scientific reports

Check for updates

OPEN

A robust deep learning attack immune MRAM-based physical unclonable function

Mohammad Javad Adel[1], Mohammad Hadi Rezayati[1], Mohammad Hossein Moaiyeri[1]✉, Abdolah Amirany[2] & Kian Jafari[3,4]

The ubiquitous presence of electronic devices demands robust hardware security mechanisms to safeguard sensitive information from threats. This paper presents a physical unclonable function (PUF) circuit based on magnetoresistive random access memory (MRAM). The circuit utilizes inherent characteristics arising from fabrication variations, specifically magnetic tunnel junction (MTJ) cell resistance, to produce corresponding outputs for applied challenges. In contrast to Arbiter PUF, the proposed effectively satisfies the strict avalanche criterion (SAC). Additionally, the grid-like structure of the proposed circuit preserves its resistance against machine learning-based modeling attacks. Various machine learning (ML) attacks employing multilayer perceptron (MLP), linear regression (LR), and support vector machine (SVM) networks are simulated for two-array and four-array architectures. The MLP-attack prediction accuracy was 53.61% for a two-array circuit and 49.87% for a four-array circuit, showcasing robust performance even under the worst-case process variations. In addition, deep learning-based modeling attacks in considerable high dimensions utilizing multiple networks such as convolutional neural network (CNN), recurrent neural network (RNN), MLP, and Larq are used with the accuracy of 50.31%, 50.25%, 50.31%, and 50.31%, respectively. The efficiency of the proposed circuit at the layout level is also investigated for simplified two-array architecture. The simulation results indicate that the proposed circuit offers intra and inter-hamming distance (HD) with a mean of 0.98% and 49.96%, respectively, and a mean diffuseness of 49.09%.

Keywords Hardware security primitives, Physical unclonable function (PUF), Magnetic tunnel junction (MTJ), Emerging technologies, Machine learning (ML)-based modeling attack, Deep learning (DL)-based modeling attack

Physical unclonable functions (PUFs) are designed to enhance hardware security by using the unique physical properties of electronic components[1–3]. These functions generate distinct and unpredictable responses, known as challenge-response pairs (CRP)[4], which serve as robust cryptographic keys[5]. PUFs play a crucial role in safeguarding electronic circuits against various security threats[6], particularly those arising from machine learning (ML)-based modeling attacks[7,8]. In the landscape of hardware security, PUFs have become integral due to their ability to withstand sophisticated attacks. However, they are not immune to challenges, with modeling attacks employing ML and deep learning (DL) posing a significant risk[9,10]. ML and DL attacks can exploit vulnerabilities in PUFs by analyzing large datasets to identify patterns and predict and generate accurate models that emulate the behavior of original PUFs. An attacker can create a mathematical simulation of the desired PUF by building an ML model from PUF, which can be trained to achieve high accuracy in prediction after obtaining a sufficient set of CRP[11]. For instance, the Arbiter PUF was initially proposed as a simple, structured, functioning secret key. However, the Arbiter PUF is vulnerable to adversaries who can access a PUF sample to CRPs and attempt to construct a mathematical framework to predict PUF response with high accuracy[12,13]. These attacks aim to replicate the behavior of PUFs, compromising their security. Therefore, there is a need for innovative approaches to construct PUFs that are resilient against such threats.

Traditional PUFs, often based on complementary metal–oxide–semiconductor (CMOS) technology, have demonstrated effectiveness in hardware security applications[14,15]. However, due to specific vulnerabilities and limitations in CMOS-based designs[16], such as high energy consumption, substantial area overhead[17], design

[1]Faculty of Electrical Engineering, Shahid Beheshti University, Tehran 1983969411, Iran. [2]Department of Electrical and Computer Engineering, The George Washington University, Washington, DC, USA. [3]Institut Interdisciplinaire d'Innovation Technologique (3IT), Université de Sherbrooke, Sherbrooke, QC, Canada. [4]Faculty of Engineering, Université de Sherbrooke, 2500 Boul. de l'Université, Sherbrooke, QC, Canada. ✉email: h_moaiyeri@sbu.ac.ir

complexity, and susceptibility to power overhead and environmental fluctuations[18–21], researchers are increasingly exploring emerging technologies to bolster PUF security. CMOS-based PUFs suffer from two major issues: high bit error rate (BER) and weak uniqueness. Most CMOS-based PUFs employ error correction codes (ECC) to improve results uncertainty; however, strong ECC results in significant area overhead and high energy consumption. Post-silicon technologies like spintronics are emerging as promising alternatives to address these challenges. Spintronic devices, particularly magnetic tunnel junctions (MTJ), offer significant advantages over traditional CMOS technology, including low-power consumption, non-volatility, and high endurance. These attributes make them highly resilient against modeling-based attacks and better at achieving ideal uniqueness, approaching the target value of 50%[22]. The pursuit of these emerging technologies is driven by their potential to address the shortcomings of CMOS-based[23] PUFs and enhance overall hardware security[24,25].

MTJ-based PUFs, utilizing magnetic materials, exhibit unique characteristics that make them challenging to model or predict. This marks a significant advancement in PUF design, offering heightened security for embedded systems. Exploring emerging technologies, particularly MTJs, signifies a shift towards more resilient PUFs[26]. Associating PUF models with adversarial training processes enhances PUF circuit security[27]. By training PUF models through adversarial attacks, circuits become more resilient against malicious input. For instance, training PUF models to detect fraudulent patterns enables them to counteract attacks like tampering with encryption keys. This adaptive training transforms PUF circuits into highly resilient security systems.

Hardware obfuscation alters hardware to hinder unauthorized analysis and comprehension, safeguarding sensitive information. The goal is to obscure the analysis process at the hardware level to protect confidential data like circuit designs or encryption algorithms. CMOS-based methods face challenges, such as accessibility of device characteristics, deterministic properties, and conventional architectural designs, leading to increased power consumption and area overhead. Alternative approaches beyond CMOS devices offer heightened security with minimal overheads. STT-MRAM presents a promising hardware obfuscation solution because it can generate random responses and prevent feature extraction. Increasing circuit complexity enhances security and prevents unauthorized analyses[22].

Numerous studies have explored PUFs' application in hardware security, significantly advancing our understanding of their potential and limitations. However, challenges persist. In the pursuit of bolstering hardware security, multiple research papers have proposed distinct implementations of PUFs. One work presented an MRAM-based PUF (MPUF) to enhance resilience against ML attacks. Nevertheless, there are opportunities for improving energy consumption and the utilization of space and increasing the CRP area[28]. Another explored a spintronics memory PUF based on STT-MRAM, emphasizing its robustness against cloning and counterfeiting; however, this circuit incurs a significant area overhead[29]. A different study also focused on MRAM PUF, utilizing geometric variations and energy tilt for heightened security and efficiency. Nevertheless, this approach requires further investigation in terms of reliability[24].

Researchers introduce a memristive crossbar PUF in another work, emphasizing ML attack resilience through circuit design enhancements. It appears that XORing the responses as post-processing has been utilized to enhance the efficiency of the presented PUF [7]. Finally, a subthreshold current array PUF in 130 nm CMOS technology demonstrates remarkable resilience to ML attacks while maintaining predictability at negligible levels. Moreover, the average bit error rate in this study is reported to be 9%, with a reduction of approximately 10% in CRP space to improve it. Additionally, the circuit's resistance against ML-based attacks slightly deviates from the expected value[30]. Despite their contributions, each work has limitations, including specific vulnerabilities, authentication overhead, power consumption, and delay considerations.

This paper proposes a strong MPUF operating based on the resistance values derived from the manufacturing process variations of MTJ cells. The performance of this circuit is evaluated using various criteria, such as the NIST statistical test suit, which is specific for uniformity. Also, under different environmental conditions, such as changes in source voltage and temperature, the presented circuit is tested in the fabrication corners, and its post-layout simulation is also considered. The main contributions of this paper can be expressed as follows:

- Proposing a grid-like structure that provides a high resilience against forgery attacks and modeling based on ML and DL algorithms
- Offering advantages regarding the occupied area
- Utilization of fewer transistors than similar work
- Being superior in power consumption compared to previous counterparts
- Providing a more vast CRP space state than previous work
- Performing appropriately in evaluations and various metrics, especially in intra-HD

Preliminaries

STT-MRAM technology

Spin-transfer torque magnetic random access memory (STT-MRAM) stands out as a promising nonvolatile memory technology that uses the principles of spintronics[31,32]. In STT-MRAM, information is stored and retrieved by manipulating the orientation of magnetic moments using spin-polarized currents. STT-MRAM devices typically consist of an MTJ. An MTJ comprises two ferromagnetic layers separated by a thin insulating barrier[33,34]. The free layer's magnetic moment can be manipulated using spin-polarized electrons generated by passing a current through the barrier. The relative alignment of magnetic moments in the free and reference layers determines the overall resistance of the MTJ. The efficiency of STT-MRAM relies on the tunnel magnetoresistance ratio (TMR)[35,36]. The TMR ratio is expressed as

$$TMR = \frac{R_{AP} - R_P}{R_P} \times 100\% \quad (1)$$

where R_{AP} is the resistance of MTJ when the magnetizations of the two layers are antiparallel, and R_P is the resistance of MTJ when the magnetizations of the two layers are parallel. Figure 1 shows the two operational modes of the MTJs[28,37]. A high TMR indicates a significant difference in resistance between the parallel (P) and antiparallel (AP) states[38,39]. A high TMR ratio is crucial for reliable and efficient operation[40,41].

STT-MRAM offers several advantages, including low power consumption, excellent scalability[42,43] nearly zero leakage power, and fast access speed[44,45]. Its nonvolatile nature ensures data retention even during sudden power outages, making it suitable for various memory-intensive applications. In the context of hardware security, STT-MRAM presents an intriguing option for implementing PUFs. The inherent variability in resistance due to manufacturing process variations[46] and thermal fluctuations within the MTJ can be used to generate unique and unpredictable responses.

