Journal of
Low Power Electronics and Applications

MDPI

Article

Comprehensive Study of Side-Channel Attack on Emerging Non-Volatile Memories †

Mohammad Nasim Imtiaz Khan [1], Shivam Bhasin [2], Bo Liu [1], Alex Yuan [1], Anupam Chattopadhyay [3,*] and Swaroop Ghosh [1]

[1] School of Electrical Engineering and Computer Science, Pennsylvania State University, State College, PA 16801, USA; nasimimtiaz.khan@gmail.com (M.N.I.K.); noobdoge@seas.upenn.edu (B.L.); ahy5028@psu.edu (A.Y.); szg212@psu.edu (S.G.)
[2] Temasek Laboratories, Nanyang Technological University, Singapore 637553, Singapore; sbhasin@ntu.edu.sg
[3] School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore
* Correspondence: anupam@ntu.edu.sg
† This paper is an extended version of our paper published in Khan, M.N.I.; Bhasin, S.; Yuan, A.; Chattopadhyay, A.; Ghosh, S. Side-Channel Attack on STTRAM Based Cache for Cryptographic Application. In Proceedings of the 2017 IEEE International Conference on Computer Design (ICCD), Boston, MA, USA, 5–8 November 2017; pp. 33–40, doi:10.1109/ICCD.2017.14.

Abstract: Emerging Non-Volatile Memories (NVMs) such as Magnetic RAM (MRAM), Spin-Transfer Torque RAM (STTRAM), Phase Change Memory (PCM) and Resistive RAM (RRAM) are very promising due to their low (static) power operation, high scalability and high performance. However, these memories bring new threats to data security. In this paper, we investigate their vulnerability against Side Channel Attack (SCA). We assume that the adversary can monitor the supply current of the memory array consumed during read/write operations and recover the secret key of Advanced Encryption Standard (AES) execution. First, we show our analysis of simulation results. Then, we use commercial NVM chips to validate the analysis. We also investigate the effectiveness of encoding against SCA on emerging NVMs. Finally, we summarize two new flavors of NVMs that can be resilient against SCA. To the best of our knowledge, this is the first attempt to do a comprehensive study of SCA vulnerability of the majority of emerging NVM-based cache.

Keywords: STTRAM; MRAM; RRAM; PCM; experimental validation; side channel attack; AES; key extraction; encoding

Citation: Khan, M.N.I.; Bhasin, S.; Liu, B.; Yuan, A.; Chattopadhyay, A.; Ghosh, S. Comprehensive Study of Side-Channel Attack on Emerging Non-Volatile Memories. *J. Low Power Electron. Appl.* **2021**, *11*, 38. https://doi.org/10.3390/jlpea11040038

Academic Editor: Luis Parrilla Roure

Received: 1 September 2021
Accepted: 24 September 2021
Published: 28 September 2021

1. Introduction

Several emerging Non-Volatile Memories (NVMs) such as Magnetic RAM (MRAM) [1,2], Spin-Transfer Torque RAM (STTRAM) [3–5], Phase Change Memory (PCM) [6,7] and Resistive RAM (RRAM) [8,9] offer high density/speed and low (static) power operation. Their application can extend various sectors. For example, SRAM in cache (L2/L3) can be replaced by STTRAM [10,11] since it offers compatible endurance and speed. RRAM and STTRAM can be considered to replace eFlash [10,12]. A Solid State Drive (SSD) using PCM, namely Optane is already sold by Intel [13]. Additionally, low-power computations and novel architectures can be realized by NVMs [10,11]. In addition to storage, NVMs are also investigated for novel applications, such as neuromorphic computing, ambient sensor, security primitive, etc., [10]. However, the unique NVM features brings new threats to data privacy and security. Therefore, they should be investigated properly before their wide adoption.

Vulnerabilities of emerging NVMs: NVMs are susceptible to ambient parameters, such as thermal and magnetic field. This can be leveraged to launch Denial-of-Service (DoS) attacks [14]. NVMs also suffer from supply voltage droop. The role of high supply noise (i.e., droop and ground bounce) to launch fault injection attack has also been investigated

in [15]. An adversary can generate deterministic supply noise in emerging NVM-based Last Level Cache (LLC) by writing a specific data pattern in his memory. The generated noise can propagate to the victim's space and affect the write/read operation (i.e., fault injection attacks [15]). Furthermore, NVMs suffer from asymmetric [16] and high read/write current (i.e., write/read current for data '1' and data '0' are different), which can be leveraged to launch Side-Channel Attack (SCA) [17,18].

SCA has been a serious threat to cryptographic chip [19] which are widely used in computer/network-based security control systems [20], as well as the secure measurement systems [21]. In prior works, secured and low-power asynchronous Advanced Encryption Standard (AES) substitution box (S-Box) has been proposed [22]. Furthermore, asymmetric mask has also been proposed to secure Data Encryption Standard (DES) circuit [19]. Traditional memories have also been investigated for their vulnerability to SCA. The work [23] shows that a geometrically regular structure can be observed in SRAM. The side channel leakage significantly correlates with the number of state transitions, which is known as Hamming Distance (HD). This is true since SRAM write current is different when a complementary data are written compared to the previously stored data. The current dependency on the data can be leveraged and information can be extracted by power analysis attacks. Therefore, the new memory technologies should be investigated for their resistance to such known threats.

This work investigates the vulnerability of various NVM-based LLC by monitoring the current drawn by the LLC from V_{DD} during write and/or read operations and performing Differential Power Analysis (DPA) based SCA. We have assumed that the CPU executes cryptographic operations (typically executed on network data). The cipher text, intermediate computations and round keys are periodically written (in LLC) and read from LLC. A small resistance in series with V_{dd} or GND can be leveraged to find the current drawn by LLC by measuring the voltage drop across it. The voltage can be sampled using precise instruments at high frequency (1 GHz) with high accuracy (<1% error). Note that accurate measurement of the power profile is essential of a successful attack [22]. Techniques such as [24] can provide accurate power consumption characterization. A picture of the system level illustration of attack setup is shown in Figure 1.

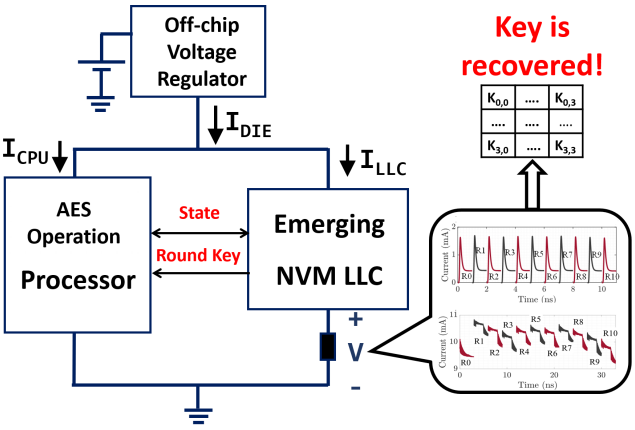


Figure 1. Picture showing CPU, Last Level Cache (LLC) and an external voltage regulator. Inserting resistor and monitoring die current through it, can be leveraged to extract key using DPA [25].

STTRAM suffers from high and asymmetric read/write current [26,27]. Therefore, STTRAM read/write current is a function of new data being written and the stored data. In [27], write current is leveraged to design a Correlation Power Analysis (CPA) on STTRAM-based cache. Ref. [25] extended the above lines of works and investigated various information leakage channels, such as write/read current for STTRAM and MRAM. In this work, we study the SCA vulnerability of read/write current of all emerging NVMs and analyze the attack model on the commercial chips such as MRAM. To the best of our

knowledge, this paper is the first work towards this direction. In summary, this work makes the following new contributions over [25]:

- Investigate SCA vulnerability of RRAM read/write operations based on simulation results;
- Investigate SCA vulnerability of PCM read/write operations based on simulation results;
- Show that encoding read/write operations cannot protect NVM from SCA attack;
- Emphasize on more device level solutions to remove the data signature.

2. General Background

2.1. Advanced Encryption Standard (AES)

AES [28] is a NIST standard for symmetric encryption. AES is a block cipher which encrypts 128-bit plaintext P into a 128 ciphertext C, with a secret key. The key size can be chosen from 128/192/256 bits, depending on the desired security. Depending on the key size, the algorithm consist of 10/12/14 rounds, respectively. 4 sub-operations are computed in each round $R_i, 0 \leq i \leq 9$ denotes round index on a state of 128 bits organized as 4×4 matrix. SubBytes (SB) is a non-linear look up table, followed by two linear permutations called ShiftRows (SR) and MixColumns (MC). A round key k_i computed from master key is added to the state in each round during AddRoundKeys (ARK) operation. All rounds except the last (skips MixColumns) are identical.

