10.1109/JXCDC.2019.2923745](https://doi.org/10.1109/JXCDC.2019.2923745).
- [15] T. Gokmen, M. Onen, and W. Haensch, "Training deep convolutional neural networks with resistive cross-point devices," *Front. Neurosci.*, vol. 11, p. 538, Oct. 2017, doi: [10.3389/fnins.2017.00538](https://doi.org/10.3389/fnins.2017.00538).
- [16] S. Park et al., "RRAM-based synapse for neuromorphic system with pattern recognition function," in *Proc. Int. Electron Devices Meeting*, 2012, pp. 10.2.1–10.2.4, doi: [10.1109/IEDM.2012.6479016](https://doi.org/10.1109/IEDM.2012.6479016).
- [17] G. W. Burr et al., "Experimental demonstration and tolerancing of a large-scale neural network (165 000 synapses) using phase-change memory as the synaptic weight element," *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3498–3507, Nov. 2015, doi: [10.1109/TED.2015.2439635](https://doi.org/10.1109/TED.2015.2439635).
- [18] D. Garbin et al., "HfO<sub>2</sub>-based OxRAM devices as synapses for convolutional neural networks," *IEEE Trans. Electron Devices*, vol. 62, no. 8, pp. 2494–2501, Aug. 2015, doi: [10.1109/TED.2015.2440102](https://doi.org/10.1109/TED.2015.2440102).
- [19] Z. Dong et al., "Convolutional neural networks based on RRAM devices for image recognition and online learning tasks," *IEEE Trans. Electron Devices*, vol. 66, no. 1, pp. 793–801, Jan. 2019, doi: [10.1109/TED.2018.2882779](https://doi.org/10.1109/TED.2018.2882779).
- [20] S. Jung et al., "A crossbar array of magnetoresistive memory devices for in-memory computing," *Nature*, vol. 601, no. 7892, pp. 211–216, Jan. 2022, doi: [10.1038/s41586-021-04196-6](https://doi.org/10.1038/s41586-021-04196-6).
- [21] T. Böske, J. Müller, D. Braeuhaus, U. Schroeder, and U. Bottger, "Ferroelectricity in hafnium oxide thin films," *Appl. Phys. Lett.*, vol. 99, pp. 102903–102903, Sep. 2011, doi: [10.1063/1.3634052](https://doi.org/10.1063/1.3634052).
- [22] M. Trentzsch et al., "A 28nm HKMG super low power embedded NVM technology based on ferroelectric FETs," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2016, pp. 11.5.1–11.5.4, doi: [10.1109/IEDM.2016.7838397](https://doi.org/10.1109/IEDM.2016.7838397).
- [23] S. Dinkel et al., "A FeFET based super-low-power ultra-fast embedded NVM technology for 22nm FDSOI and beyond," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2017, pp. 19.7.1–19.7.4, doi: [10.1109/IEDM.2017.8268425](https://doi.org/10.1109/IEDM.2017.8268425).
- [24] H. Mulaosmanovic, E. T. Breyer, T. Mikolajick, and S. Slesazek, "Ferroelectric FETs with 20-nm-thick HfO<sub>2</sub> layer for large memory window and high performance," *IEEE Trans. Electron Devices*, vol. 66, no. 9, pp. 3828–3833, Sep. 2019, doi: [10.1109/TED.2019.2930749](https://doi.org/10.1109/TED.2019.2930749).
- [25] S. Chatterjee, S. Thomann, K. Ni, Y. S. Chauhan, and H. Amrouch, "Comprehensive variability analysis in dual-port FeFET for reliable multi-level-cell storage," *IEEE Trans. Electron Devices*, vol. 69, no. 9, pp. 5316–5323, Sep. 2022, doi: [10.1109/TED.2022.3192808](https://doi.org/10.1109/TED.2022.3192808).
- [26] S. Mueller, S. Slesazek, T. Mikolajick, J. Müller, P. Polakowski, and S. Flachowsky, "Next-generation ferroelectric memories based on FeHfO<sub>2</sub>," in *Proc. Joint IEEE Int. Symp. Appl. Ferroelectr. (ISAF), Int. Symp. Integr. Funct. (ISIF), Piezoelectr. Force Microsc. Workshop (PFM)*, 2015, pp. 233–236, doi: [10.1109/ISAF.2015.7172714](https://doi.org/10.1109/ISAF.2015.7172714).
- [27] M. Seo et al., "First demonstration of a logic-process compatible junctionless ferroelectric FinFET synapse for neuromorphic applications," *IEEE Electron Device Lett.*, vol. 39, no. 9, pp. 1445–1448, Sep. 2018, doi: [10.1109/LED.2018.2852698](https://doi.org/10.1109/LED.2018.2852698).
- [28] M. Rafiq, S. S. Parihar, Y. S. Chauhan, and S. Sahay, "Efficient implementation of max-pooling algorithm exploiting history-effect in ferroelectric-FinFETs," *IEEE Trans. Electron Devices*, vol. 69, no. 11, pp. 6446–6452, Nov. 2022, doi: [10.1109/TED.2022.3207114](https://doi.org/10.1109/TED.2022.3207114).