Previous work and challenges

In[28], a strong PUF based on STT-MRAM and the intrinsic properties and process variations present in MTJs was proposed. Introducing an array selection circuit (ASC) to enhance nonlinear characteristics also increases resistance against machine learning-based attacks. The design proposed in[28] demonstrates good performance, with MTJ cells effectively interacting with 28 nm CMOS technology in the evaluated metrics. However, there is potential for reducing the number of transistors and, consequently, reducing area and energy consumption while increasing CRP space.

An STT-MRAM is utilized as a PUF for secure and anti-counterfeit storage, as proposed in[29]. This work proposes a temper-resilient solution that remains resistant to tampering with a detection rate of 100%. The authors of[29] study the advantage of STT-MRAM over attacks such as threat models and cloning attacks on SRAM PUFs by examining major attacks on CMOS circuits. However, there is a significant area overhead in the examined circuit.

An MRAM-based PUFs and CMOS integrated circuits design in[24] has shown notable advantages, such as its high entropy, a crucial feature for system security, and a smaller footprint. Furthermore, it has demonstrated performance in terms of area and power consumption. However, there may be challenges in terms of initial setup and complexity of fabrication. Additionally, evaluating PUF responses may require specific hardware and protocols, leading to increased authentication overhead in terms of delay and power consumption.

In[7], a memristive crossbar-based PUF proposed, attempting to increase its resistance against machine learning-based attacks by XORing response bits and swapping columns, shows impressive entropy levels. Nevertheless, the reliability and energy consumption of the proposed circuit in the[7] require improvement.

A strong PUF based on subthreshold current is proposed in[30], formed by an array of two-dimensional cells capable of providing 2^{65} challenge-response pairs, leading to high reliability. However, the bit error rate is generally high [30]. In this case, the BER has been significantly reduced by using a calibration-based CRP filtering, with 10% CRP loss. In addition, the proposed design in[30] is susceptible to some ML-based attacks.

The proposed ML and DL attack immune MPUF

The architecture of the proposed MPUF

Figure 2 illustrates the proposed MPUF circuit, including signal decoders for discharge path selection (A, B) for odd arrays, and signal decoders for discharge path selection (D, F) for even arrays, as well as signal decoders for end-node determination (C, S) and (V, Z). An ASC circuit is also placed to select among arrays with control signals P and Q. By the dimensions and topology of the circuit and the available signals, using Algorithm 1, the necessary set of challenges to apply to the circuit and extract responses for use in analyses and evaluations is provided.

A pre-charged sense amplifier (PCSA) is also embedded in the proposed MPUF circuit. The PCSA circuit measures the predetermined path resistance value and compares two selected arrays in a challenge, determining the output. To prevent the input voltage offset effect in PCSA, two transistors controlled by the clock signal are connected from the voltage source to the input of PCSA (Fig. 6)[47]. The ASC circuit enhances resilience against modeling CRP attacks due to its nonlinear properties and ability to choose from multiple arrays. Moreover, the grid-like proposed circuit architecture offers greater complexity, higher resilience against attacks, and fewer

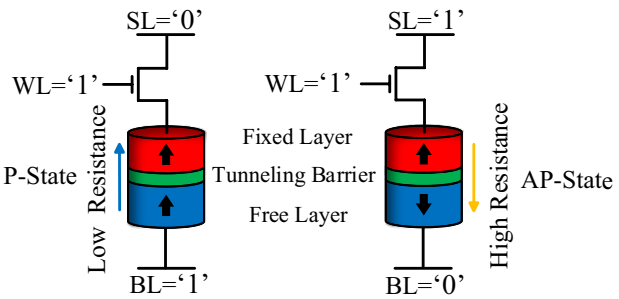

Fig. 1. 1MTJ/1T's structure and its switching mechanism.

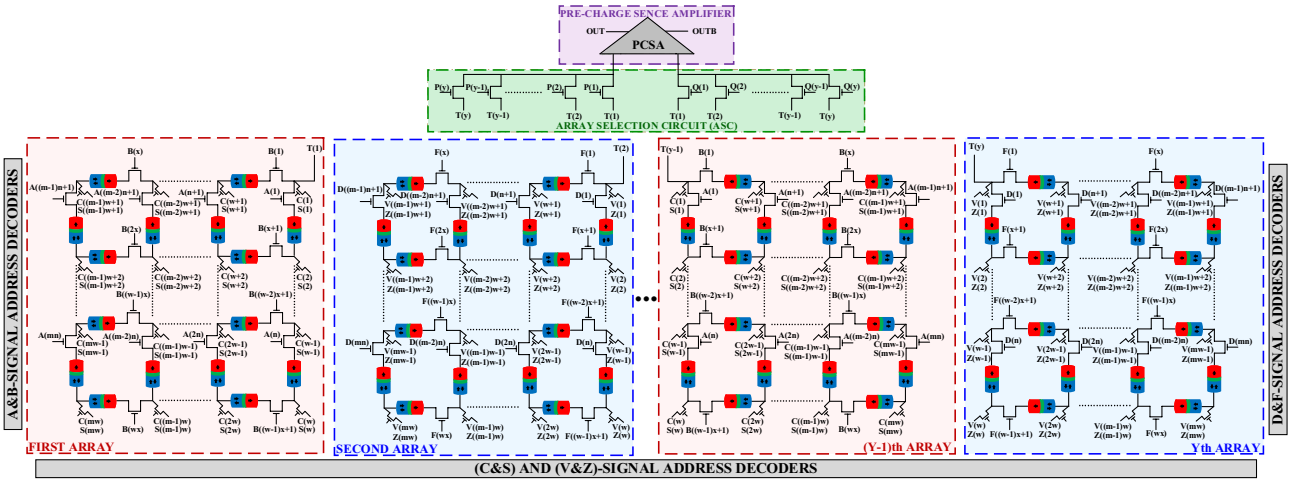

Fig. 2. Schematic of the proposed MPUF.

transistors, resulting in reduced area. In general, the number of MTJ cells in each array can be calculated from the following equation:

$$(x+1)\times n+(n+1)\times x \tag{2}$$

where, x represents cells in length and n represents cells in width. For an array with the dimensions of 3×2, by the way of example, the number of MTJ cells is 17. The critical point is that transistors with gate signals C(1) and V(1) do not participate in the set of PUF challenges.

The layout of the proposed circuit, designed using the 7 nm FinFET technology design kit[48], for the two arrays, each containing 60 MRAM cells, is shown in Fig. 3. Two transistors associated with the Clk signal are embedded after the end-node path transistors.

```
START
  INPUT number_of_challenges
  INPUT circuit_dimension
  challenges_created = 0
  challenges_array = []
  WHILE challenges_created < number_of_challenges DO
    (A,B) ← generate_random_bitstream(circuit_dimension)
    (D,F) ← generate_random_bitstream(circuit_dimension)
    IF ASC_exist THEN
      (P,Q) ← Select_two_Arrays()
    ENDIF
    IF is_suitable_challenge((A,B),(D,F)) THEN
      IF is_unique( (A,B),(D,F),challenges_array ) THEN
        (C,V) ← every end-node could be on the discharge path
        ADD [A,B,D,F,P,Q,C,V] challenges to challenges_array
        challenges_created ← challenges_created + the number
                             of challenges that have just created
      ENDIF
    ENDIF
  ENDWHILE
  OUTPUT (generate HSPICE signal codes (.sp files) based on
          created challenges)
END
```

Algorithm 1. Creating challenges and signal code files.

The proposed circuit's grid-like architecture encompasses many series resistance paths, introducing a specific complexity. Additionally, including sub-paths not connected to the main path has increased the circuit's resilience against modeling-based attacks. Furthermore, using ASC also provides another effective means of maintaining circuit flexibility in using multiple arrays and contributes to the increased complexity of the proposed circuit. Moreover, considering the circuit's topology, employing fewer transistors is possible.

Operation of the Proposed MPUF

An example of a four-array circuit, with each array having dimensions of 3×2, is illustrated in Fig. 4. In a challenge involving five MTJ cells from the first array and six MTJ cells from the fourth array for comparison by the PSCA, control signals B(1), B(6), B(7), A(4), and A(5) from the first array, as well as F(1), F(4), F(7), D(3), D(4), and D(5) from the fourth array, are enabled. Consequently, MTJ cells on either side are selected and arranged in

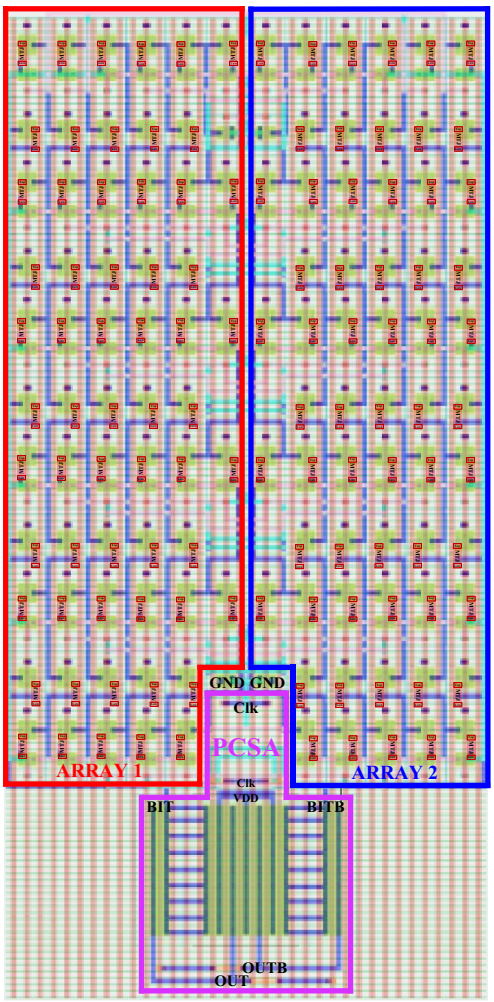

Fig. 3. Layout of the 2-array proposed MPUF with 60 cells in each array without ASC.

series. Additionally, transistors with their gate connected to C(11) in the first array and their gate connected to V(10) in the fourth array are turned on as end-node leading to the ground [using S(11) and Z(10)], completing the discharge path. As depicted in Fig. 4, signals B(7) in the first array and F(7) and D(3) in the fourth array are, although '1' are considered deviated inputs because they do not reside in the discharge path. Each MTJ cell, whether in a parallel or antiparallel state at the moment of challenge, is considered, and the basis for comparison between the two arrays is the resistance value of each MTJ, resulting from fabrication process variations, which are engaged in the challenge.

