

Data Remanence Vulnerabilities in Commercial SRAM at Low Temperature

*Farzana Hoque, †Izak Halseide, *Aleksandar Milenkovic and †Biswajit Ray

*Department of Electrical and Computer Engineering, The University of Alabama in Huntsville

†Department of Electrical and Computer Engineering, Colorado State University

Email: *{fh0016, milenka}@uah.edu

†{Izak.Halseide, biswajit.ray}@colostate.edu

Abstract—This paper presents an in-depth analysis of data remanence vulnerabilities in both embedded and standalone Static Random-Access Memory (SRAM). We demonstrate a data remanence attack that leverages prolonged Data Remanence Time (DRT) in low-temperature environments. Our experimental results show nearly perfect data retention of almost 100% for up to 600 milliseconds at -25°C for known images, highlighting the real-world implications of this phenomenon for predicting and mitigating SRAM-based security risks. Finally, we develop a conceptual framework as part of our root cause analysis, linking leakage current behavior to DRT, enabling predictive modeling of DRT.

Index Terms—SRAM, Data Remanence, Low Temperature, Security, Attack

I. INTRODUCTION

SRAM is ubiquitous in CPU caches and embedded systems, often storing sensitive information such as cryptographic keys, passwords, and other confidential data. Being a volatile memory, it is generally assumed that SRAM contents are lost immediately after the power is turned off. However, recent studies indicate that data may persist for microseconds to seconds after power-off but the duration ranging from microseconds to seconds, depends on the SRAM technology node, operating temperature, and other environmental factors [1]–[6]. This phenomenon is known as data remanence. It poses a significant security risk, especially in resource-constrained Internet of Things (IoT) devices, where an adversary might gain physical access to the SRAM chip for a brief period of time. Given the widespread deployment of IoT devices in modern life, data remanence-based attacks on SRAM could cause substantial damage [2], [4], [7]. Therefore, a fundamental understanding of the data remanence of commercial SRAM memory is crucial for assessing the security of SRAM-based computing systems.

In this paper, we evaluate data remanence vulnerabilities in embedded SRAM memory as well as stand-alone SRAM chips. Embedded SRAM memory is evaluated using on-chip SRAM of the 16-bit MSP430F5529LP microcontroller [8]. The stand-alone SRAM chips are evaluated using commercial-off-the-shelf SRAM chips from Cypress Semiconductor. We first demonstrate data-remanence based attack by writing a

This work was supported in part by the National Science Foundation under Grant 2403540, 2423249 and 2346853.

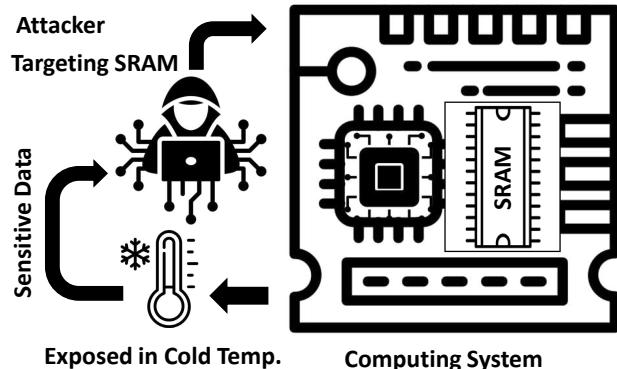


Fig. 1. Conceptual cold boot attack on SRAM.

known pattern (a pre-selected bitmap image) to the SRAM and monitoring its retention over time after performing a power OFF/ON cycle. We examine the effects of operating temperature on data remanence time (DRT) under varying temperatures, from ambient conditions to as low as -25°C. Additionally, we propose a conceptual framework to understand DRT depending on transistor leakage current which provides a predictive model of DRT as a function of temperature and SRAM technology nodes. Finally, we provide a bit-by-bit DRT analysis from the SRAM array illustrating the impact of process variability on SRAM data remanence. The findings from this study not only expose SRAM's data remanence vulnerabilities from an attacker's viewpoint but also provide valuable insights for designing more resilient SRAM-based security in IoT devices.