The prime target of this study is the last round key k_9 of AES-128, which computes $C = SR(SB(R_9))\hat{}K_9$. This operation is byte-wise over 16 bytes (128 bits). Given a hardware implementation with round based architecture, the final ciphertext overwrites the state register containing R_9. Therefore, C (1 byte) can be computed by finding R_9 (1 byte) and K_9 (1 byte). This observation can be used to implement a divide and conquer approach to extract all 16bytes, 1byte at a time. The key expansion algorithm is public and computes the master key from k_9.

2.2. Side-Channel Attack

Side-channel attacks (SCA [29]) target the physical implementation of otherwise theoretically secure cryptographic algorithms. The attacks exploit observable physical traits of the target implementation, which are related to underlying sensitive computation. The observable trait which is linked to underlying CMOS technology can be timing [30], power consumption [23] or electromagnetic (EM) emanation [31], etc. In terms of power analysis, a CMOS gate consumes significant power whenever the input changes (also known as transition), while the power consumed with no input change is minimal. Moreover, there could be secondary power consumption differences in different input transition ($0 \rightarrow 1$ or $1 \rightarrow 0$). These differences can be easily captured on a standard oscilloscope. If the differences are related to sensitive computation, it leaks information to the attacker. HD model is used for memory components. Note that HD is equal to the total bit flips during write operation. Next, a statistical analysis is performed to find a dependency between the simulated or observed side channel measurements and the hypothetical leakage (calculated by the leakage model using hypotheses on the key). As stated before, SCA can work in a divide and conquer setting, like byte-wise in AES, allowing recovery of small parts of the key independently and combination of the recovered key bytes to form the full key. This makes the attack complexity linear For SCA to be successful, the correct key hypothesis, compared to others, shows a significantly higher statistical dependency. Some common statistical tools that are used for SCA are Pearson correlation coefficient [32] and difference of means (DoM [29]).

2.3. Attack Model

The attack targets memory updates (read/write) in NVM with sensitive data from a cryptographic computation.

2.3.1. Attacking Write Data

For memory write operations, the target is last round update of AES-128 encryption ($R_9 \rightarrow C$), where C is public. The leakage L can be modeled as: $L = HD(C, R_9) + N = HW(C \hat{} R_9) + N$ where N is the noise of measurement. HW(.) denotes the Hamming Weight function, which is equal to the total number of '1' in a data [33].

As k_9 is not known, the attacker can guess it bytewise to recover individual bytes of l_9. Thus, knowing one byte of C and guessing one byte of k_9, the attacker computes hypothetical value of R_9, leading to L. The correlation between L and the actual side-channel traces will be highest for the correct value of k_9, thus statistically revealing its value. The process is repeated for each byte of k_9 independently. If the algorithm is changed, the attack remains fairly the same, only functions to compute R_9 will be updated, which are anyways public.

2.3.2. Attacking Read Data

The attack process targeting read data are similar to the one for write. The main differences are:

The data bus is generally pre-charged to '0' each clock cycle before performing a new read. This makes the initial state of the target data bus where sensitive information is read, as all 0's. The attack targets the moment when R_9 of AES-128 encryption is loaded for encryption of ciphertext C. As before, C is public and R_9 can be guessed based on k_9 hypotheses. The leakage L can be modeled as: $L = HW(R_9) + N$, where N is the noise in the measurement. Apart from the leakage computation, the rest of the attack settings remain the same.

2.4. Write/Read Trace Generation

For key extraction from emerging NVMs using DPA attack, 5000 write traces using the same secret key with different plaintexts for all the 10 AES-128 rounds were generated. As NVM write current depends on previous stored data, when writing to the memory array, the initial value of all 128 bits (i.e., for R_0) were kept as 0, and for the following AES rounds, the stored data are written with new round outputs. Similarly, 5000 read traces were generated when reading R_9 output stored in the memory.

3. Case Study: STTRAM

3.1. Basics of STTRAM

STTRAM bitcell (Figure 2a) contains one MTJ and one access (NMOS) transistor. MTJ has two ferromagnetic layers separated by an oxide layer. They are known as Pinned Layer (PL) and Free Layer (FL). If the magnetic orientation of the layers are parallel to each other, the MTJ resistance is low. If the orientation are antiparallel to each other, the MTJ resistance is high. Typically, data '1' and '0' are represented by the high and low resistance, respectively. Magnetic orientation of FL can be toggled from the one state to another by passing the write current (>critical current) from Sourceline (SL) to Bitline (BL) for data '0' to '1' (or vice versa). Figure 2b shows the energy barrier between two stable state of MTJ.

A 22 nm NMOS and one MTJ based on [28] are used for the simulation. Various simulation parameters used in this work are summarized in Table 1.

3.2. Asymmetric Write and Read Current

STTRAM write current is a function of the polarity of the stored data. The equivalent resistance of MTJ is low (high) in state '0' ('1'). Figure 3a shows I_{write} for 1→0. During write, the initial current is high since MTJ resistance is low. However, the current goes low after successful write. Similarly, write current for 1→0 goes from low to high (Figure 3b).

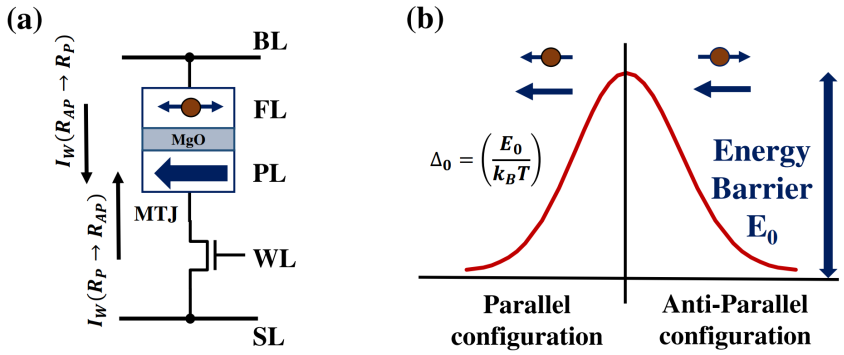


Figure 2. (**a**) STTRAM bitcell; (**b**) MTJ two states separated by an energy barrier [25].

Table 1. STTRAM Simulation Parameters used in this work.

Parameter	Value
Access transistor (nMOS) (W/L)	100 nm/30 nm
MTJ FL volume	1.04×10^{-17} cm^3
MTJ anisotropy (uniaxial), K_u	150,150 erg/cc
Magnetization saturation, M_s	790 Oe
MTJ anisotropic magnetic field, H_k	380 Oe
MTJ Thermal barrier, Δ	37.99
Tunnel magnetoresistance, TMR	2
Write/read latency	3 ns/1 ns
nMOS/pMOS width ratio (read circuit)	1/6

Note that three phases can be identified in the waveform for I_{write} of 0→1 and 1→0; 'Old data', 'Transition from old data to new data' and finally, 'New data'. I_{write} of 1→1 and 0→0 are fairly constant since the MTJ state does not change. The current magnitude difference of low and high states of current waveform (for 0→1 and 1→0) is significant and therefore, reveals the information about new and previous data.

STTRAM read current (Figure 3c) also depends on the present state of the bit. The amplitude difference between two read currents depend on the Tunnel Magneto Resistance [16] of MTJ. For a good sense margin during read, a higher TMR is desirable. However, this adversely affects the data security.

3.3. SCA on STTRAM Write Operation

Figure 4a shows the results to extract the first key bytes from write operation of STTRAM. The figure presents the correlation evolution for the entire key candidates against the trace number. SCA is successful when the correct key (denoted by black line) emerges out from the band of all wrong keys (denoted by grey). The first key byte extraction takes around 800 traces. It could be concluded that signal to noise ratio, SNR of STTRAM write current is low, which makes SCA harder. Similar result is obtained for other key bytes (with minor statistical discrepancy). With the given 2000 traces, it is possible to recover only 8 bytes from 2000 traces. All key bytes are expected to be retrieved with more measurements. However, it is evident that the SCA performance is poor. The reason behind sub-optimal SCA result either could be poor attack setting or low SNR. We can rule out low SNR since simulation traces are very precise. Therefore, SCA on STTRAM write operation provides a false impression of higher resiliency compared to SRAM [25] since the attack is not developed considering the underlying technology.