It is noteworthy that during the writing phase, the clock (Clk) signal is '0', and the state of each MTJ cell is adjustable. The output is not considered a valid response when the clock signal is zero. Therefore, using control signals, the state of MTJ cells can be set to either the parallel or antiparallel state. However, the worst-case scenario is where all states of the MTJs are identical. The proposed circuit is simulated under conditions where all MTJs have the same state (all are parallel and antiparallel). The results indicate that the proposed circuit remains resilient to the attacks mentioned in these scenarios.

On the contrary, when the Clk signal is '1', the read operation is performed, and the response is obtained from the circuit. As stated in[28], the ASC enhances nonlinearity and increases the CRP state space, thereby improving resilience against attacks like ML and DL modeling. Furthermore, for this proposed circuit, the presence of the ASC in this configuration leads to the subdivision of arrays into smaller dimensions, increasing the CRP space. In this case, using ASC, the first array is selected by signal P(1), and the fourth array is selected by signal Q(4). Subsequently, they will participate in the comparison operation. Notably, no signals P(i) and Q(i) with the same index will be concurrently high to prevent connecting a single array to both PCSA inputs. Once the arrays are identified, MTJ cells selected by the challenge are placed at the two PCSA inputs, and their resistance is compared. For instance, if array T(1) has a lower resistance value, the OUT terminal is '1'; if array T(4) has a lower resistance value, the OUTB terminal discharges to '0'.

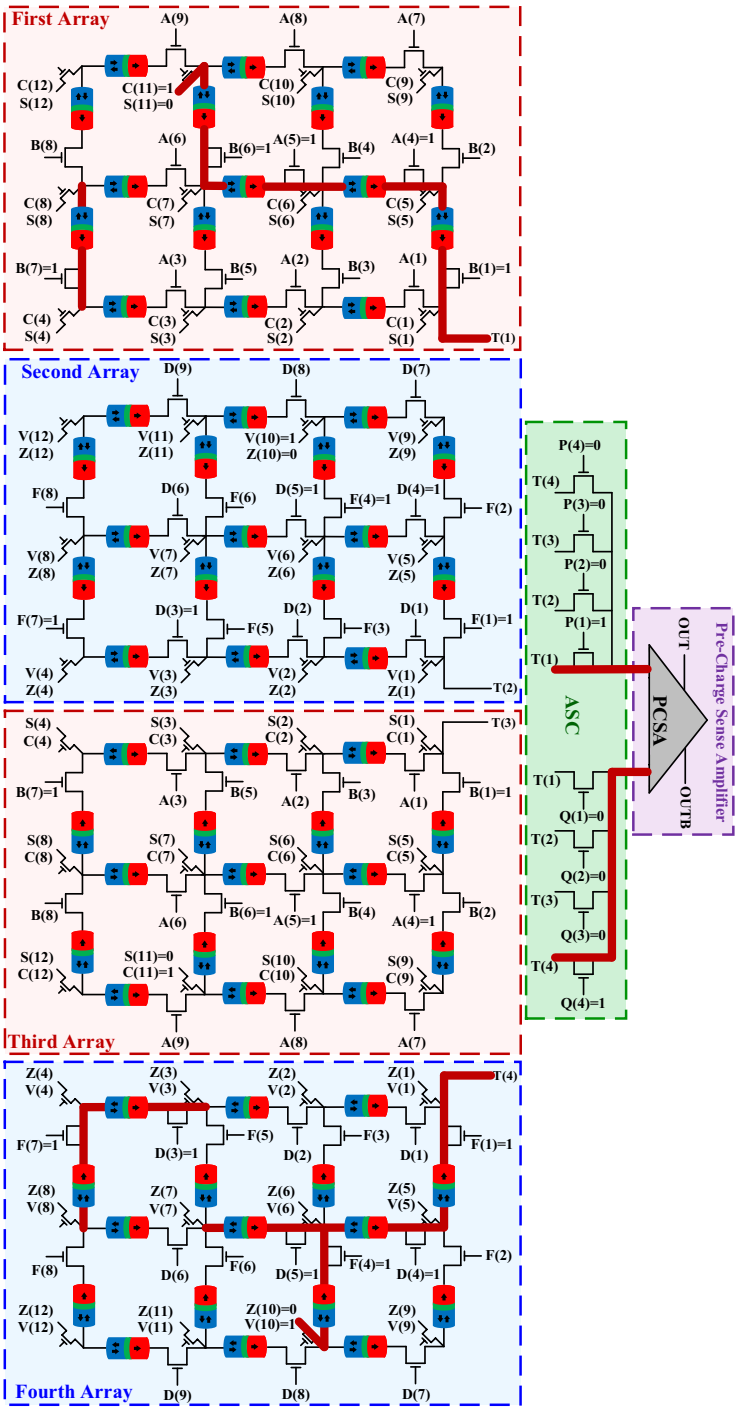

Fig. 4. Sample operation of the proposed 4-array MPUF with ASC.

Simulation results

The performance of the proposed circuit has been evaluated using the experimentally validated MTJ model presented in[49] and the ASAP 7 nm FinFET technology design kit presented in[48]. The specifications and defined parameters are given in Table 1. The post-layout circuit simulations are conducted using Cadence Virtuoso and HSPICE tools. The Python programming language is also utilized to generate challenges and the code related to the circuit.

Function Simulation

Figure 5 depicts the timing diagram for the elementary 1×1 dual-array proposed MPUF. The structure of this array is shown in Fig. 6, comprising four MRAM cells in each array. When the Clk signal is '0', the circuit's outputs, OUT and OUTB, are '1'. Challenges are applied when the Clk signal transitions to '1', and responses are

Symbol	Description	Value
MTJ		
TMR	Tunnel magnetoresistance ratio	200%
t_b	Thickness of the oxide barrier	0.85 nm
t_{sl}	Thickness of the free layer	1.3 nm
d	Diameter of the MTJ	64 nm
RA	Resistance area product	10 Ω.μm^2
Vh	Voltage bias when the real TMR is 0.5 × TMR_0	0.5 V
φ	Energy barrier height for MgO	0.4 eV
H_k	Effective anisotropy field	1433 A m^{-1}
M_s	Saturation magnetization	15,800 A m^{-1}
α	Magnetic damping constant	0.027
FinFET		
T_{fin}	Fin thickness	7 nm
H_{fin}	Fin height	32 nm
P_{fin}	Fin pitch	27 nm
L	Gate length	21 nm
T_{oxp}	Oxide thickness	2.1 nm
EOT	Equivalent oxide thickness	1 nm

Table 1. Device parameters.

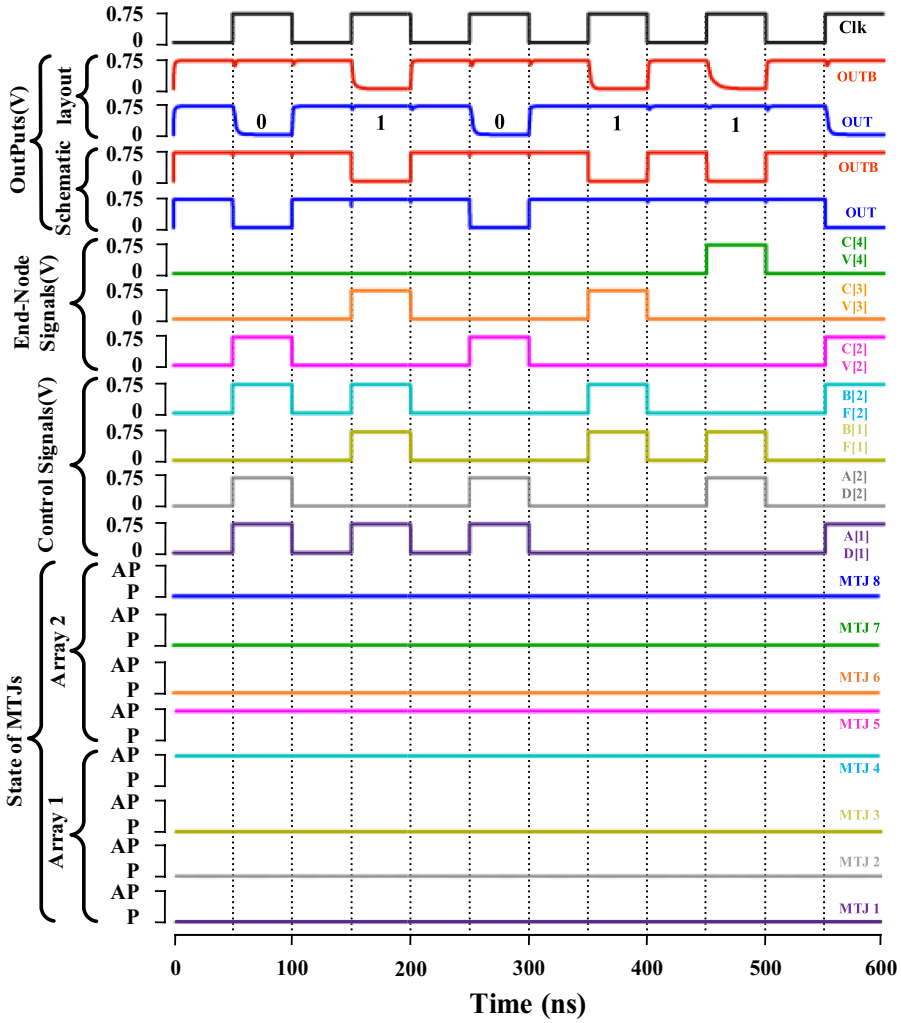

Fig. 5. Timing Diagram of the proposed MPUF.