The rest of the paper is structured as follows: Section II provides background on the architecture and operational dynamics of the 6T SRAM cell, explores the SRAM power-up state, and defines the threat model. Section III contrasts our approach with related work, underscoring the contribution of this study. Section IV details the experimental setup and methodology employed to investigate the temperature-induced vulnerabilities in SRAM. Section IV further discusses our results and provides root cause analysis. Section V concludes the paper with the implications for future research and applications.

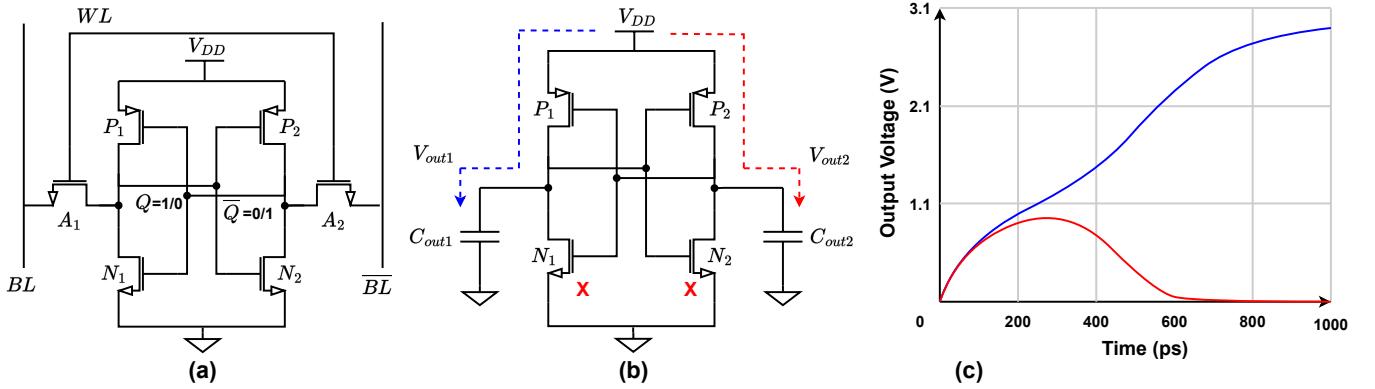


Fig. 2. Schematic of an SRAM cell showing power up current distribution: (a) a typical 6T SRAM cell, (b) current distribution of power-up transients, (c) transient power-up characteristics of the output node [9].

II. BACKGROUND

A. Static Random Access Memory (SRAM)

Fig. 2(a) illustrates a typical six-transistor (6T) SRAM bit cell. SRAM is a volatile semiconductor memory that uses bistable latching circuitry to store each bit. A 6T SRAM bit cell typically consists of two cross-coupled inverters, which create a stable positive feedback loop, and two access transistors to perform read/write operation on the cell. The cross-coupled inverters ensure that the bit cell remains in one of two stable states, representing either a binary '0' or '1'. The cross-coupled inverters are formed by four transistors: P_1 and N_1 form one inverter, and P_2 and N_2 form the other inverter. The access transistors are denoted as A_1 and A_2 , which connect the cell to the bit lines during read and write operations. In the logic '1' state of the cell, the P_1 (PMOS) transistor of the first inverter remains turned ON charging the node Q close to V_{DD} or logic high state. At the same time, the N_2 (NMOS) transistor of the second inverter remains turned ON ensuring that the complement node, labeled as \bar{Q} is at a low voltage (close to ground). The access transistors, A_1 and A_2 , remain off during the data hold state, isolating the cell from the bit lines and thus preserving the stored data.