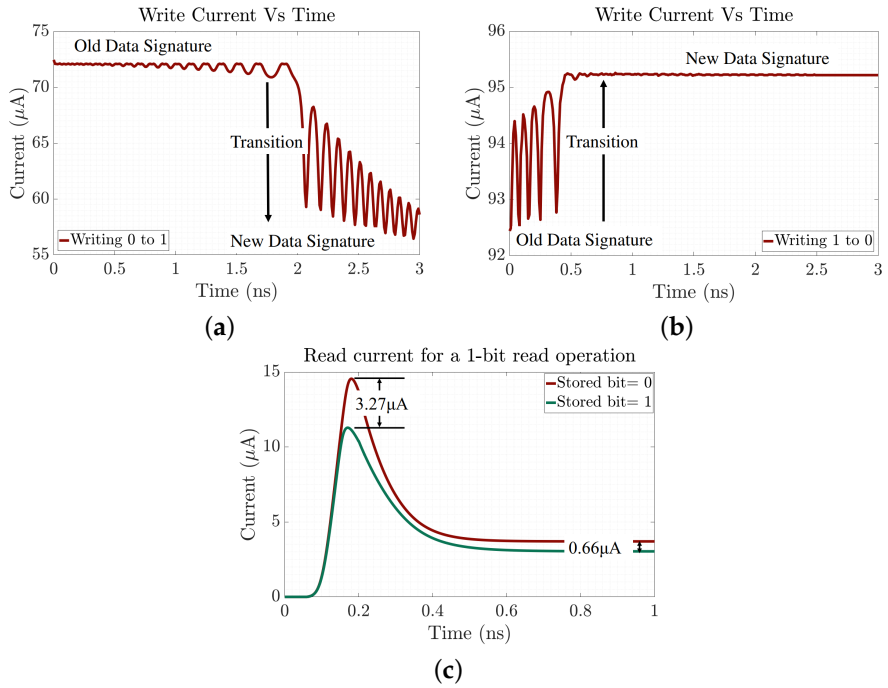


Figure 3. (**a**) I_{write} for 0→1, (**b**) I_{write} for 1→0 and (**c**) I_{read} for '0' and '1'. Write current magnitude and waveform is a function of stored data and a distinctive gap is present between write '1' and '0' and as well as read '1' and '0' currents, which can be leveraged as side channel signature [25].

Improved Attack: We have tried to pre-process the traces since the difference in write current between the final and initial phase is analogous to a switching activity and thereby, should correlate with HD model. Therefore, we generated new traces from the old ones by simply subtracting the current values at final time samples from the initial time samples. The choice of time samples was made considering the attack settings. This type of pre-processing also compresses the traces and thereby, accelerating the SCA computation.

The result of SCA on the pre-processed traces is shown in Figure 4b. The first key byte can be extracted with 300 traces whereas the non-pre-processed version required 800 traces. Note that the entire key is retrieved by 1600 traces. Therefore, we can conclude that the efficiency of attack has significantly improved.

3.4. SCA on STTRAM Read Operation

Further investigation is done to identify the vulnerability of the read. The main difference between STTRAM write vs. read operation is that a constant read voltage is applied across the bitcell. The cell draws current based on the stored data which is sensed to determine the data. In case of read, SCA can exploit the time window when R_9 is read in order to compute the final ciphertext. Therefore, the read current leakage model is: $L = HW(R_9) + N$.

The rest of the SCA setting is similar. Figure 4c summarizes the result. The attack is very efficient and can extract the first key byte in 40 traces and the entire key extraction needs only 400 traces.

In summary, our investigation reveals that although the characteristics of side channel leakage of STTRAM are different from SRAM, the distinction cannot be considered as a protection. By tweaking the attack setting and applying very basic pre-processing, similar exploitability can be found. Furthermore, read current revealed a high vulnerability. Therefore, designers can target efficient countermeasures to emphasize securing the read operation.

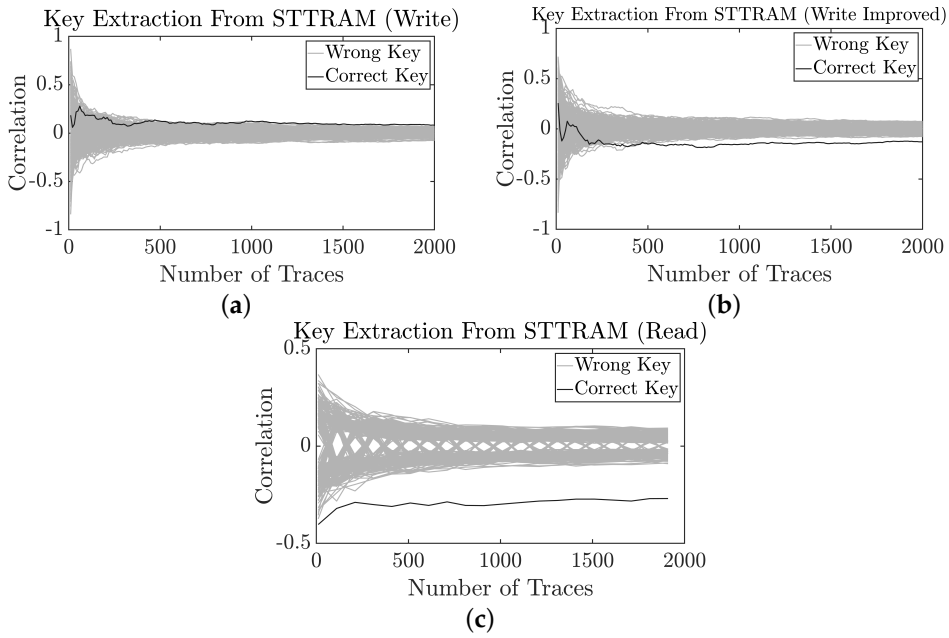


Figure 4. Attack result for (**a**) STTRAM (write); (**b**) STTRAM (write) after pre-processing; (**c**) STTRAM (read). An attack is considered successful when the correct key (black) emerges from the cloud of all wrong keys (grey) [25].

4. Case Study: MRAM

4.1. Basics of MRAM

MRAM cell (Figure 5a) is consist of one MTJ and one nMOS access transistor (similar to STTRAM bitcell (Figure 2a)). However, the MRAM MTJ lies between a couple of metal lines known as bit-line (BL) and digit-line (DL). The BL and DL are orthogonal to each other, parallel to the cell plane, one above and one below the MTJ. The write current is passed through BL and DL with appropriate polarity which induces a magnetic field and exerts a torque on the FL magnetic orientation. This flips the FL magnetic orientation. The access transistor has no part in the write operation. Although MRAM write is magnetic field driven and STTRAM write is current driven, MRAM and STTRAM read operations are similar. Therefore, MRAM read current also shows the data dependency since the asymmetry of the read is due to the difference of the MTJ resistance of parallel/anti-parallel states. Ref. [25] evaluated a DPA based SCA on MRAM read. For this study, a MRAM commercial chip [34] is used. The features of the chip is shown in Table 2.

Table 2. MRAM Chip Characteristic.

Parameter	Value
Capacity	16 Mbit
Write/Read latency	35 ns
Data/Address bus length	8/21
V_{DD}	3.3 V
Data retention time	>20 years
AC standby current	9–14 mA
AC active current (read)	60–68 mA
AC Active Current (write)	152–180 mA
Output enable access time	15 ns