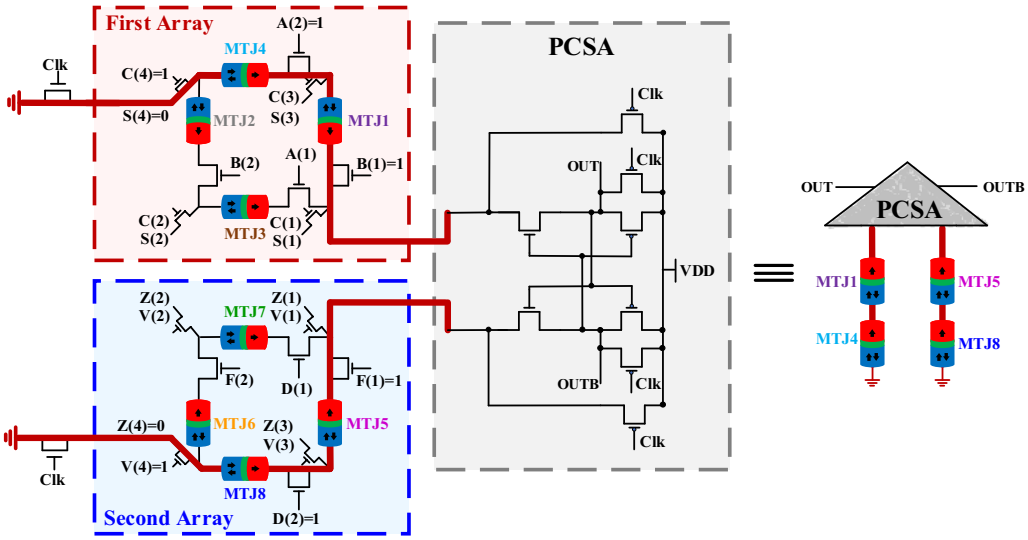

Fig. 6. Schematic of Simplified 2-array MPUF without ASC during the fifth Clk signal pulse.

obtained from the circuit as outputs. For instance, during the fifth Clk signal pulse, inputs A(2) and B(1) with end-node signal C(4) from the first array and D(2) and F(1) with end-node signal V(4) from the second array. The discharge path forms through A(2) and B(1) in the first array and D(2) and F(1) in the second array. In this example, the MTJ cells named MTJ1 and MTJ8 were in a parallel state, while MTJ4 and MTJ5 cells were in an antiparallel state. The resistance values of the corresponding cells in these two paths ((MTJ1, MTJ4) and (MTJ5, MTJ8)) will be of importance. Since the path connected to the second array has a lower resistance than the cells connected to the first, the OUTB of the second array will discharge.

Performance Evaluation

Simulations have been conducted on multiple PUF chips with different process variations for a more accurate evaluation.

Various metrics such as reliability, uniqueness, diffuseness, and uniformity are calculated to assess the proposed circuit's performance. These metrics are explained as follows:

Reliability

Reliability evaluates the circuit's performance and efficiency in adverse environmental conditions, including voltage and temperature variations. For this purpose, the intra-hamming distance (HD) is utilized. The ideal value for the difference between responses is 0%[50]. Equation (3) is used for measuring this metric[51].

$$HD_{\text{intra}} = \frac{1}{m}\sum_{i=1}^{m}\frac{HD(R_0(x), R_i(x))}{n} \times 100\% \tag{3}$$

The HD function calculates the intra-hamming distance between the reference response ($R_0(x)$) and the response at a different condition ($R_i(x)$) for the same challenge of x.

Bit Error Rate (BER) is another reliability metric that quantifies the number of response changes due to environmental conditions[8]. BER is calculated using (4).

$$BER = \frac{1}{m}\sum_{t=1}^{m}\frac{HD(R_i, R'_{i,t})}{n} \times 100\% \tag{4}$$

R_i represents the reference response at nominal conditions, and R'i,t is the t-th response extracted at different conditions using an n-bit response from m samples.

Uniqueness

PUFs are generally recognized as fingerprint-like entities in identity verification tasks. The capability of a PUF to generate unique CRPs is assessed using the uniqueness metric. The inter-HD typically measures uniqueness with an ideal value of 50%[50]. Uniqueness is calculated using (5).

$$HD_{\text{inter}} = \frac{2}{d(d-1)}\sum_{i=1}^{d-1}\sum_{j=i+1}^{d}\frac{HD(R(x), R'(x))}{n} \times 100\% \tag{5}$$

R(x) and R'(x) are n-bit responses from two among *d* PUF instances using the same challenge of x. If multiple PUFs exist, and the same set of challenges is applied to them, the average HD between their responses should ideally be 50%.

Diffuseness
Gauges the variability in responses when different challenges are applied to the same PUF. The diffuseness metric is calculated by determining the mean HD across all possible responses generated by the PUF. A random subset of responses is assessed in practical scenarios if the total number of CRPs is extensive. In short, the diffuseness evaluates differences between responses while different challenges are applied to the same PUF. The ideal diffuseness value is half of the length of the responses, ideally reaching 50%[18]. Diffuseness quantifies the information richness derived from a PUF, indicating the number of distinct identifiers (IDs) the PUF can produce[18]. Diffuseness is calculated using (6).

$$\text{Diffuseness} = \frac{2}{d(d-1)} \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \frac{HD(R_i, R_j)}{n} \times 100\% \tag{6}$$

In (6), R_i and R_j ($i \neq j$) represent n-bit responses for two among d distinct PUFs.

Uniformity
The randomness uniformity criterion assesses the stochastic nature of the output of a PUF circuit by examining the balance and equilibrium between '0' and '1' states[52].

Simulations have been conducted on multiple PUF chips with different process variations for a more accurate evaluation. Monte Carlo simulations were conducted to assess the impact of process variations. Gaussian distribution and variations at the $\pm 3\sigma$ level were considered for the MTJ and FinFET critical device parameters. For the MTJs, 10% variations in the TMR ratio, 15% variation in the resistance-area product (R_{AP}), 5% variation in the barrier thickness (t_b) and the thickness of the free layer (t_{sl}), and 15% for the surface area were considered[23]. Furthermore, for the FinFETs, 10% variations have been considered for the gate length (L_g), fin height (H_{fin}), fin thickness (T_{fin}), and gate oxide thickness (T_{ox}) parameters[35].

Figure 7 shows the above metrics for the proposed two-array circuit, each containing 60 MTJ cells. Figure 7a shows the intra-HD, with an average response difference of 0.98% and a standard deviation of 0.56% for 512 challenges applied at different temperatures ranging from – 25 to 100 °C and supply voltages ranging from 0.65 to 0.85 V for 100 different PUF devices. Figure 7a also displays the inter-HD for 40 chips with proposed PUF for a 1024-challenge set. These measurements were taken at the supply voltage of 0.75V and temperature of 27 °C. Figure 7a shows that the average inter-HD is 49.96%, with a standard deviation of 7.40%. Considering how close these values are to the ideal value and their low standard deviation, Fig. 7a firmly indicates the high reliability of the proposed MPUF.

In Fig. 7b, the BER is displayed in different temperatures for the supply voltage of 0.75 V. The worst-mean value is 2.52%, occurring at – 25 °C. Figure 7.c also shows the BER for supply voltage variations at a temperature of 27 °C. The highest mean value is 1.99%, observed at a supply voltage of 0.65 V. Figure 7b and c together indicate the reliable performance of the proposed MPUF in different temperatures and supply voltage. These Figures are the result of simulations on 200 different PUF chips.

Finally, Fig. 7d depicts the diffuseness distribution of the proposed MPUF. To evaluate diffuseness, 200 sets of challenges, each set containing 128-bit output, have been applied to a PUF, and the HDs of each set of responses have been calculated. The average diffuseness is 49.09%, with a standard deviation of 4.39%. This result indicates that each PUF circuit constructed will produce different outputs than other PUF circuits.

In addition, Table 2 investigates the uniformity of the proposed MPUF employing the National Institute of Standards and Technology (NIST) statistical test suit. The proposed two-array circuit generated responses for the NIST statistical test. Each array contains 60 MTJ cells. Test conducted over 80 million-bit. A test with a portion exceeding 0.96 and a *p*-value greater than 0.01 is considered successful[33]. This table indicates that the proposed circuit is highly random and unpredictable. Notably, the standard deviation of energy consumption in the presence of process variations is 3.2%, with a mean value of 9.49 pj/bit.

It is crucial for a PUF circuit used in security applications to operate significantly reliably and randomly. Based on evaluation criteria, the ideal scenario is for it to be completely resilient to environmental changes, with the responses of each PUF chip being unique and specific to that chip. Based on the results obtained, the circuit presented in this paper has achieved these objectives satisfactorily.

ML-Based modeling for attack simulation

Machine learning (ML)-based attacks, which predict responses to previously unseen challenges, have increasingly threatened various types of PUFs. These attacks do not directly map the transformation between challenges and responses; instead, they predict the outcome of this transformation after learning from a set of CRPs collected from a specific PUF[53]. ML-based modeling, including techniques like logistic regression (LR), support vector machine (SVM), and multilayer perceptron (MLP), has proven effective against conventional PUFs, compromising their robustness. In these attacks, an adversary first acquires a small set of CRPs and constructs a model of the PUF's characteristics. They then attempt to generate additional unknown CRPs with high accuracy[28]. Consequently, despite the initial assumption that PUFs are unpredictable and irreproducible, ML-based modeling attacks can undermine the security of PUFs by enabling identity forgery and application falsification.