B. SRAM Power-up State

Upon powering up, the state of an SRAM cell can be somewhat indeterminate due to random fluctuations and variations in the threshold voltages of the transistors [9], [10]. However, once the power stabilizes, the feedback loop formed by the cross-coupled inverters forces the cell into one of its two stable states (0 or 1), depending on the slight imbalances between the inverter pair induced by the process variations. The current flows and the node voltages during power-up transients are shown in Fig. 2(b) and Fig. 2(c), respectively. After powering up, initially, as V_{DD} is just turned ON, both the PMOS transistors are turned ON whereas both the NMOS transistors remain off. The difference in PMOS (P_1 and P_2) V_{th} decides which node Q or \bar{Q} charges faster. The node that charges faster determines which NMOS (N_1 or N_2) turns ON. Once the NMOS transistor turns ON the corresponding

node discharges to ground providing a positive feedback to stabilize the cell state. The power-up state of an SRAM cell is thus determined by the slight asymmetries in the initial current transients induced by the inherent process variations. These factors cause each SRAM cell to power up into a state that is consistent across power cycles, providing a unique and reproducible pattern. This property is utilized in various security applications, such as generating physical fingerprints for device authentication [11]–[13].

When the device is powered ON, the initial states of the SRAM cells are read, capturing the inherent randomness present in the initial states. This random pattern can be used as a source of entropy for various applications, including secure key generation and device authentication. The power-up state of an SRAM cell is therefore an essential feature that leverages the physical properties of the memory cells to enhance security and reliability in modern computing systems.

C. Threat Model

Our threat model details a specific attack on embedded systems that utilize SRAM, employing a cold boot technique to exploit data retention characteristics at low temperatures. In this attack, the attacker forcefully reboots the system while it is in a cold environment, a condition that significantly prolongs the retention of data within the SRAM, thus exposing its vulnerability to unauthorized data extraction. After inducing the reboot, the attacker employs an external boot device or relocates the SRAM module to a different system to facilitate access to the retained data. Utilizing memory analysis tools, sensitive information such as cryptographic keys can be extracted. Such an attack can lead to severe security implications, including the exposure of encryption keys and unauthorized access to confidential data.

The model assumes that the attacker has several key advantages: physical access to the target system, the ability to maintain or create a cold environment to maximize data retention in SRAM, and the requisite technical skills to perform sophisticated memory analysis and data extraction. Additionally, the attacker has access to appropriate external

boot devices or alternative systems capable of reading SRAM contents without initiating an overwrite. Timing is also crucial, as the extraction must occur swiftly enough to leverage the extended data retention period. Finally, the model assumes that the SRAM in the targeted system is unencrypted and lacks other protective measures that might otherwise hinder successful data extraction.

III. RELATED WORKS

Low-temperature attacks on Random Access Memory were first widely discussed by Halderman et al. [14], who explored the data remanence in DRAM. This study demonstrated that during a low-temperature start-up (cold boot), DRAM retains data for several seconds, thereby limiting the operating system's ability to protect cryptographic keys from attackers with physical access [14]. While cold boot vulnerabilities in DRAM have been extensively studied and subsequent research has shown data remanence on off-the-shelf SRAM chips [2], on-chip embedded SRAM has generally been considered more robust due to its inherent design, smaller memory size, and integration within SoCs or microcontrollers. Recent studies, however, have highlighted vulnerabilities in on-chip SRAMs. For example, Volt Boot [5] exploits asymmetrical power states to force SRAM state retention across power cycles, bypassing traditional cold boot attack enablers such as low temperature or intrinsic data retention time. A comprehensive study on SRAM data remanence indicates that traces of data could persist at -40°C [1], suggesting that cold temperatures could be used to retrieve cryptographic keys from SRAM, thus challenging the assumed security of these systems under extreme conditions. However, this research primarily focused on commercial off-the-shelf stand-alone SRAM devices with very old technology nodes.

In contrast, our work extends the data remanence analysis to a broader temperature range on both embedded and stand-alone SRAM chips. In addition, we provide a detailed bit-wise assessment of SRAM data remanence and provide a comprehensive modeling framework to assess data remanence vulnerability based on operating temperature and technology nodes. Moreover, our study applies this analysis to modern SRAM technologies, providing insights and recommendations relevant to the latest generation of SRAM devices [2], [7].