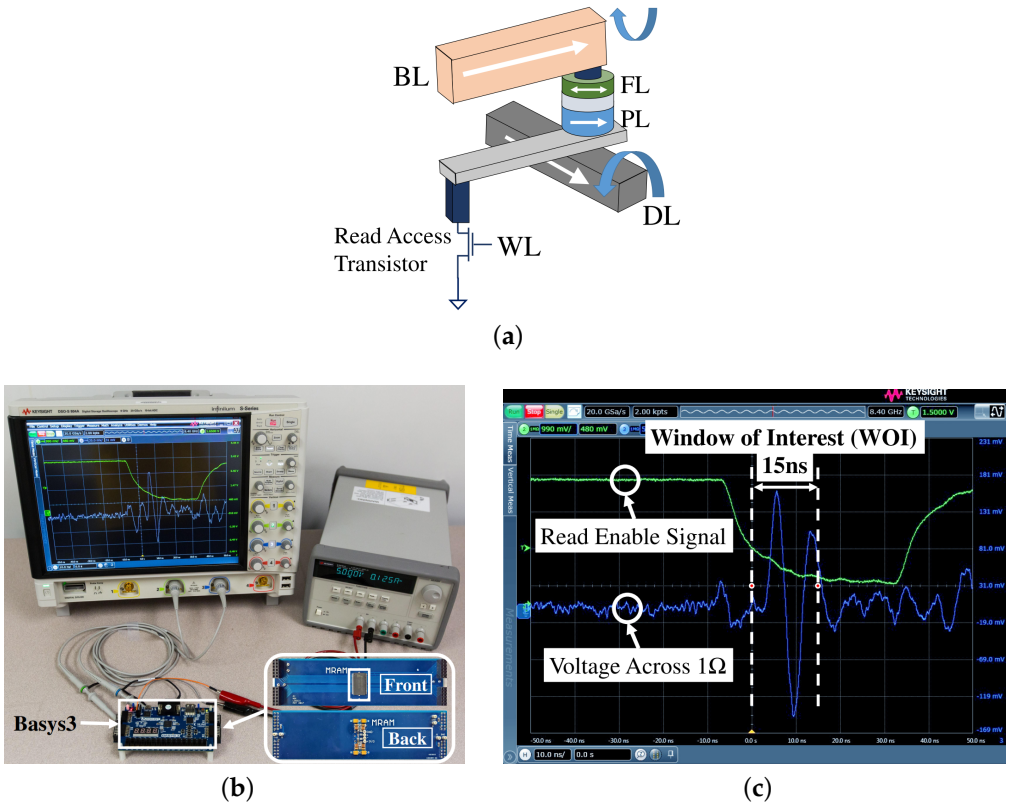


Figure 5. (**a**) MRAM bitcell; (**b**) experimental setup showing the MRAM chip, the PCB (inset) and the FPGA; (**c**) snapshot of the oscilloscope capture: green line shows the read enable signal and the blue line shows the voltage across 1 Ω resistor [25].

4.2. Experimental Setup

MRAM chip [34] was interfaced with a Basys 3 [35] board with an Artix-7 FPGA [36] using a PCB with four layers. The design of the PCB optimized the wire capacitance, resistance and stray inductance to keep the read signature integrity consistent. A capacitor with ~19 µF was used between the GND and V_{dd} to keep the DC supply power stabilized. A current shunt with 1 Ω resistance was connected between the MRAM chip ground and the GND of the PCB to measure the current drawn by the chip as shown in Figure 1. The traces for read current were captured by a Keysight DSOS-804A High Definition Oscilloscope [37] with a bandwidth of 8 GHz and a sampling frequency of 20 GSa/s. Figure 5b shows the experimental setup. Note, that the FPGA frequency is 100 MHz and the MRAM minimum read cycle is 35 ns. Therefore, we have divided the 100 MHz clock frequency by 4 and thereby, implementing 40 ns for read/write cycle time.

4.3. SCA on MRAM Write Operation

In [27], a Correlation Power Analysis (CPA) on MRAM write operation has been performed. The work proposes a hypothetical power model that considers the difference of $0 \rightarrow 1$ and $1 \rightarrow 0$ transitions to estimate the post-alignment power consumption while writing to MRAM. They considered the stream cipher MICKEY-128 2.0 to validate the proposed attack model. The results show that the secret key can be retrieved from MRAM write operation traces.

4.4. SCA on MRAM Read Operation

The data bus length of MRAM is 8. Therefore, data 0 to 255 can be read in one read cycle. Data 0 to 255 was written sequentially on a fixed address and 20 read operations were performed to capture the read current waveform for each data sample. A total of 15 ns is required to enable the output after the read enable signal is asserted (timing reference

voltage is 1.5 V [34]). This indicates that the actual read sensing is done in 15 ns although the read cycle is 40 ns. Similar to the method proposed in [23] where read/write signature for SRAM is experimentally found by taking the average voltage at a Point of Interest (POI), we have calculated the average current for the Window of Interest (WOI) for each data that is being read from MRAM. A single read waveform captured on the oscilloscope is shown in Figure 5c. The start of WOI is considered when the read enable signal (green) crosses 1.5 V since the timing reference voltage is 1.5 V [29]. The average read current in the proposed WOI for data 0 to 255 is plotted in Figure 6a. It is clear that the average I_{read} in the WOI depends on the number of ones in 8-bit read data. The average current reduces with the increasing HW of the 8-bit data. This proves that MRAM read current which is asymmetric in nature reveals the sensitive read data.

Similar to the simulation case, DPA-based SCA is implemented on MRAM read by exploiting the time window when R_9 is read to compute the final cipher text. Therefore, the read signature leakage model is: $L = HW(R_9) + N$. However, a restriction is only 8 bit data can be read in one cycle from the MRAM chip. Therefore, R_9 round output is read in total of 16 cycles. However, the read current in simulation is measured for all 16 bytes. The read current corresponding to the byte under attack is the signal of interest since we attack one byte at a time. Therefore, the current related to the other 15 bytes is considered as noise. For MRAM chip, we read only one byte in each cycle and therefore, noise related to other 15 bytes become zero. This leads to 15× less (algorithmic) noise in the experimental measurement compared to the case in simulation.

The correlation against the time samples within the WOI is shown in Figure 6b. The correlation for the correct key hypothesis clearly stands out from all the wrong key hypotheses, confirming the practical side channel vulnerability of MRAM read current. Note that 15 traces were enough to perform the SCA (Figure 6c, correlation vs. the number of traces). A successful SCA with 15 traces is faster compared to 40 traces for STTRAM in simulated setting (Section 3.4). This is because the experimental measurements contained 15× less algorithmic noise.

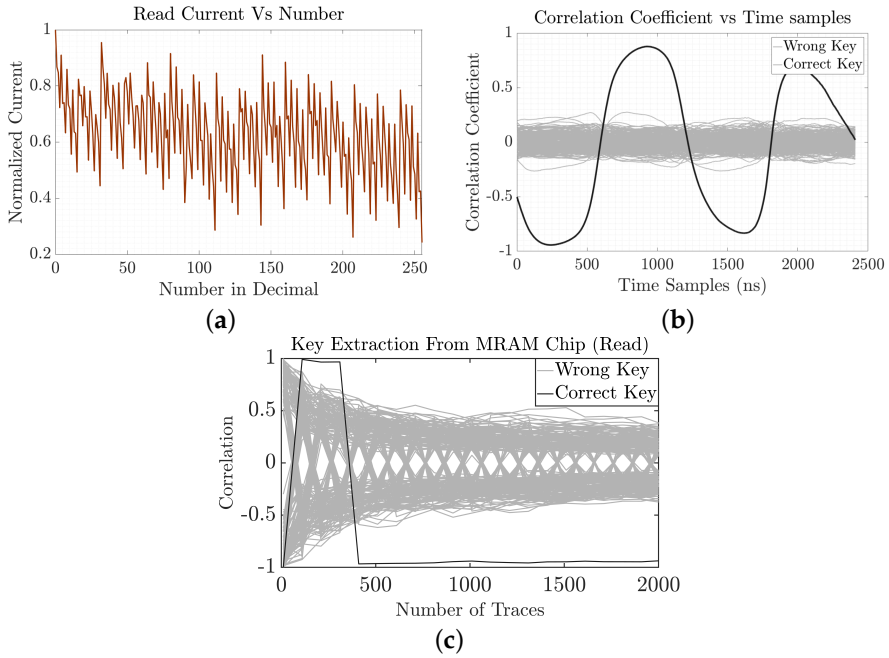


Figure 6. (**a**) Average current at Window of Interest (WOI) with respect to the data being read from the MRAM chip; (**b**) correlation with respect to time samples; (**c**) correlation with respect to the number of traces captured during MRAM read [25].

5. Case Study: RRAM

5.1. Basics of RRAM

RRAM storage element is mainly an oxide material between two electrodes namely, Top Electrode (TE) and Bottom Electrode (BE) (Figure 7a). The filament between the electrodes can be formed or ruptured based on the direction and magnitude of the electric field through it. If a filament is formed between the two electrodes, the resistance of the cell is low (Low Resistance State, LRS) and that can be considered as data '0'. However, if the filament is ruptured, the resistance of the cell is high (High Resistance State, HRS) and that can be considered as data '1'.

5.2. Asymmetric Read/Write Current

Figure 7b shows RRAM write current for writing data 0→1, 0→0, 1→0 and 1→1. Similar to STTRAM, current is almost constant for writing data 0→0 and 1→1 while the new data, old data and the transition regions are distinguishable for writing data 0→1 and 1→0. Therefore, the total write current for writing a full data word is a function of data pattern. Similarly, RRAM read current is also asymmetric (Figure 7c). The average current drawn by the bitcell is more if the stored datum is '0'.