ML algorithms consider PUF outputs a classification problem and model it using supervised learning classification. In this paper, similar to[28], the resilience of the proposed MPUF circuit, with a CRP space of 25,000

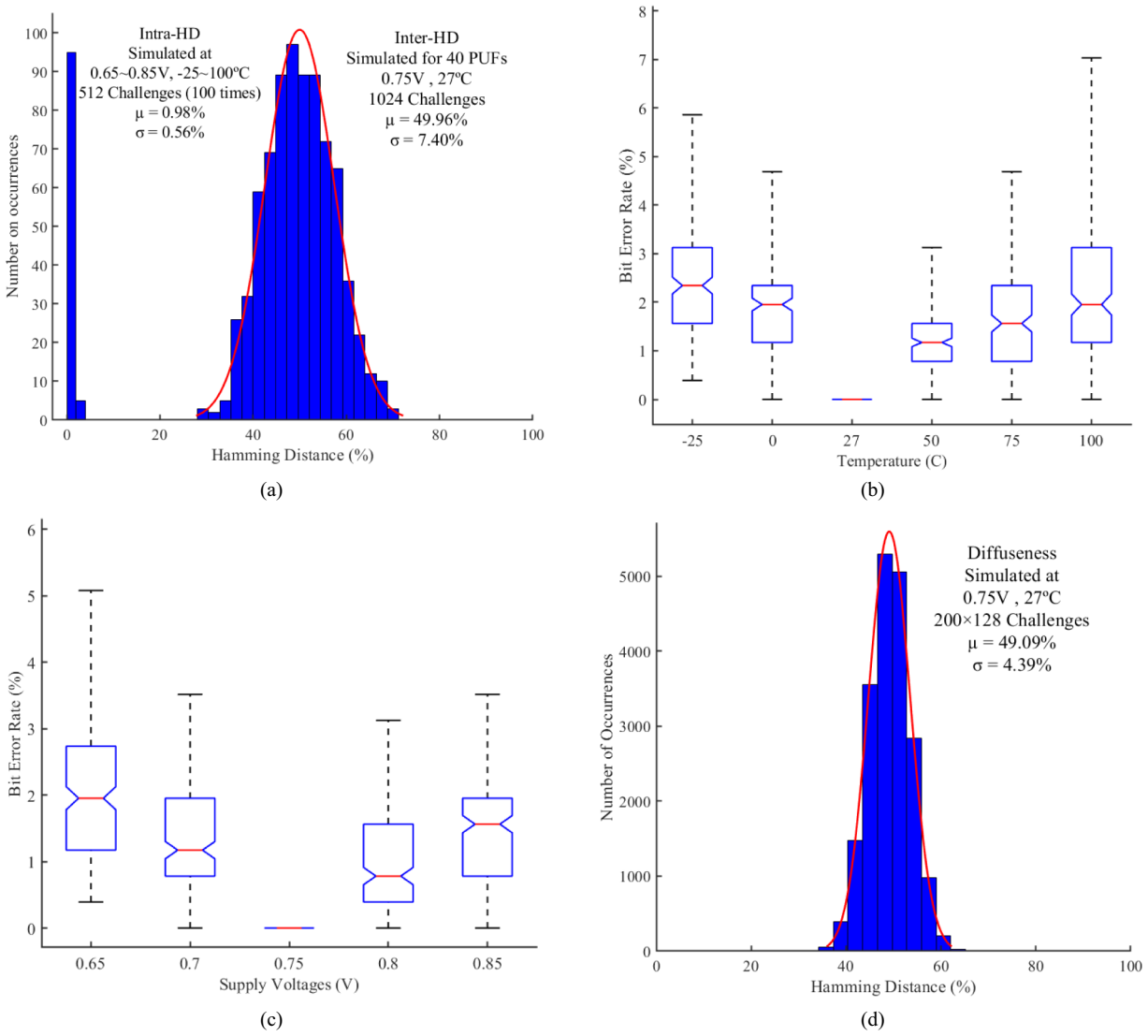

Fig. 7. Results of performance evaluation (**a**) Intra and inter-HD (**b**) BER under temperature variation when the supply voltage is 0.75V (**c**) BER under supply voltage variations when the temperature is 27°C (**d**) diffuseness when the supply voltage is 0.75V and temperature is 27°C for 100 sets of challenges under the fabrication process variation.

samples while 75% of those used for training against ML attacks, is examined for a two-array circuit with 60 MTJs in each array and four-array with ASC as shown in Fig. 8. To this end, the three most common algorithms are used as follows:

Support vector machine (SVM)
SVM is a widely utilized algorithm in machine learning for classification tasks. The primary objective of SVM is to discover a hyperplane within the feature space that effectively separates data belonging to different classes[54]. The algorithm strategically positions the hyperplane by identifying support vectors—the data points closest to the decision boundary to maximize the margin between these classes. This capability of SVM in class separation is particularly valuable in analyzing attacks on PUFs, as it enables the algorithm to forecast responses and mimic behavior in attack scenarios. In this work, similar to[28], an SVM with a nonlinear RBF kernel was used. The simulation results for the proposed two-array and four-array circuits were 54.04% and 49.60%, respectively. These results indicate that the SVM algorithm did not achieve the expected success in finding the optimal hyperplane to predict the responses accurately.

Logistic regression (LR)
LR is a widely acknowledged supervised learning technique typically applied in binary classification tasks[55]. It predicts the probability of a certain outcome based on specific input variables. The LR algorithm relies on the sigmoid function and weights learned from the training data. By focusing on the likelihood of a sample belonging

Test	*P*-value	Proportion	Pass/Fail
Frequency	0.021999	97/100	Pass
Block frequency	0.419021	100/100	Pass
Cumulative sums	0.026948	98/100	Pass
Cumulative sums	0.249284	97/100	Pass
Runs	0.191687	96/100	Pass
Longest runs	0.911413	98/100	Pass
Rank	0.554420	100/100	Pass
FFT	0.616305	98/100	Pass
Overlapping template	0.048716	99/100	Pass
Approximate entropy	0.935716	100/100	Pass
Serial	0.455937	100/100	Pass
Serial	0.816537	100/100	Pass
linear complexity	0.534146	98/100	Pass
Universal	0.798139	98/100	Pass
Non-overlapping template	Pass		
Random excursions variant	Pass		
Random excursions	Pass		

Table 2. Results of the NIST statistical test.

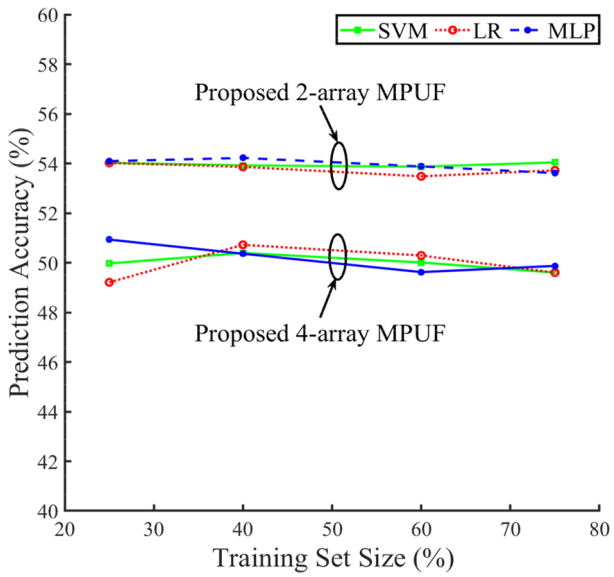

Fig. 8. Results of ML-attacks on 2 and 4 array proposed MPUF with different training set sizes.

to one of two classes, logistic regression uses the logistic function to convert a linear output into a probability between 0 and 1[56]. Despite its simplicity, logistic regression can be an effective and interpretable model for predicting the behavior of PUFs, such as the binary output of Arbiter PUFs. Several studies have employed this method to model the probability of correct or incorrect responses, capturing nonlinearity within binary datasets. In this paper, similar to[28], an LR model with the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) solver was used for simulation. The simulation results for the two-array and four-array circuits yielded 53.72% and 49.60%, respectively, suggesting that LR did not effectively analyze PUF behavior.

Multilayer perceptron (MLP)

MLP is a neural network used to understand complex relationships within data, particularly in deep learning applications. Unlike single-layer perceptrons (SLPs), which can only handle linear data, MLPs incorporate hidden layers with nonlinear activation functions like ReLU to capture and predict intricate patterns[27,53]. In the realm of PUF modeling, MLPs play a crucial role in replicating PUF behavior, aiding in the probing of advanced PUF structures, and evaluating their resilience against machine learning attacks. Like[28], an assault simulation was conducted using an MLP-based model with three hidden layers, each containing 300 neurons, utilizing the ReLU activation function and the Adam solver on the two-array and four-array PUF circuits. The analysis results

indicated that the applied model could not decipher the pattern of the proposed PUF circuit, with prediction accuracies of 53.61% and 49.87%, respectively.

For every challenge, the circuit's output is only one bit. When there is an even distribution of cases, if the ML-based algorithms cannot precisely predict the correct responses and instead choose randomly, the prediction rate should remain around 50%. From this, we can conclude that the proposed circuit can withstand these types of attacks.

Figure 9 depicts the results of ML-based modeling attacks on two 2-array circuits proposed in[28] and the proposed MPUF in this paper. It is observed that with the same number of CRPs, the circuit presented in[28] reaches a prediction accuracy of approximately 80%. In contrast, the proposed circuit maintains an accuracy of about 54%, indicating its robustness against the conducted attacks.

Figure 10 also illustrates prediction accuracy results for the proposed MPUF, arbiter PUF, and MPUF suggested in[28], and the current array PUF proposed in[30]. For small training sets, arbiter PUF reaches an accuracy close to 65%, while other PUFs are around 50%. As the training set size increases, arbiter PUF's accuracy quickly reaches 99%, while the current array PUF maintains an accuracy of around 60%. Arbiter PUF's significant increase in accuracy is due to its poor strict avalanche criterion (SAC), which makes it easily modeled. On the contrary, the proposed MPUF shows high resistance to ML modeling due to its nonlinearity and grid-like architecture. The proposed MPUF limits accuracy to about 49.87%, maintaining resistance even with an extensive training dataset.

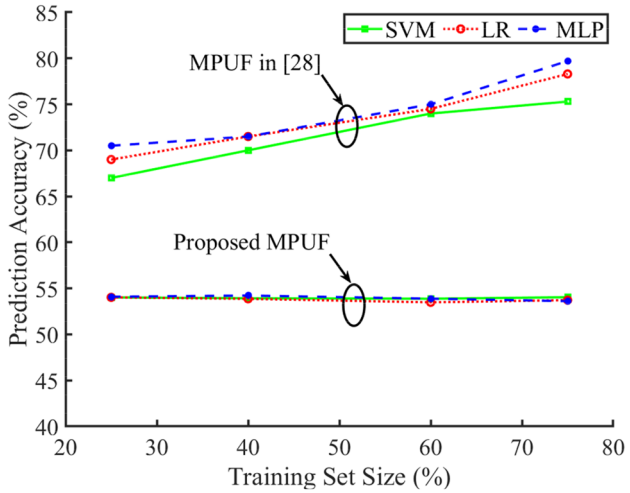

Fig. 9. Results of the ML-attacks on 2-array MPUF proposed in[28] and our proposed 2-array MPUF with different training set sizes.