IV. EXPERIMENTAL SETUP AND RESULTS

This section details the experimental setup and methodology of our experiment aimed at assessing the data remanence of embedded and stand-alone SRAM chips under varying temperature conditions. Moreover, the results of the attack based on data remanence and the root cause analysis are also discussed.

A. Setup and Methodology

Our experimental setup involves the 16-bit MSP430F5529LP microcontroller (MCU) embedded in an MSP430F5529 LaunchPad development board, which provides dual voltage outputs of 5V and 3.3V through a

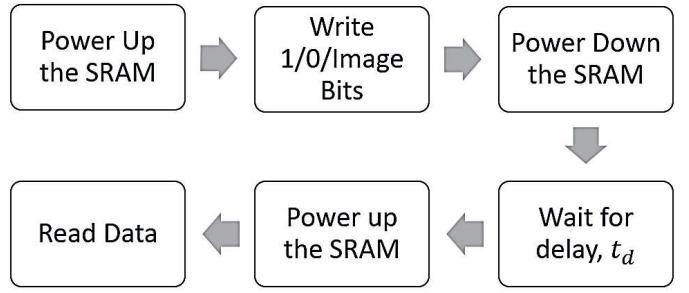


Fig. 3. Data collection procedure from the embedded SRAM during experiment.

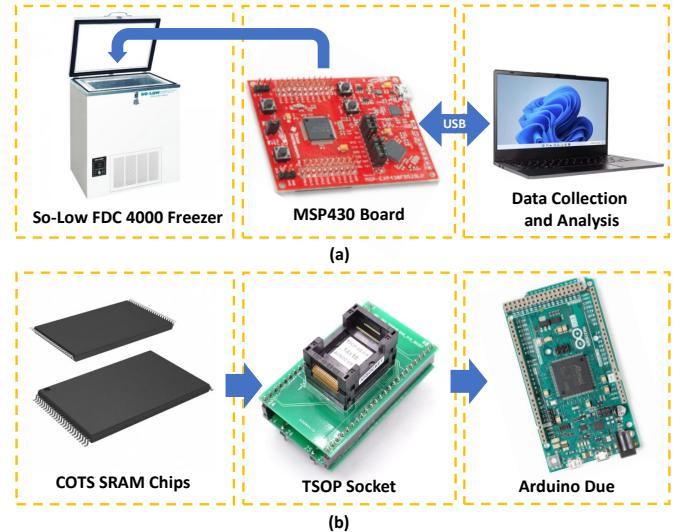


Fig. 4. Experimental Setup: (a) for Embedded SRAM, (b) for the COTS SRAM chips.

USB-connected high-efficiency DC/DC converter, ensuring stable power supply. The MCU features 128KB of Flash and 8KB of SRAM, divided into eight sectors; our experiments utilized the 2KB of general-purpose USB SRAM in sector 7. The board supports LPMx.5 low power mode, which cuts power supply by disconnecting the MCU core from the supply voltage and clearing SRAM contents. The low-power mode allows us to control the power on/off cycle of the SRAM array with precise delay during our retention experiments. We used Code Composer Studio as our Integrated Development Environment (IDE) and a SO-LOW FDC-4000 Freezer to control temperature conditions.

We used C program for the microcontroller to capture uninitialized SRAM contents immediately upon power-up and transmit these values to a host system via a simulated UART over USB, setting a baseline for analysis. The embedded SRAM is initialized programmatically by specific bit patterns to evaluate retention characteristics and simulate real-world power cycling effects. We replicated this procedure on two identical MSP430F5529LP units, cooled according to research protocols, ensuring controlled temperature maintenance during our trials. Our methodology involved initializing SRAM cells to '0', then systematically powering down with a delay from