- [29] J. Hoffman et al., "Ferroelectric field effect transistors for memory applications," *Adv. Mater.*, vol. 22, nos. 26–27, pp. 2957–2961, 2010, doi: [10.1002/adma.200904327](https://doi.org/10.1002/adma.200904327).
- [30] B. Zeng et al., "2-bit/cell operation of Hf<sub>0.5</sub>Zr<sub>0.5</sub>O<sub>2</sub> based FeFET memory devices for NAND applications," *IEEE J. Electron Devices Soc.*, vol. 7, pp. 551–556, May 2019, doi: [10.1109/JEDS.2019.2913426](https://doi.org/10.1109/JEDS.2019.2913426).
- [31] P. Wang et al., "Drain-erase scheme in ferroelectric field effect transistor—Part II: 3-D-NAND architecture for in-memory computing," *IEEE Trans. Electron Devices*, vol. 67, no. 3, pp. 962–967, Mar. 2020, doi: [10.1109/TED.2020.2969383](https://doi.org/10.1109/TED.2020.2969383).
- [32] M. Rafiq, T. Kaur, A. Gaidhane, Y. S. Chauhan, and S. Sahay, "Ferroelectric FET-based time-mode multiply-accumulate accelerator: Design and analysis," *IEEE Trans. Electron Devices*, vol. 70, no. 12, pp. 6613–6621, Dec. 2023, doi: [10.1109/TED.2023.3323261](https://doi.org/10.1109/TED.2023.3323261).



- [33] M. Jain, R. K. Singh, M. Rafiq, and S. Sahay, "Hybrid CMOS-ferroelectric FET-based image sensor with tunable dynamic range," *IEEE Trans. Electron Devices*, vol. 71, no. 1, pp. 624–629, Jan. 2024, doi: [10.1109/TED.2023.3331677](https://doi.org/10.1109/TED.2023.3331677).
- [34] S. Chatterjee, Y. S. Chauhan, and H. Amrouch, "Programmable delay element using dual-port FeFET for post-silicon clock tuning," *IEEE Electron Device Lett.*, vol. 44, no. 11, pp. 1907–1910, Nov. 2023, doi: [10.1109/LED.2023.3317316](https://doi.org/10.1109/LED.2023.3317316).
- [35] M. K. Q. Jooq, M. H. Moaiyeri, A. Al-Shidaifat, and H. Song, "Ultra-efficient and robust auto-nonvolatile Schmitt trigger-based latch design using ferroelectric CNTFET technology," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 5, pp. 1829–1840, May 2022, doi: [10.1109/TUFFC.2022.3158822](https://doi.org/10.1109/TUFFC.2022.3158822).
- [36] M. H. Moaiyeri, M. K. Q. Jooq, A. Al-Shidaifat, and H. Song, "Breaking the limits in ternary logic: An ultra-efficient auto-backup/restore nonvolatile ternary flip-flop using negative capacitance CNTFET technology," *IEEE Access*, vol. 9, pp. 132641–132651, 2021, doi: [10.1109/ACCESS.2021.3114408](https://doi.org/10.1109/ACCESS.2021.3114408).
- [37] Y. Jeong, J. Lee, J. Moon, J. H. Shin, and W. D. Lu, "K-means data clustering with memristor networks," *Nano Lett.*, vol. 18, no. 7, pp. 4447–4453, Jul. 2018, doi: [10.1021/acs.nanolett.8b01526](https://doi.org/10.1021/acs.nanolett.8b01526).
- [38] H. Zhou et al., "Energy-efficient memristive Euclidean distance engine for brain-inspired competitive learning," *Adv. Intell. Syst.*, vol. 3, no. 11, 2021, Art. no. 2100114, doi: [10.1002/aisy.202100114](https://doi.org/10.1002/aisy.202100114).
- [39] A. D. Gaidhane, R. Dangi, S. Sahay, A. Verma, and Y. S. Chauhan, "A computationally efficient compact model for ferroelectric switching with asymmetric nonperiodic input signals," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 42, no. 5, pp. 1634–1642, May 2023, doi: [10.1109/TCAD.2022.3203956](https://doi.org/10.1109/TCAD.2022.3203956).
- [40] A. D. Gaidhane, R. Dangi, S. Sahay, A. Verma, and Y. S. Chauhan, "Verilog-a code of ferroelectric capacitor." 2023. [Online]. Available: <https://home.iitk.ac.in/~chauhan/ferro.va>
- [41] Y. S. Chauhan et al., *FinFET Modeling for IC Simulation and Design: Using the BSIM-CMG Standard*. Cambridge, MA, USA: Academic, 2015.
- [42] K. Ni, M. Jerry, J. A. Smith, and S. Datta, "A circuit compatible accurate compact model for ferroelectric-FETs," in *Proc. IEEE Symp. VLSI Technol.*, 2018, pp. 131–132, doi: [10.1109/VLSIT.2018.8510622](https://doi.org/10.1109/VLSIT.2018.8510622).