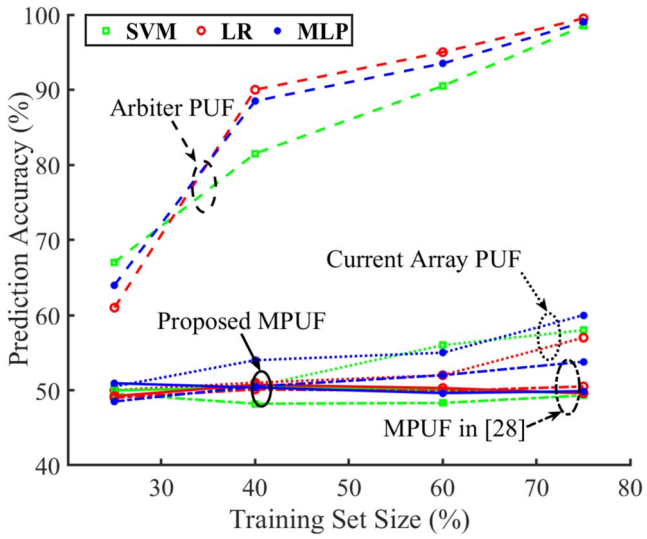

Fig. 10. The ML modeling robustness of the proposed 4-array MPUF in comparison to the other PUFs.

DL-based modeling for attack simulation

Deep learning is a special and more complex ML type used in applications with large datasets. CNN, RNN, MLP, and Larq[57] are among the most important deep-learning networks.

Convolutional Neural Networks (CNNs)

The CNN-based model has demonstrated superior potential in modeling highly nonlinear data, making it an effective tool for attacking PUFs by learning patterns and correlations in CRPs. These networks, composed of convolutional layers, pooling layers, and fully connected layers[58], automatically extract features from raw data for final classification, making them well-suited for modeling PUFs without the need to understand specific characteristics. CNNs achieve higher prediction accuracy and faster convergence by utilizing resilient back-propagation as the training algorithm. Although CNN-based attack engines require more computational power and complexity than other classifiers, they can be deployed on powerful servers where the trained models are hosted[59]. This capability allows attackers to replicate authorized nodes, creating malicious nodes that can compromise PUF security by accurately predicting responses to unseen challenges, especially when the spatial arrangement of bits or signals is critical.

Larq

Larq is a deep learning framework designed specifically for training and deploying Binarized Neural Networks (BNNs)[57]. BNNs are a type of neural network where weights and activations are constrained to binary values, making them highly efficient in memory and computation. In the context of PUFs, Larq can be used to develop lightweight models capable of predicting PUF responses with reduced computational resources. Despite their simplicity, BNNs trained using Larq can be quite effective in modeling PUF behavior, particularly when the PUF structure is relatively simple or when the attacker can access many CRPs. This approach can lead to efficient and fast attacks on PUFs, making them a significant threat.

Recurrent Neural Networks (RNNs)

RNNs are artificial neural networks designed to work with sequential information[60]. In the context of PUFs, RNNs can help simulate PUFs whose behaviors are time-dependent or sequentially influenced by earlier responses. When responding to certain challenges, earlier responses might influence subsequent ones, and RNNs can adjust for this, thereby improving prediction power. Due to their ability to retain memory from past inputs, adversaries could develop more sophisticated models for predicting future responses of a PUF, making it a significant security risk.

Deep Multilayer Perceptrons (Deep MLPs)

MLPs with multiple layers can handle nonlinear dependencies and correlations present in large datasets, thanks to their fully interconnected structures. When used for PUF attacks, deep MLPs can learn the complexities of challenge-response interactions[61]. The increased depth of these networks enables attackers to capture more subtle behaviors or inconspicuous correlations in a true PUF prototype. This is particularly effective against complex PUFs or when the relationship between challenge and response is strongly nonlinear.

The number of parameters used in deep learning is more than that used in machine learning, so it requires a more extensive training dataset and more training time[62]. Several hyperparameters have been carefully selected to optimize the model's performance in training a deep learning network for binary classification. Table 3 illustrates the critical hyperparameters used in this simulation, tailored to the dataset, which is of binary and single-class type. A learning rate of 0.001 facilitates the adjustment of model parameters during training iterations. The Adam optimizer, known for its efficiency and adaptability, helps to optimize the network's weights and biases. Binary cross-entropy, chosen as the loss function, quantifies the disparity between predicted and actual class labels. Rectified Linear Unit (ReLU) activation functions are employed in the hidden layers due to their simplicity and effectiveness in mitigating the vanishing gradient problem. The Sigmoid activation function is utilized for the output layer to produce probabilistic outputs within the range (0, 1), suitable for binary classification tasks.

For this reason, in the proposed 25×25 two-array MPUF, a CRP-balanced dataset of 868,000 with a training set size of 650,000 is used to simulate the DL-based modeling attack (similar to what was done in ML). Table 4 outlines the structure of networks and presents the simulation results of the proposed circuit's resilience against DL attacks. From the accuracy value, it can be concluded that the network was unsuccessful in attacking the proposed design, and the predictions were random.

Hyperparameter	Values
Learning rate	0.001
Loss function	Binary cross entropy
Optimizer	Adam
Activation function of hidden layers	ReLU
Activation function of the last layer	Sigmoid

Table 3. Important hyperparameters for DL-based attack.

Network	Structure	Prediction accuracy (%)
MLP	Nine Fully hidden connected layers and one output layer	50.31
CNN	Three Convolutional layers, three MaxPooling layers, eight fully connected layers, and one output layer	50.31
RNN	Two LSTM layers, and eight fully connected layers, and one output layer	50.25
Larq	Three Convolutional layers, two MaxPooling layers, twelve BatchNorm. layers, eight fully connected Layers, and one output layer	50.31

Table 4. Structure of networks and prediction accuracies.

Corner simulations

In PUF design, it is imperative to assess the robustness and reliability of the circuit in the various fabrication corners[63], considering the intricate nature of MRAM-based PUFs and their vulnerability to process variations. For corner simulations, critical parameters of the proposed MPUF are chosen using the corner value indicated in Table 5.

Table 6 illustrates the simulation results evaluating resistance against ML-based classification attacks for the proposed circuits in both the 60-cell 2-array and the 31-cell 4-array architecture in all eight possible corners. These results signify that the proposed MPUF maintains robust performance even under the worst-case scenarios arising from the variations in the fabrication process.

Comparison

Table 7 shows the simulation results of the proposed MPUF and the other state-of-the-art PUFs. The results of this table show the proposed MPUF superiority, particularly in terms of energy and area compared to other PUFs. Regarding the CRP space and transistor count, for a proposed circuit without ASC consisting of 4 MTJ cells, similar to Fig. 6, there are 625 CRP states and 192 transistors. In comparison, for the circuit in[28] with an equal number of MTJ cells and no ASC, the CRP space is limited to 16 with 240 transistors, indicating the advantage of the proposed circuit over the previous work (without considering the state of MTJs).

Description	Typical value (T)	Corner value	
		Slow (S)	Fast (F)
MTJ			
Tunnel magnetoresistance ratio	200%	220%	180%
Thickness of the oxide barrier	0.85 nm	0.8925 nm	0.8075 nm
The thickness of the Free layer	1.3 nm	1.235 nm	1.365 nm
Diameter of the MTJ	64 nm	59.392 nm	68.61 nm
Resistance area product	10 $\Omega\ \mu m^2$	11.5 $\Omega\ \mu m^2$	8.5 $\Omega\ \mu m^2$
FinFET			
Fin thickness	7 nm	6.5 nm	7.5 nm
Fin height	32 nm	28 nm	36 nm
Gate length	21 nm	25 nm	17 nm
Oxide thickness	2.1 nm	2.31 nm	1.89 nm

Table 5. Corner values of important device parameters.

MTJ	FinFET	Proposed 2-array			Proposed 4-array		
		MLP (%)	LR (%)	SVM (%)	MLP (%)	LR (%)	SVM (%)
S	SS	53.62	53.31	54.38	51.88	48.16	51.37
	SF	53.62	53.31	53.31	51.88	48.16	51.37
	FS	53.62	53.31	53.31	51.88	48.16	51.37
	FF	53.62	53.31	53.31	51.88	50.12	50.32
T	TT	53.61	53.72	54.04	49.87	49.60	49.60
F	SS	57.31	57.39	57.39	50.57	48.16	50.32
	SF	57.31	57.18	57.18	50.57	49.40	50.32
	FS	57.31	57.18	57.18	50.57	49.50	50.32
	FF	56.80	57.47	57.47	50.37	49.36	48.98

Table 6. Results of the corner simulations.

PUF	[30]	[7]	[24]	[29]	[28]		Proposed in this paper
Technology	**130 nm**	**65 nm**	**40 nm**	**40 nm**	**28 nm**	**7 nm**	**7 nm**
PUF Type	Current array	Memristive crossbar	STT-MRAM	STT-MRAM	STT-MRAM		STT-MRAM
Number of CRPs	3.7×10^{19}	NA	NA	NA	$2^m\times2^n\times z$		$2^x\times2^y\times z$
Inter-HD (%)	49.9	47.5	47	49 ~ 51	49.76	49.30	49.96
Intra-HD (%)	5.8	10	2.25	3 ~ 4	0.447	2.04	0.98
ML-Prediction Accuracy (%)	60	58	NA	NA	53.8		49.87
Energy (fJ/bit)	11000	106600	20000	3400	870	24.42	9.57
Area (μm^2)	44700	NA	6.74	6061.2	4.5	13.12	7.96

Table 7. Comparative analyses of the proposed MPUF. m = Number of MTJ cells in even arrays. n = Number of MTJ cells in odd arrays. x = Number of charge–discharge paths in even arrays. y = Number of charge–discharge paths in odd arrays. z = Number of possible ASC challenge configurations.