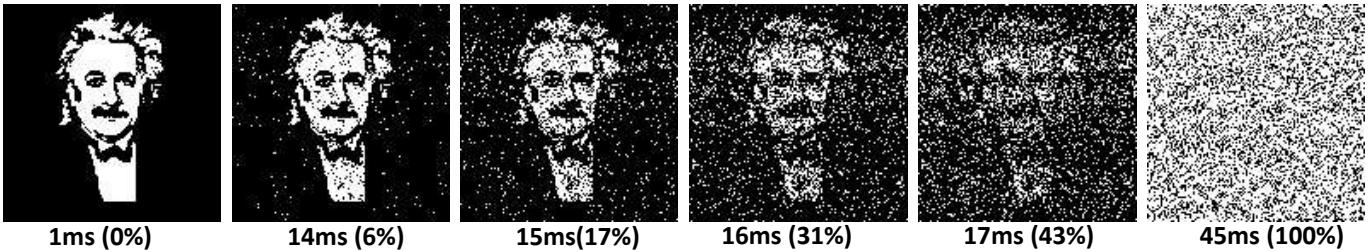


Fig. 5. Data Loss in binary Einstein Image at room temperature.

1ms to 60ms, during which most bits stabilized. We recorded the indices of bits stable at ‘0’ and ‘1’, then tested their resilience by inverting their states to confirm their ability to revert to original stable states. For image retention tests, a binary image was written to SRAM. The board was placed inside the freezer, and upon stabilizing the temperature, we commenced data collection as shown in Fig. 3. We collected data points up to the maximum retention time in our experiments, 800ms. The experimental setup for embedded SRAM has been depicted in Fig. 4(a).

Fig. 4(b) depicts the experimental setup used for the stand-alone SRAM chips. We utilized commercial-of-the-shelf (COTS) SRAM chips of 4-megabit capacity from Cypress Semiconductor. In this experimental setup, the SRAM chip is mounted on a socket, which connects to an Arduino DUE board through a custom designed PCB board. This board is linked to a workstation via a USB cable, enabling data transfer. The PCB incorporates a MOSFET switch, which facilitates rapid toggling of the V_{CC} power line to the SRAM chip. The Arduino executes a program that manages the SRAM operations and transmits experimental data to the laptop. During one trial for measuring image remanence, the Arduino program writes an image to the specified memory addresses with the SRAM turned on, turns the SRAM off for a specified duration, turns it back on, and then records the state of the memory.

B. Attack Demonstration

To demonstrate the attack, we write a binary representation of Einstein’s image into the 2KB SRAM sector, with the original image comprising 17.93% white (‘1’s) and 82.07% black (‘0’s) pixels. We conducted 60 power-up cycles, increasing the DRT by 1ms per cycle. Fig. 5 illustrates the progressive distortion of the image as DRT increases. Visible distortion begins at a power-down delay of approximately 14ms. Beyond this, the distortion continues to worsen, with complete data loss occurring after 45ms. For DRT values exceeding 45ms, the image degrades into a salt-and-pepper pattern, resembling the natural power-up state of the SRAM.

In order to quantify the data loss percentage with DRT, we use the following formula,

$$\text{Data Loss \%} = \frac{\# \text{ of set bits}(\text{Image}_{\text{original}} \oplus \text{PU}_{\text{read}})}{\# \text{ of set bits}(\text{Image}_{\text{original}} \oplus \text{PU}_{\text{ref}})} \times 100\%, \quad (1)$$

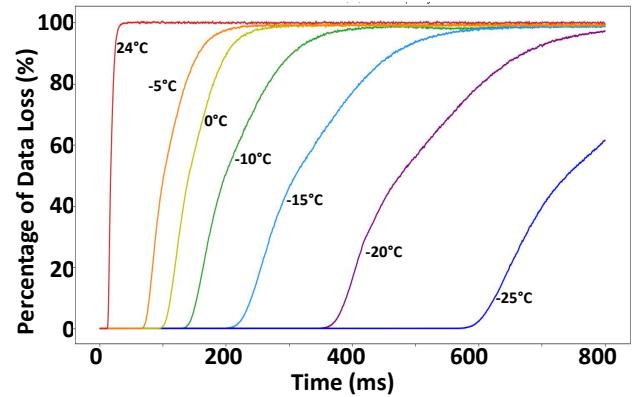


Fig. 6. Data Loss percentage for 800 power-up cycle at seven different temperatures.

where $\text{Image}_{\text{original}}$ is the original binary image written to the SRAM, PU_{read} is the power-up state read from the SRAM after a specified time, PU_{ref} is the reference power-up state. The reference power-up state is derived from 51 instances of power-up state for the same chip and by using majority voting for each bit position to assign a bit value for the noisy or unstable PUF bits. Note that the data loss percentage is zero if the PU_{read} retains the original image whereas it is 100% when PU_{read} resembles the natural or reference power-up state.