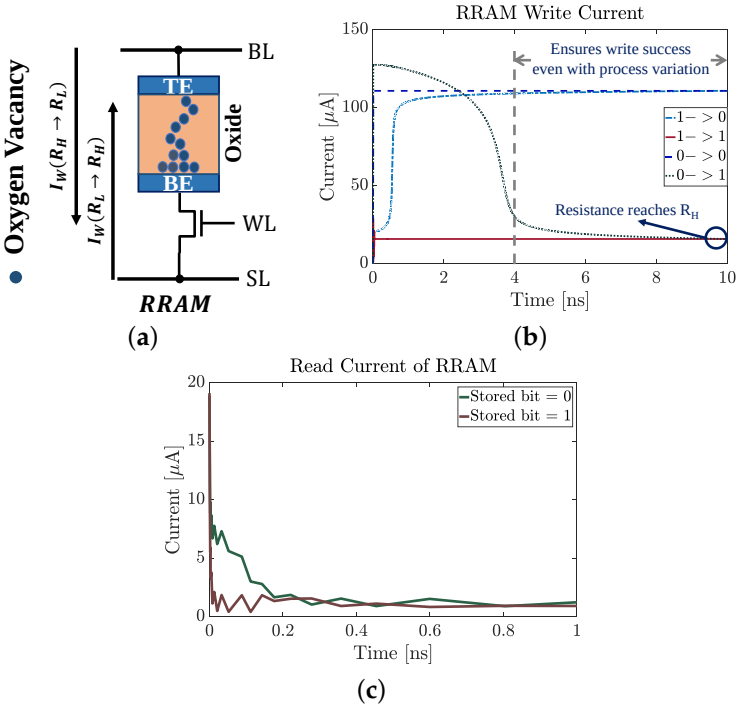


Figure 7. RRAM (**a**) bitcell; (**b**) asymmetric write current; and (**c**) asymmetric read current.

A 65 nm nMOS along with ASU (bipolar HfO_x-based) RRAM Verilog-A model [38] is used in this work. Simulation parameters are provided in Table 3.

Table 3. Parameters used for RRAM Simulation.

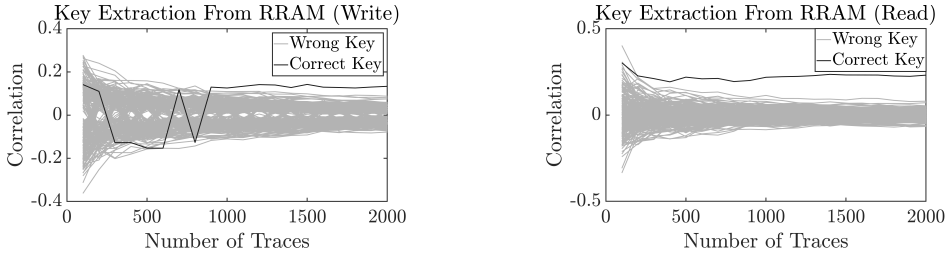
Parameter	Value
Access Transistor W/L/V_T	195 nm/65 nm/0.423 V
RRAM Gap for R_L/R_H	0.53 nm/1.368 nm
Unit Cell Size	12 F^2
System Clock Frequency/V_{dd}	2 GHz/2.2 V
Read/Write Latency	0.5 ns (1 cycle)/10 ns (20 cycle)

5.3. SCA on RRAM Write Operation

The attack model follows the description provided in Section 2.3, i.e., HD leakage model with Pearson Correlation. Figure 8a demonstrate the attack results to extract the first byte of the key from RRAM write operation. The attack is efficient and can reveal the correct key in only 900 traces.

5.4. SCA on RRAM Read Operation

The attack follows the description provided in Section 2.3. Figure 8b demonstrate the attack results to extract the first byte of the key from RRAM read operation. The attack is efficient and can reveal the correct key in only 200 traces.



(**a**) (**b**)

Figure 8. Attack result from RRAM (**a**) write current and, (**b**) read current.

6. Case Study: PCM

6.1. Basics of PCM

There are two major PCM designs (Figure 9a); heater-based and self-heating-based [39]. The first one contains a material layer (e.g., titanium nitride or tungsten [39–41]) that act as a heat source to heat the adjacent layer of phase-change material [39]. The later one relies on the internal generated heat within the PCM and helps the state change of the material [39]. $Ge_2Sb_2Te_5$ (GST) [40,42] is used as a phase-change material for both designs. Another example of a similar phase-change material is $In_3Sb_1Te_2$ [43].

A current is applied through the PCM cell to force the cell to either SET (low resistance) or RESET (high resistance) during write. A small voltage is applied across the cell during read operation and the resistance is sensed to read the stored data.

6.2. Asymmetric Read and Write Current

Figure 9b shows PCM write current for writing data 0→1, 0→0, 1→0 and 1→1. Unlike STTRAM and RRAM, the write current of PCM does not depend on the old data, rather depends on the new data that is being written to the cell. This reduces the 4 write current combinations to 2 which an adversary can leverage to launch a stronger SCA attack. Similarly, PCM read current is also asymmetric (Figure 9c). The current drawn by the bitcell is significantly more if the stored datum is '0' compared to case of '1'.

A 65 nm nMOS along with ASU RRAM Verilog-A model [44] is used for analysis. Simulation parameters are provided in Table 4.

Table 4. Parameters used for PCM Simulation.

Parameter	Value
Access Transistor W/L/V_T	195 nm/65 nm/0.423 V
Bottom Contact Width, CW	28 nm
GST thickness	49 nm
$R_{SET}/R_{RESET}/R_{WRITE}$	9 kΩ/3.6 MΩ/1 kΩ
Read/Write Latency	20 ns (40 cycle)/150 ns (300 cycle)

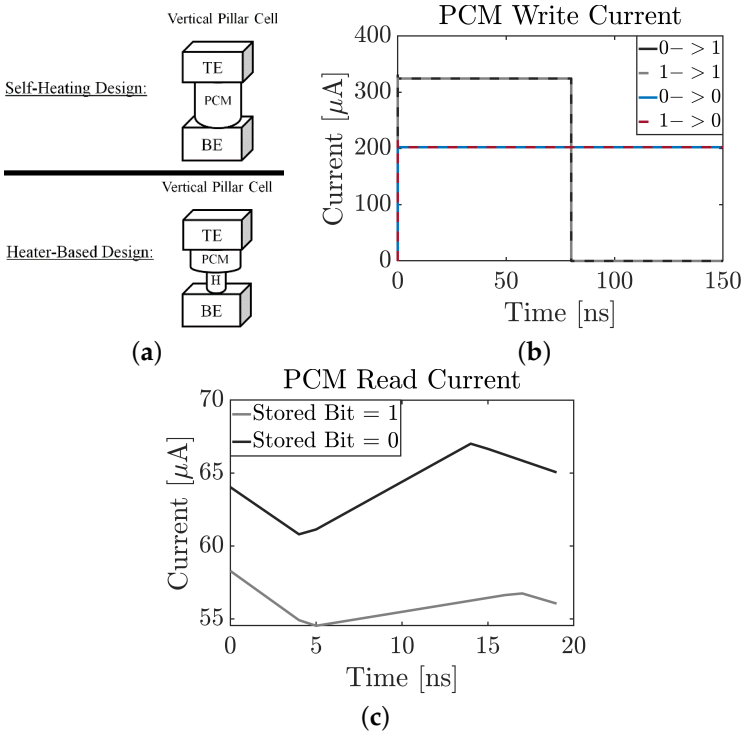


Figure 9. PCM (**a**) bitcell: (**b**) asymmetric write current and (**c**) asymmetric read current.

6.3. SCA on PCM Write Operation

Figure 10a demonstrates the results to extract the first key byte from PCM write operation. The attack is successful, efficient and only takes 200 traces to reveal the key. Furthermore, the full key can be extracted in 400 traces.

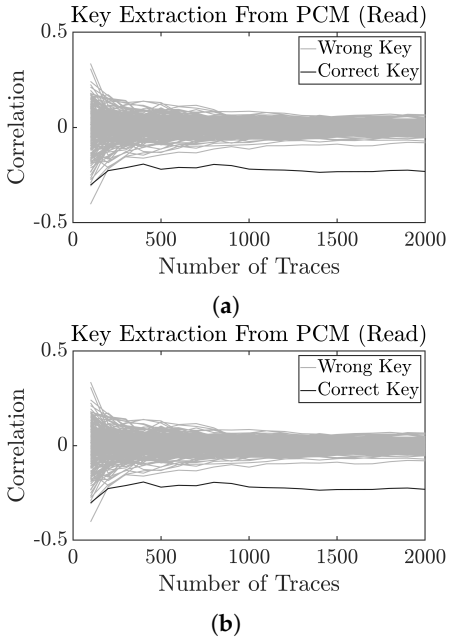


Figure 10. Attack result from PCM: (**a**) write current; (**b**) read current.