The findings presented in Table 7 underscore the proposed MPUF's significant advantages compared to the design outlined in[28]. Specifically, the intra-HD of the proposed MPUF is demonstrated to be at least two times lower than the design mentioned above in the same technology, highlighting superior performance in terms of reliability. Moreover, due to the value of inter-HD in the proposed circuit, a slight improvement in uniqueness is observed. Furthermore, the energy efficiency of the proposed MPUF stands out, as it consumes less energy per bit. Additionally, using the ASC and a judicious number of MTJs enables the proposed MPUF to provide an impressive quantity of CRPs, further enhancing its versatility and potential applications.

Conclusion and future work

This paper proposed an ML and DL modeling attack immune MPUFs circuit with a large number of CRPs. The proposed MPUF is based on the intrinsic variation of the MTJs during fabrication. This feature provides unique characteristics in each fabricated PUF. Thanks to the grid-like structure and the utilization of ASC of the proposed MPUF, the proposed MPUF offers high security and ML and DL modeling immunity. ML attack simulation shows a prediction accuracy of 53.61% for the two-array circuit and 49.87% for the four-array circuit, indicating the immunity of the proposed MPUF to ML modeling. In addition, DL modeling attacks are also simulated to demonstrate the reliability of the circuit against CNN, RNN, MLP, and Larq with an accuracy result of 50.31%, 50.25%, 50.31%, and 50.31%, respectively. Considering other evaluation metrics such as reliability, uniqueness, and uniformity, the proposed MPUF offers intra- and inter-HD of 0.98% and 49.96%, respectively, and diffuseness with a mean of 49.09%. Additionally, the proposed MPUF excels the state-of-the-art PUFs in energy consumption. Moreover, corner simulation validates the robust performance of the proposed MPUF even in the presence of the fabrication process variation.

Despite these strengths, the proposed MPUF faces scalability challenges, environmental sensitivity, and increased complexity. However, its robust security features make it highly suitable for applications in IoT device authentication, secure key storage, anti-counterfeiting, and supply chain security. Its energy efficiency and compact design also make it ideal for integration into mobile devices such as wearable devices, particularly in healthcare, where the amount of available energy is limited. However, at the same time, secure and reliable data transmission is critical. Future research can explore enhancing the capabilities of this circuit and increasing the CRP to improve its performance and application further.

Data availability

Data related to the current study are available from the corresponding author upon reasonable request.

Received: 8 February 2024; Accepted: 30 August 2024
Published online: 04 September 2024

References

1. Ma, Q., Gu, C., Hanley, N., Wang, C., Liu, W. & O'Neill M. A machine learning attack resistant multi-PUF design on FPGA," in *23rd Asia and South Pacific Design Automation Conf. (ASP-DAC), IEEE*, pp. 97–104, https://doi.org/10.1109/ASPDAC.2018.8297289 (2018).
2. Gargari, M. A., Eslami, N. & Moaiyeri, M. H. A reconfigurable nonvolatile memory architecture for prolonged wearable health monitoring devices. *IEEE Trans. Consumer Electr.* https://doi.org/10.1109/tce.2024.3399223 (2024).
3. Schultz, T., Jha, R., Casto, M. & Dupaix, B. Vulnerabilities and reliability of ReRAM based PUFs and memory logic. *IEEE Trans. Reliab.* **69**(2), 690–698. https://doi.org/10.1109/tr.2019.2910793 (2020).
4. Böttger, S., Frank, F., Anagnostopoulos, N. A., Mohamed, A., Hartmann, M., Arul, T., Hermann, S. & Katzenbeisser, S. Cnt-pufs: highly robust physical unclonable functions based on carbon nanotubes," in *2023 IEEE 23rd International Conf. on Nanotechnology (NANO)*, Sep. 2023: IEEE, pp. 1–6, https://doi.org/10.1109/NANO58406.2023.10231160 (2023)
5. Lim, S., Song, B. & Jung, S.-O. Highly independent MTJ-based PUF system using diode-connected transistor and two-step post-processing for improved response stability. *IEEE Trans. Inf. Forens. Secur.* **15**, 2798–2807. https://doi.org/10.1109/tifs.2020.2976623 (2020).

6. Zheng, Y., Zhang, F. & Bhunia, S. ScanPUF: a delay-based physical unclonable function built into scan chain. *IEEE Trans. Very Large Scale Integrat. (VLSI) Syst* **24**(3), 1059–1070. https://doi.org/10.1109/tvlsi.2015.2421933 (2016).
7. Uddin, M., Majumder, M. B. & Rose, G. S. Robustness analysis of a memristive crossbar PUF against modeling attacks. *IEEE Trans. Nanotechnol.* **16**(3), 396–405. https://doi.org/10.1109/tnano.2017.2677882 (2017).
8. Lin, C.-C. & Chen, M.-S. Enhancing reliability and security: A configurable poisoning PUF against modeling attacks. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **41**(11), 4301–4312 (2022).
9. Mohammadinodoushan, M., Cambou, B., Afghah, F., Philabaum, C. R. & Burke, I. Reliable, secure, and efficient hardware implementation of password manager system using SRAM PUF. *IEEE Access* **9**, 155711–155725. https://doi.org/10.1109/access.2021.3129499 (2021).
10. Suvizi, A., Subramaniam, S., Lan, T. & G. Venkataramani, "Exploring In-Memory Accelerators and FPGAs for Latency-Sensitive DNN Inference on Edge Servers," in *2024 IEEE Cloud Summit*, 2024: IEEE, pp. 1–6, https://doi.org/10.1109/Cloud-Summit61220.2024.00007.
11. Rai, V. K. Tripathy, S. & Mathew, J. "2SPUF: Machine learning attack resistant SRAM PUF," in *2020 Third ISEA conference on security and privacy (ISEA-ISAP)*, 2020: IEEE, pp. 149–154, https://doi.org/10.1109/ISEA-ISAP49340.2020.235013.
12. Shahrakht, A.-A., Hajirahimi, P., Rostami, O. & Martín, D. A novel attack on complex APUFs using the evolutionary deep convolutional neural network. *Intell. Automat. Soft Comput.* **37**(3), 3059–3081. https://doi.org/10.32604/iasc.2023.040502 (2023).
13. Mursi, K. T., Thapaliya, B., Zhuang, Y., Aseeri, A. O. & Alkatheiri, M. S. A fast deep learning method for security vulnerability study of XOR PUFs. *Electronics* https://doi.org/10.3390/electronics9101715 (2020).
14. Hiller, M., Yu, M.-D. & Sigl, G. Cherry-picking reliable PUF bits with differential sequence coding. *IEEE Trans. Inf. Forensics Secur.* **11**(9), 2065–2076 (2016).
15. Divyanshu, D., Kumar, R., Khan, D., Amara, S. & Massoud, Y. Physically unclonable function using GSHE driven SOT assisted p-MTJ for next generation hardware security applications, (in English). *IEEE Access* **10**, 93029–93038. https://doi.org/10.1109/Access.2022.3203817 (2022).
16. Srinivas, M. B. R. & Elango, K. Era of sentinel tech: Charting hardware security landscapes through post-silicon innovation, threat mitigation and future trajectories. *IEEE Access* **12**, 68061–68108. https://doi.org/10.1109/access.2024.3400624 (2024).
17. Japa, A., Majumder, M. K., Sahoo, S. K. & Vaddi, R. Tunnel FET-based ultra-lightweight reconfigurable TRNG and PUF design for resource-constrained internet of things. *Int. J. Circuit Theory Appl.* **49**(8), 2299–2311. https://doi.org/10.1002/cta.3030 (2021).
18. Gao, Y., Ranasinghe, D. C., Al-Sarawi, S. F., Kavehei, O. & Abbott, D. Memristive crypto primitive for building highly secure physical unclonable functions. *Sci. Rep.* **5**, 12785. https://doi.org/10.1038/srep12785 (2015).
19. Ghosh, S. Spintronics and security: Prospects, vulnerabilities, attack models, and preventions. *Proc. IEEE* **104**(10), 1864–1893. https://doi.org/10.1109/jproc.2016.2583419 (2016).
20. Das, K. K., Soni, S. Moradi, F., Shreya, S. & Kaushik, B. K. Vortex spin torque nano oscillator-based PUF and TRNG design for lightweight security solutions, in *2024 8th IEEE Electron Devices Technology & Manufacturing Conference (EDTM)*, (2024)
21. Kang, J. *et al.* Highly reliable magnetic memory-based physical unclonable functions. *ACS Nano* https://doi.org/10.1021/acsnano.4c00078 (2024).
22. Japa, A., Majumder, M. K., Sahoo, S. K., Vaddi, R. & Kaushik, B. K. Hardware security exploiting post-CMOS devices: Fundamental device characteristics, state-of-the-art countermeasures, challenges and roadmap. *IEEE Circ. Syst. Mag.* **21**(3), 4–30. https://doi.org/10.1109/mcas.2021.3092532 (2021).
23. Amirany, A., Jafari, K. & Moaiyeri, M. H. DDR-MRAM: double data rate magnetic RAM for efficient artificial intelligence and cache applications. *IEEE Trans Magnet* https://doi.org/10.1109/tmag.2022.3162030 (2022).
24. Das, J., Scott, K., Rajaram, S., Burgett, D. & Bhanja, S. MRAM PUF: A novel geometry based magnetic PUF with integrated CMOS. *IEEE Trans. Nanotechnol.* **14**(3), 436–443 (2015).
25. Rezayati, M. H., Moaiyeri, M. H., Amirany, A. & Jafari, K. A new paradigm for immunization of deep neural networks against replication attacks based on spintronics. *IEEE Trans. Circ. Syst. II: Expr. Briefs* https://doi.org/10.1109/tcsii.2024.3371154 (2024).
26. Hou, Z. *et al.* Reconfigurable and dynamically transformable In-Cache-MPUF system with true randomness based on the SOT-MRAM. *IEEE Trans. Circ Syst I: Regular Paper* **69**(7), 2694–2706. https://doi.org/10.1109/TCSI.2022.3168133 (2022).
27. Khalfaoui, S. *et al.* Security analysis of machine learning-based PUF enrollment protocols: a review. *Sensors (Basel)* https://doi.org/10.3390/s21248415 (2021).
28. Ali, R., Zhang, D., Cai, H., Zhao, W. & Wang, Y. A machine learning attack-resilient strong PUF leveraging the process variation of MRAM. *IEEE Trans. Circ. Syst. II Expr. Briefs* **69**(6), 2712–2716. https://doi.org/10.1109/tcsii.2022.3144497 (2022).
29. Ben Dodo, S., Bishnoi, R., Mohanachandran Nair, S. & Tahoori, M. B. A spintronics memory PUF for resilience against cloning counterfeit. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **27**(11), 2511–2522. https://doi.org/10.1109/tvlsi.2019.2931481 (2019).
30. Xi X., Zhuang H., Sun, N. & Orshansky, M. Strong subthreshold current array PUF with 2 65 challenge-response pairs resilient to machine learning attacks in 130nm CMOS," in *2017 Symp. on VLSI Circuits*, Kyoto, Japan, IEEE, pp. C268-C269, https://doi.org/10.23919/VLSIC.2017.8008503 (2017)
31. Akbari, M., Mirzakuchaki, S., Jamshidi, V., Fazeli, M. & Tarihi, M. R. An Ultra-compact pure magnetic arbiter PUF with high reliability and low power consumption. *IEEE Trans. Nanotechnol.* **22**, 449–456. https://doi.org/10.1109/tnano.2023.3292481 (2023).
32. Gajaria, D., Antony Gomez, K. & Adegbija, T. A study of STT-RAM-based In-memory computing across the memory Hierarchy. in *2022 IEEE 40th International Conference on Computer Design (ICCD)*, (2022).
33. Amirany, A., Jafari, K. & Moaiyeri, M. H. True random number generator for reliable hardware security modules based on a neuromorphic variation-tolerant spintronic structure. *IEEE Trans. Nanotechnol.* https://doi.org/10.1109/tnano.2020.3034818 (2020).
34. Amirany, A., Moaiyeri, M. H. & Jafari K. MTMR-SNQM: multi-tunnel magnetoresistance spintronic nonvolatile quaternary memory," in *2021 IEEE 51st International Symposium on Multiple-Valued Logic (ISMVL)*, *IEEE*, pp. 172–177, https://doi.org/10.1109/ISMVL51352.2021.00037 (2021).
35. Gargari, M. A., Eslami, N. & Moaiyeri, M. H. An energy efficient in-memory computing architecture using reconfigurable magnetic logic circuits for big data processing. *IEEE Trans. Magnet.* https://doi.org/10.1109/tmag.2023.3322731 (2023).
36. Wu, B., Cheng, Y., Yang, J., Todri-Sanial, A. & Zhao, W. Temperature impact analysis and access reliability enhancement for 1T1MTJ STT-RAM. *IEEE Trans. Reliab.* **65**(4), 1755–1768 (2016).
37. Ali, R. *et al.* A reconfigurable arbiter MPUF with high resistance against machine learning attack. *IEEE Trans. Magn.* **57**(10), 1–7. https://doi.org/10.1109/tmag.2021.3102838 (2021).
38. Chen, Y.-S. *et al.* On the hardware implementation of MRAM physically unclonable function. *IEEE Trans. Electron Devices* **64**(11), 4492–4495. https://doi.org/10.1109/ted.2017.2755867 (2017).
39. Mohseni, A., Moaiyeri, M. H., Amirany, A. & Hadi Rezayati, M. Protecting the intellectual property of binary deep neural networks with efficient spintronic-based hardware obfuscation. *IEEE Trans. Circu. Syst. I: Regular Pap.* **71**(7), 3146–3156. https://doi.org/10.1109/tcsi.2024.3397925 (2024).
40. Nasab, M. T., Amirany, A., Moaiyeri, M. H. & Jafari, K. High-performance and robust spintronic/CNTFET-based binarized neural network hardware accelerator. *IEEE Trans. Emerg. Top. Comput.* https://doi.org/10.1109/TETC.2022.3202113 (2022).
41. Jamshidi, V. & Fazeli, M. Pure magnetic logic circuits: A reliability analysis. *IEEE Trans. Magn.* **54**(10), 1–10. https://doi.org/10.1109/tmag.2018.2846623 (2018).
42. Hu, Y. *et al.* STT-MRAM-based reliable weak PUF. *IEEE Trans. Comput.* **71**(7), 1564–1574. https://doi.org/10.1109/TC.2021.3095657 (2021).