- [43] S. De et al., "Ultra-low power robust 3bit/cell Hf<sub>0.5</sub>Zr<sub>0.5</sub>O<sub>2</sub> ferroelectric FinFET with high endurance for advanced computing-in-memory technology," in *Proc. Symp. VLSI Technol.*, 2021, pp. 1–2. [Online]. Available: <https://ieeexplore.ieee.org/document/9508674>.
- [44] S. Oh et al., "HfZrO<sub>x</sub>-based ferroelectric synapse device with 32 levels of conductance states for neuromorphic applications," *IEEE Electron Device Lett.*, vol. 38, no. 6, pp. 732–735, Jun. 2017, doi: [10.1109/LED.2017.2698083](https://doi.org/10.1109/LED.2017.2698083).
- [45] 2017, D. Dua and C. Graff, "UCI machine learning repository," Dataset, UCI Machine Learning Repository. [Online]. Available: <http://archive.ics.uci.edu/ml>
- [46] H. Kim, M. R. Mahmoodi, H. Nili, and D. B. Strukov, "4K-memristor analog-grade passive crossbar circuit," *Nat. Commun.*, vol. 12, no. 1, p. 5198, Aug. 2021, doi: [10.1038/s41467-021-25455-0](https://doi.org/10.1038/s41467-021-25455-0).
- [47] M. D. Giles et al., "High sigma measurement of random threshold voltage variation in 14nm logic FinFET technology," in *Proc. Symp. VLSI Technol.*, 2015, pp. T150–T151, doi: [10.1109/VLSIT.2015.7223657](https://doi.org/10.1109/VLSIT.2015.7223657).
- [48] P. Oldiges et al., "Critical analysis of 14nm device options," in *Proc. Int. Conf. Simul. Semicond. Processes Devices*, 2011, pp. 5–8, doi: [10.1109/SISPAD.2011.6035034](https://doi.org/10.1109/SISPAD.2011.6035034).
- [49] A. Gupta, N. Chauhan, O. Prakash, and H. Amrouch, "Variability effects in FinFET transistors and emerging NC-FinFET," in *Proc. Int. Conf. IC Design Technol. (ICICDT)*, 2021, pp. 1–4, doi: [10.1109/ICICDT51558.2021.9626531](https://doi.org/10.1109/ICICDT51558.2021.9626531).
- [50] T. Ali et al., "High endurance ferroelectric hafnium oxide-based FeFET memory without retention penalty," *IEEE Trans. Electron Devices*, vol. 65, no. 9, pp. 3769–3774, Sep. 2018, doi: [10.1109/TED.2018.2856818](https://doi.org/10.1109/TED.2018.2856818).
- [51] H. Mulaosmanovic et al., "Evidence of single domain switching in hafnium oxide based FeFETs: Enabler for multi-level FeFET memory cells," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2015, pp. 26.8.1–26.8.3, doi: [10.1109/IEDM.2015.7409777](https://doi.org/10.1109/IEDM.2015.7409777).
- [52] T. Ali et al., "A novel dual ferroelectric layer based MFMFIS FeFET with optimal stack tuning toward low power and high-speed NVM for neuromorphic applications," in *Proc. IEEE Symp. VLSI Technol.*, 2020, pp. 1–2, doi: [10.1109/VLSITechnology18217.2020.9265111](https://doi.org/10.1109/VLSITechnology18217.2020.9265111).
- [53] S.-J. Yoon, D.-H. Min, S.-E. Moon, K. S. Park, J. I. Won, and S.-M. Yoon, "Improvement in long-term and high-temperature retention stability of ferroelectric field-effect memory transistors with metal-ferroelectric-metal-insulator-semiconductor gate-stacks using al-doped HfO<sub>2</sub> thin films," *IEEE Trans. Electron Devices*, vol. 67, no. 2, pp. 499–504, Feb. 2020, doi: [10.1109/TED.2019.2961117](https://doi.org/10.1109/TED.2019.2961117).
- [54] Q. Xie et al., "5nm FinFET standard cell library optimization and circuit synthesis in near-and super-threshold voltage regimes," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, 2014, pp. 424–429, doi: [10.1109/ICICDT51558.2021.9626531](https://doi.org/10.1109/ICICDT51558.2021.9626531) [10.1109/ISVLSI.2014.101](https://doi.org/10.1109/ISVLSI.2014.101).
- [55] J. Deng and H.-S. P. Wong, "A compact SPICE model for carbon-nanotube field-effect transistors including nonidealities and its application—Part I: Model of the intrinsic channel region," *IEEE Trans. Electron Devices*, vol. 54, no. 12, pp. 3186–3194, Dec. 2007, doi: [10.1109/TED.2007.909030](https://doi.org/10.1109/TED.2007.909030).
- [56] K. Ni et al., "In-memory computing primitive for sensor data fusion in 28 nm HKMG FeFET technology," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2018, pp. 16.1.1–16.1.4, doi: [10.1109/IEDM.2018.8614527](https://doi.org/10.1109/IEDM.2018.8614527).