6.4. SCA on PCM Read Operation

Figure 10b demonstrates the results to extract the first key byte from PCM read operation. The attack is successful, efficient and only takes 200 traces to reveal the key.

6.5. Comparative Analysis of NVM Vulnerability to SCA

Finally, we compare all the previous results in Table 5. While STTRAM (write) and RRAM (write) offer the highest resistance, it is obvious that all the tested memory technologies are sooner or later vulnerable. This brings the need of strong countermeasures. In the following section, we test some commonly used countermeasures or mitigation techniques from the side-channel domain.

Table 5. Comparative Analysis of SCA Vulnerability of NVMs w.r.t their side-channel resistance in terms of minimum number of traces to disclose (MTD) first byte of the key.

NVM	MTD
STTRAM (write)	300
STTRAM (read)	40
MRAM (read)	15
RRAM (write)	900
RRAM (read)	200
PCM (write)	200
PCM (read)	200

7. Analysis of Encoding as a Mitigation Technique

The authors of [45] propose that encoding could obfuscate the data signature and make the key extraction difficult. Therefore, we have performed four types of encoding for write operation and two types of encoding for read operation to analyze the impact of encoding. We have considered RRAM as a test case.

- Write Encoding Try 1: Out of 128-bit write data, first MSB 8 bits are encoded with reverse polarity. This means that for those 8 bits, high resistance state is considered as data '0' and low resistance state is considered as data '1'. The key extraction result is shown in Figure 11a. It is evident that the attack is successful and the first byte of the key can be retrieved in roughly 800 traces.
- Write Encoding Try 2: Out of 128-bit write data, first MSB 16 bits are encoded with reverse polarity. Figure 11b shows the corresponding successful attack result where the first byte of the key can be retrieved in roughly 950 traces.
- Write Encoding Try 3: Out of 128-bit write data, first MSB 32 bits are encoded with reverse polarity. The attack result is summarized in Figure 11c. The first byte of the key can be retrieved in roughly 600 traces.
- Write Encoding Try 4: Out of 128-bit write data, first MSB 64 bits are encoded with reverse polarity. Figure 12a shows that the first byte of the key can be retrieved in roughly 600 traces.
- Read Encoding Try 1: Out of 128-bit read data, first MSB 32 bits are encoded with reverse polarity. Figure 12b shows that the first byte of the key can be retrieved in roughly 350 traces.
- Read Encoding Try 2: Out of 128-bit read data, first MSB 64 bits are encoded with reverse polarity. The attack result is summarized in Figure 12c. We note that the first byte of the key can be retrieved in roughly 200 traces.

Without encoding, it took 900 traces to extract the first byte of the key from RRAM write operation. After encoding, 950 traces are required (with Try 2). Therefore, we conclude that encoding does improve the resistance but is not an effective mitigation technique. A similar conclusion can be drawn for encoding on read operation.

The underlying reason for such vulnerability is emerging NVMs basically use different resistance states to store different data. Therefore, there would always be a minuscule difference between data '0' and '1' read/write currents. Thus, the mitigation needs to

identify, (possibly by profiling) if a better codeword exists that reduces the difference (e.g., 0001 and 1110). Even then, with large number of traces the attack might be feasible. These results are in line with that on SRAM as shown in [46].

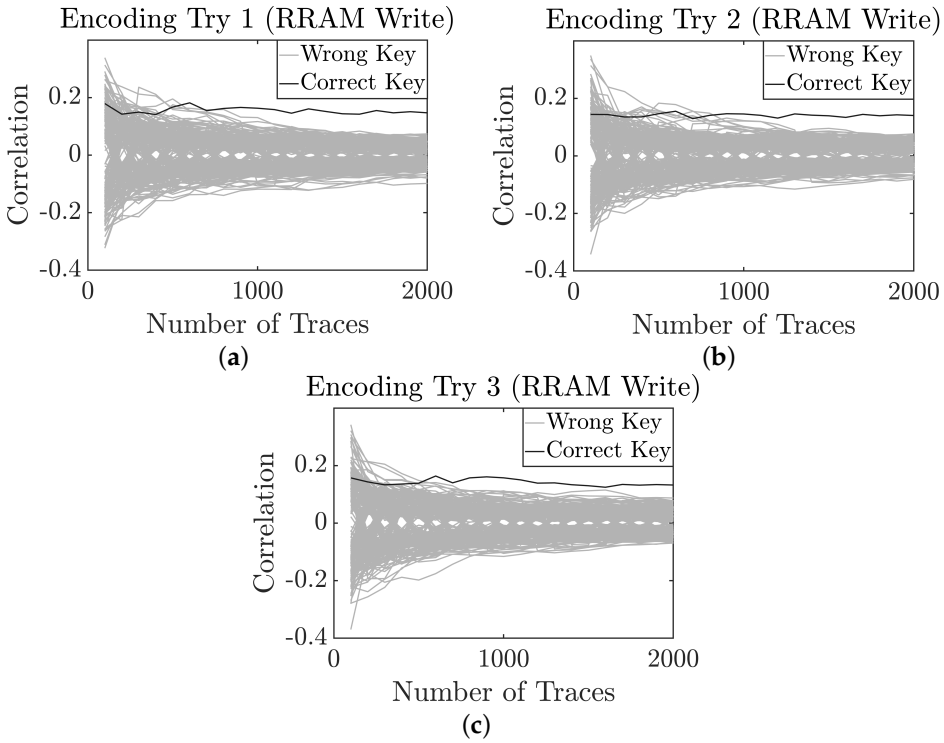


Figure 11. Key extraction from RRAM write operation after implementation of data (**a**) Encoding Try 1; (**b**) Encoding Try 2; and (**c**) Encoding Try 3.

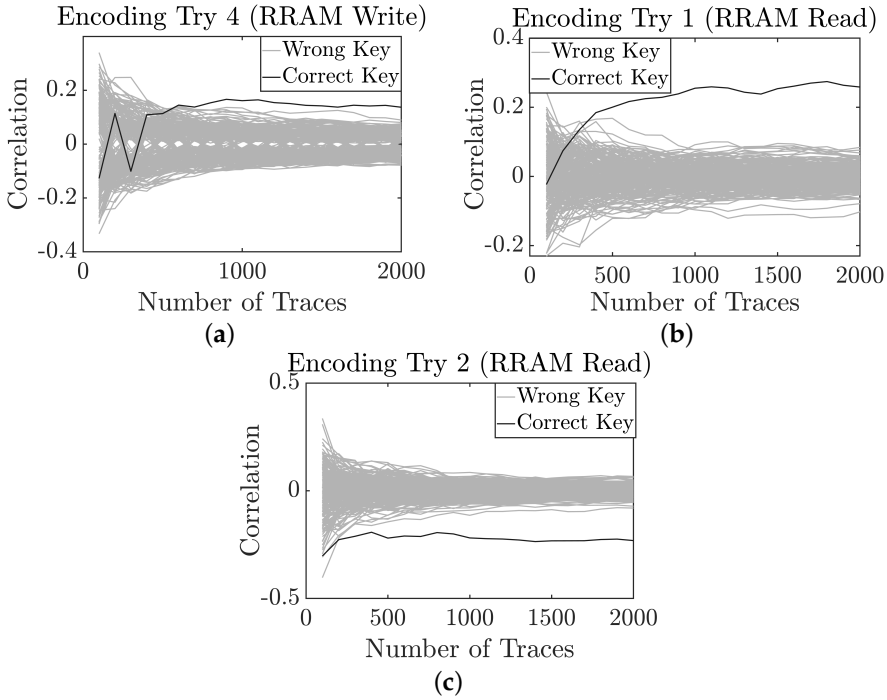


Figure 12. Key extraction from RRAM write operation after implementation of data (**a**) Encoding Try 4; Key extraction from RRAM read operation after implementation of data (**b**) Encoding Try 1; and (**c**) Encoding Try 2.

8. Discussion

8.1. Assumptions Used in This Work

Following assumptions are made in this work: (i) the current drawn by the LLC can be isolated from the total system current. Note that the total system current will also include the CPU current. However, the adversary can filter out the high frequency components of the total current by applying signal processing [47] since the CPU current frequency is much higher compared to cache current (long NVM read and write latency); (ii) the proposed SCA model is independent of cache size. This assumption is true as long as the cache is large enough to hold the round keys. The unused cache will only contribute to constant leakage which can be filtered out; (iii) the CPUs considered are laptop and PCs although the attack can be investigated for Internet of Things (IoTs) as well; and (iv) AES is the cryptographic application. Other algorithms such as MICKEY 2.0 (shown in [27]) are also vulnerable.