43. Amirany, A., Meghdadi, M., Moaiyeri, M. H. & Jafari, K. Stochastic spintronic neuron with application to image binarization, in *presented at the 2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, (2021).
44. Wu, L., Rao, S., Taouil, M., Marinissen, E. J., Sankar Kar, G. & Hamdioui, S. Testing STT-MRAM: manufacturing defects, fault models, and test solutions. in *presented at the 2021 IEEE International Test Conference (ITC)*, (2021).
45. Amirany, A., Jafari, K. & Moaiyeri, M. H. A task-schedulable nonvolatile spintronic field-programmable gate array. *IEEE Magn. Lett.* **12**, 1–4. https://doi.org/10.1109/lmag.2021.3092995 (2021).
46. Mei, T., Meng, Z., Zhao, K. & Chen, C. Q. A mechanical metamaterial with reprogrammable logical functions. *Nat. Commun.* **12**(1), 7234. https://doi.org/10.1038/s41467-021-27608-7 (2021).
47. Gandhi, P. P. & Devashrayee, N. M. A novel low offset low power CMOS dynamic comparator. *Analog Integr. Circ. Sig. Process* **96**(1), 147–158. https://doi.org/10.1007/s10470-018-1166-9 (2018).
48. Clark, L. T. *et al.* ASAP7: A 7-nm finFET predictive process design kit. *Microelectron. J.* **53**, 105–115. https://doi.org/10.1016/j.mejo.2016.04.006 (2016).
49. Wang, Y. *et al.* Compact model of dielectric breakdown in spin-transfer torque magnetic tunnel junction. *IEEE Trans. Electron Devices* **63**(4), 1762–1767. https://doi.org/10.1109/TED.2016.2533438 (2016).
50. Maes, R. Physically unclonable functions: concept and constructions, in *Physically unclonable functions: constructions, Properties and applications*. Berlin, Heidelberg: Springer, pp. 11–48 (2013).
51. Hou, S. *et al.* A dynamically configurable LFSR-based PUF design against machine learning attacks. *CCF Trans. High Perform. Comput.* **3**(1), 31–56. https://doi.org/10.1007/s42514-020-00060-7 (2020).
52. Yao, J. *et al.* Design and evaluate recomposited or-and-xor-puf. *IEEE Trans. Emerg Top. Comput.* **10**(2), 662–677. https://doi.org/10.1109/TETC.2022.3170320 (2022).
53. Henderson, E. R., Henderson, J. M., Shahoei, H., Oxford, W. V., Larson, E. C., MacFarlane, D. L. & Thornton, M. A. Designing a photonic physically unclonable function having resilience to machine learning attacks. in *Quantum Information Science, Sensing, and Computation XVI*, 2024, vol. 13028: SPIE, pp. 64–77, https://doi.org/10.1117/12.3013126.
54. Khalfaoui, S. *et al.* Security analysis of machine learning-based puf enrollment protocols: A review. *Sensors* **21**(24), 8415. https://doi.org/10.3390/s21248415 (2021).
55. Rührmair, U. *et al.* PUF modeling attacks on simulated and silicon data. *IEEE Trans. Inform. Foren. Secur.* **8**(11), 1876–1891. https://doi.org/10.1109/TIFS.2013.2279798 (2013).
56. Rührmair, U., Sehnke, F., Sölter, J., Dror, G., Devadas, S. & Schmidhuber, J. Modeling attacks on physical unclonable functions, in *Proceedings of the 17th ACM conference on Computer and communications security*, pp. 237–249, https://doi.org/10.1145/1866307.1866335(2010).
57. Geiger, L. & Team, P. Larq: An open-source library for training binarized neural networks. *J. Open Sour. Softw.* https://doi.org/10.21105/joss.01746 (2020).
58. Zhang, J. & Shen, C. Set-based obfuscation for strong PUFs against machine learning attacks. *IEEE Trans. Circ. Syst. I: Regular Papers* **68**(1), 288–300. https://doi.org/10.1109/tcsi.2020.3028508 (2021).
59. Sajadi, A., Shabani, A. & Alizadeh, B. DC-PUF: Machine learning-resistant PUF-based authentication protocol using dependency chain for resource-constraint IoT devices. *J. Netw. Comput. Appl.* https://doi.org/10.1016/j.jnca.2023.103693 (2023).
60. Lalouani, W., Younis, M., Ebrahimabadi, M. & Karimi, N. Countering modeling attacks in PUF-based IoT security solutions. *ACM J. Emerg. Technol. Comput. Syst.* **18**(3), 1–28. https://doi.org/10.1145/3491221 (2022).
61. Wisiol, N., Thapaliya, B., Mursi, K. T., Seifert, J.-P. & Zhuang, Y. Neural Network modeling attacks on arbiter-PUF-based designs. *IEEE Trans. Inform. Forens. Secur.* **17**, 2719–2731. https://doi.org/10.1109/tifs.2022.3189533 (2022).
62. Khalifa, N. E., Loey, M. & Mirjalili, S. A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artif. Intell. Rev.* **55**(3), 2351–2377. https://doi.org/10.1007/s10462-021-10066-4 (2022).
63. Wang, Y., Cai, H., Naviner, L. & Zhao, W. A non-monte-carlo methodology for variability analysis of magnetic tunnel junction based circuits. *IEEE Trans. Magnet.* https://doi.org/10.1109/tmag.2016.2638913 (2016).

Author contributions

M.J.A. Proposed the idea, performed and reviewed the experiments, and wrote the paper. M.H.R. and A.A. helped to develop the idea, discussed and interpreted the results, and edited the paper. M.H.M. reviewed the simulations, discussed and interpreted the results, edited the paper, supervised the study, and proofed the paper. K.J. discussed and interpreted the results and proofed the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.H.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.