8.2. Considerations for Improving SCA Efficiency

Note that asymmetric NVM read and write current can serve as two isolated side channels to steal the encryption key. The attack can be improved by noting that read and write are performed on the same key during AES rounds. Thus, the attack accuracy and speed of key extraction can be improved by a cross-correlation between asymmetric read and write currents.

8.3. Considerations for SCA Resiliency

To weaken SCA, the data dependent asymmetry in read/write current should be eliminated or masked. In [48], a technique to eliminate side channel signature of emerging NVM using on-chip capacitor and very steady and robust voltage regulator [49] is proposed. The on-chip capacitor hides the side channel from being captured at the supply voltage and the robust voltage regulator provides steady supply to the capacitor during charging phase. Although the paper shows example of RRAM, the design can be extended to other NVMs as well.

9. Conclusions

In this work, we summarized a thorough study of SCA on various emerging NVMs such as STTRAM, MRAM, RRAM and PCM-based LLC, respectively. We have assumed AES-128 operation being performed on LLC and leveraged the asymmetric write/read current during the AES round operations. Results revealed that the read operation is more susceptible to leak the key although write current showed greater asymmetry than read current. Our investigation also shows that applying basic pre-processing resulted in much improvement of the attack. The proposed attack model is also experimentally validated using a commercial MRAM chip. We have also investigated mitigation techniques proposed in prior works. We conclude that techniques like encoding is not sufficient to protect the key and we need more device level solutions to hide the data signature.

Author Contributions: Conceptualization, M.N.I.K., S.B., A.C. and S.G.; methodology, M.N.I.K., S.B., A.C. and S.G.; software, M.N.I.K. and S.B; validation, M.N.I.K. and S.B.; formal analysis, M.N.I.K. and S.B.; investigation, M.N.I.K. and S.B.; resources, S.G.; data curation, M.N.I.K., S.B., B.L. and A.Y.; writing—original draft preparation, M.N.I.K. and S.B.; writing—review and editing, A.C. and S.G.; visualization, M.N.I.K. and S.B.; supervision, A.C. and S.G.; project administration, A.C. and S.G.; funding acquisition, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Semiconductor Research Corporation (SRC) (2847.001), National Science Foundation (NSF) (CNS-1722557, CCF-1718474, DGE-1723687 and DGE-1821766) and DARPA Young Faculty Award (D15AP00089).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Worledge, D.C.; Hu, G.; Trouilloud, P.L.; Abraham, D.W.; Brown, S.; Gaidis, M.C.; Nowak, J.; O'Sullivan, E.J.; Robertazzi, R.P.; Sun, J.Z.; et al. Switching distributions and write reliability of perpendicular spin torque MRAM. In Proceedings of the 2010 International Electron Devices Meeting, San Francisco, CA, USA, 6–8 December 2010; pp. 12.5.1–12.5.4. [CrossRef]
2. Lee, T.Y.; Yamane, K.; Hau, L.Y.; Chao, R.; Chung, N.L.; Naik, V.B.; Sivabalan, K.; Kwon, J.; Lim, J.H.; Neo, W.P.; et al. Magnetic Immunity Guideline for Embedded MRAM Reliability to Realize Mass Production In Proceedings of the 2020 IEEE International Reliability Physics Symposium (IRPS), Dallas, TX, USA, 28 April–30 May 2020. [CrossRef]
3. Nigam, A.; Smullen, C.W.; Mohan, V.; Chen, E.; Gurumurthi, S.; Stan, M.R. Delivering on the promise of universal memory for spin-transfer torque RAM (STT-RAM). In Proceedings of the IEEE/ACM International Symposium on Low Power Electronics and Design, Fukuoka, Japan, 1–3 August 2011; pp. 121–126. [CrossRef]
4. Baranwal, M.; Chugh, U.; Dalal, S.; Agarwal, S.; Kapoor, H.K. DAMUS: Dynamic Allocation based on Write Frequency in MUlti-Retention STT-RAM based Last Level Caches In Proceedings of the 2021 22nd International Symposium on Quality Electronic Design (ISQED), Santa Clara, CA, USA, 7–9 April 2021. [CrossRef]
5. Kuan, K.; Adegbija, T. A Study of Runtime Adaptive Prefetching for STTRAM L1 Caches In Proceedings of 2020 IEEE 38th International Conference on Computer Design (ICCD), Hartford, CT, USA, 18–21 October 2020. [CrossRef]
6. Pirovano, A.; Lacaita, A.L.; Pellizzer, F.; Kostylev, S.A.; Benvenuti, A.; Bez, R. Low-field amorphous state resistance and threshold voltage drift in chalcogenide materials. *IEEE Trans. Electron Devices* **2004**, *51*, 714–719. [CrossRef]
7. Gong, H.; Ume, R.; Tokranov, V.; Yakimov, M.; Sadana, D.; Brew, K.; Cohen, G.; Schujman, S.; Beckmann, K.; Cady, N.; et al. Bilayer Ga-Sb Phase Change Memory with Intermediate Resistance State In Proceedings of the 2021 Device Research Conference (DRC), Santa Barbara, CA, USA, 20–23 June 2021. [CrossRef]
8. Wu, Y.; Yu, S.; Guan, X.; Wong, H.S.P. Recent progress of resistive switching random access memory (RRAM). In Proceedings of the 2012 IEEE Silicon Nanoelectronics Workshop (SNW), Honolulu, HI, USA, 10–11 June 2012; pp. 1–4. [CrossRef]
9. Xu, M.; Gao, B.; Xu, F.; Wu, W.; Tang, J.; Chen, J.; Qian, H. A Compact Model of Analog RRAM Considering Temperature Coefficient for Neural Network Evaluation In Proceedings of the 2021 5th IEEE Electron Devices Technology & Manufacturing Conference (EDTM), Chengdu, China, 8–11 April 2021. [CrossRef]
10. Chen, A. A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electron.* **2016**, *125*, 25–38.10.1016/j.sse.2016.07.006. [CrossRef]
11. Xue, C.J.; Sun, G.; Zhang, Y.; Yang, J.J.; Chen, Y.; Li, H. Emerging non-volatile memories: Opportunities and challenges. In Proceedings of the 2011 Proceedings of the Ninth IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), Taipei, Taiwan, 9–14 October 2011; pp. 325–334. [CrossRef]
12. De, A.; Khan, M.N.I.; Park, J.; Ghosh, S. Replacing eFlash with STTRAM in IoTs: Security Challenges and Solutions. *J. Hardw. Syst. Secur.* **2017**, *1*, 328–339. [CrossRef]
13. Intel Optane Memory Series. 2015. Available online: https://ark.intel.com/products/97544/Intel-Optane-Memory-Series-16GB-M_2-80mm-PCIe-3_0-20nm-3D-Xpoint (accessed on 3 May 2018).
14. Ghosh, S.; Khan, M.N.I.; De, A.; Jang, J.W. Security and privacy threats to on-chip Non-Volatile Memories and countermeasures. In Proceedings of the 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), Austin, TX, USA, 7–10 November 2016; pp. 1–6. [CrossRef]
15. Khan, M.N.I.; Ghosh, S. Fault Injection Attacks on Emerging Non-volatile Memory and Countermeasures. In Proceedings of the 7th International Workshop on Hardware and Architectural Support for Security and Privacy, HASP '18, Los Angeles, CA, USA, 2 June 2018; ACM: New York, NY, USA, 2018; pp. 10:1–10:8. [CrossRef]
16. Diao, Z.; Li, Z.; Wang, S.; Ding, Y.; Panchula, A.; Chen, E.; Wang, L.C.; Huai, Y. Spin-transfer torque switching in magnetic tunnel junctions and spin-transfer torque random access memory. *J. Phys. Condens. Matter* **2007**, *19*, 165209. [CrossRef]
17. Shamsi, K.; Jin, Y. Security of emerging non-volatile memories: Attacks and defenses. In Proceedings of the 2016 IEEE 34th VLSI Test Symposium (VTS), Las Vegas, NV, USA, 25–27 April 2016; pp. 1–4. [CrossRef]
18. Sugawara, T.; Suzuki, D.; Saeki, M.; Shiozaki, M.; Fujino, T. On measurable side-channel leaks inside ASIC design primitives. In *International Workshop on Cryptographic Hardware and Embedded Systems*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 159–178.
19. Shan, W.; Chen, X.; Li, B.; Cao, P.; Li, J.; Gao, G.; Shi, L. Evaluation of Correlation Power Analysis Resistance and Its Application on Asymmetric Mask Protected Data Encryption Standard Hardware. *IEEE Trans. Instrum. Meas.* **2013**, *62*, 2716–2724. [CrossRef]
20. Lazzaroni, M.; Piuri, V.; Maziero, C. Computer security aspects in industrial instrumentation and measurements. In Proceedings of the 2010 IEEE Instrumentation Measurement Technology Conference Proceedings, Austin, TX, USA, 3–6 May 2010; pp. 1216–1221. [CrossRef]
21. Bilski, P.; Winiecki, W.; Adamski, T. Implementation of symmetric cryptography in embedded systems for secure measurement systems. In Proceedings of the 2011 IEEE International Instrumentation and Measurement Technology Conference, Hangzhou, China, 10–12 May 2011; pp. 1–6. [CrossRef]
22. Wu, J.; Shi, Y.; Choi, M. Measurement and Evaluation of Power Analysis Attacks on Asynchronous S-Box. *IEEE Trans. Instrum. Meas.* **2012**, *61*, 2765–2775. [CrossRef]

23. Fong, X.; Choday, S.H.; Roy, K. Design and optimization of spin-transfer torque mrams. In *More than Moore Technologies for Next Generation Computer Design*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 49–72.
24. Konstantakos, V.; Kosmatopoulos, K.; Nikolaidis, S.; Laopoulos, T. Measurement of Power Consumption in Digital Systems. *IEEE Trans. Instrum. Meas.* **2006**, *55*, 1662–1670. [CrossRef]
25. Khan, M.N.I.; Bhasin, S.; Yuan, A.; Chattopadhyay, A.; Ghosh, S. Side-Channel Attack on STTRAM Based Cache for Cryptographic Application. In Proceedings of the 2017 IEEE International Conference on Computer Design (ICCD), Boston, MA, USA, 5–8 November 2017; pp. 33–40. [CrossRef]
26. Lee, D.; Gupta, S.K.; Roy, K. High-performance low-energy STT MRAM based on balanced write scheme. In Proceedings of the 2012 ACM/IEEE International Symposium on Low Power Electronics and Design, Redondo Beach, CA, USA, 30 July–1 August 2012; ACM: New York, NY, USA, 2012; pp. 9–14.
27. Chakraborty, A.; Mondal, A.; Srivastava, A. Correlation Power Analysis Attack against STT-MRAM Based Cyptosystems. *IACR Cryptol. ePrint Arch.* **2017**, *2017*, 413.
28. Srikant, S. All Spin Logic: Modeling Multi-Magnet Networks Interacting via Snin Currents. Ph.D. Dissertation, Purdue University, West Lafayette, IN, USA, 2012.
29. Kocher, P.; Jaffe, J.; Jun, B. Differential power analysis. In *Annual International Cryptology Conference*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 388–397.
30. Bernstein, D.J. Cache-Timing Attacks on AES. 2005. Available online: https://cr.yp.to/antiforgery/cachetiming-20050414.pdf (accessed on 15 September 2021).
31. Gandolfi, K.; Mourtel, C.; Olivier, F. Electromagnetic analysis: Concrete results. In *International Workshop on Cryptographic Hardware and Embedded Systems*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 251–261.
32. Brier, E.; Clavier, C.; Olivier, F. Correlation power analysis with a leakage model. In *International Workshop on Cryptographic Hardware and Embedded Systems*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 16–29.
33. Rauzy, P.; Guilley, S.; Najm, Z. Formally proved security of assembly code against power analysis. *J. Cryptogr. Eng.* **2016**, *6*, 201–216. [CrossRef]
34. MR4A08B Datasheet. Available online: https://www.everspin.com/getdatasheet/MR4A08B (accessed on 13 March 2019).
35. Basys 3 FPGA Board Reference Manual. Available online: https://reference.digilentinc.com/_media/basys3:basys3_rm.pdf (accessed on 13 March 2019).
36. Artix-7 FPGAs Data Sheet. Available online: https://www.xilinx.com/support/documentation/data_sheets/ds181_Artix_7_Data_Sheet.pdf (accessed on 13 March 2019).
37. Infiniium S-Series, The Standard for Superior Measurements Data Sheet. Available online: https://literature.cdn.keysight.com/litweb/pdf/5991-3904EN.pdf?id=2447379 (accessed on 13 March 2019).
38. Chen, P.Y.; Yu, S. Compact Modeling of RRAM Devices and Its Applications in 1T1R and 1S1R Array Design. *IEEE Trans. Electron Devices* **2015**, *62*, 4022–4028. [CrossRef]
39. Boniardi, M.; Redaelli, A.; Cupeta, C.; Pellizzer, F.; Crespi, L.; D'Arrigo, G.; Lacaita, A.L.; Servalli, G. Optimization metrics for Phase Change Memory (PCM) cell architectures. In Proceedings of the 2014 IEEE International Electron Devices Meeting, San Francisco, CA, USA, 15–17 December 2014; pp. 29.1.1–29.1.4. [CrossRef]
40. Russo, U.; Ielmini, D.; Redaelli, A.; Lacaita, A.L. Modeling of Programming and Read Performance in Phase-Change Memories—Part I: Cell Optimization and Scaling. *IEEE Trans. Electron Devices* **2008**, *55*, 506–514. [CrossRef]
41. Servalli, G. A 45nm generation Phase Change Memory technology. In Proceedings of the 2009 IEEE International Electron Devices Meeting (IEDM), Baltimore, MD, USA, 7–9 December 2009; pp. 1–4. [CrossRef]
42. Pellizzer, F.; Pirovano, A.; Ottogalli, F.; Magistretti, M.; Scaravaggi, M.; Zuliani, P.; Tosi, M.; Benvenuti, A.; Besana, P.; Cadeo, S.; et al. Novel /spl mu/trench phase-change memory cell for embedded and stand-alone non-volatile memory applications. In Proceedings of the Digest of Technical Papers. 2004 Symposium on VLSI Technology, 2004, Honolulu, HI, USA, 15–17 June 2004; pp. 18–19. [CrossRef]
43. Kim, E.T.; Lee, J.Y.; Kim, Y.T. Investigation of electrical characteristics of the $In_3Sb_1Te_2$ ternary alloy for application in Phase Change Memory. *Phys. Status Solidi RRL-Rapid Res. Lett.* **2009**, *3*, 103–105. [CrossRef]
44. Xu, Z.; Sutaria, K.B.; Yang, C.; Chakrabarti, C.; Cao, Y. Hierarchical modeling of Phase Change memory for reliable design. In Proceedings of the 2012 IEEE 30th International Conference on Computer Design (ICCD), Montreal, QC, Canada, 30 September–3 October 2012; pp. 115–120. [CrossRef]
45. Maghrebi, H.; Servant, V.; Bringer, J. There is wisdom in harnessing the strengths of your enemy: Customized encoding to thwart side-channel attacks. In *International Conference on Fast Software Encryption*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 223–243.
46. Bhasin, S.; Jap, D.; Peyrin, T. Practical Evaluation of FSE 2016 Customized Encoding Countermeasure. *IACR Trans. Symmetric Cryptol.* **2017**, 2017, 108–129. [CrossRef]
47. Kar, M.; Singh, A.; Mathew, S.; Rajan, A.; De, V.; Mukhopadhyay, S. Exploiting fully integrated inductive voltage regulators to improve side channel resistance of encryption engines. In Proceedings of the 2016 International Symposium on Low Power Electronics and Design, San Francisco, CA, USA, 8–10 August 2016; ACM: New York, NY, USA, 2016; pp. 130–135.

48. Nagarajan, K.; Ahmed, F.U.; Khan, M.N.I.; De, A.; Chowdhury, M.H.; Ghosh, S. SecNVM: Power Side-Channel Elimination Using On-Chip Capacitors for Highly Secure Emerging NVM. *IEEE Trans. Very Large Scale Integr. VLSI Syst.* **2021**, *29*, 1518–1528. [CrossRef]
49. Ahmed, F.U.; Sandhie, Z.T.; Chowdhury, M.H. An Implementation of External Capacitor-less Low-DropOut Voltage Regulator in 45 nm Technology with Output Voltage Ranging from 0.4 V–1.2 V. In Proceedings of the 2020 IEEE 38th International Conference on Computer Design (ICCD), Hartford, CT, USA, 18–21 October 2020; pp. 453–456. [CrossRef]