Our attack entailed writing the known Einstein’s image in binary form, across seven different temperatures to the SRAM sector and calculating the data loss percentage over DRT. Fig. 6 summarizes our characterization results on the temperature-dependency of the DRT, highlighting the vulnerability of SRAM to data remanence attack for lower temperature. As shown in Fig. 6, the known image data can be retrieved up to 600ms without any significant distortion which illustrates that the embedded SRAM is vulnerable to data remanence attacks in cold temperatures.

The results indicate that lower temperatures effectively slow down the data loss process, extending the retention times considerably. The findings from our experiments underscore the significant impact of temperature on SRAM data remanence. Low temperatures enhance the data retention capacity of SRAM cells, potentially extending the time window during which residual data can be recovered post-power-off. This extended retention poses a critical security risk, particularly

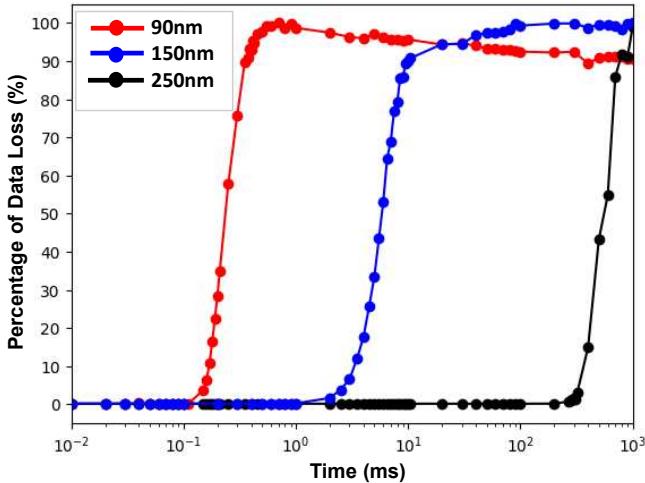


Fig. 7. Data Loss percentage in SRAM chips fabricated in different process technology shows an increase in data remanence effect in larger (older) technology nodes.

for IoT devices relying on embedded SRAM for sensitive data storage, such as cryptographic keys.

Fig. 7 presents a comparative analysis of data loss at room temperature across three stand-alone COTS SRAM chips from Cypress Semiconductor, fabricated using 250nm, 150nm, and 90nm processes respectively. Each of the 3 chip's data loss plots shown in Fig. 7 are normalized separately such that each chip's PU read that has the largest Hamming distance from the original image is taken to have a data loss of 100%. The findings indicate that the SRAM chip fabricated with the 250nm process exhibits the longest data remanence time. In contrast, the 150nm chip shows reduced remanence time, and the 90nm chip records the shortest duration of data retention. This pattern underscores the significant influence of transistor technology leakage on SRAM data remanence; larger process nodes, characterized by lower leakage, tend to preserve data longer, thereby enhancing data remanence. Conversely, smaller process technologies, such as the 90nm, with higher leakage rates, demonstrate a diminished data remanence effect. This inverse relationship between process size and data remanence is consistent with our findings on the impact of reduced temperatures on data retention. Here, the leakage current is directly proportional to temperature; thus, lower temperatures increase data remanence.

C. Root Cause Analysis

The root cause of the observed data remanence vulnerability in SRAM stems from the inherent leakage characteristics of the memory cells as illustrated in Fig. 8. Assume that an SRAM cell is holding logic '1' state and thereby the node Q is held at high voltage. We illustrate the discharge process of node Q upon power down using a effective capacitance that is connected between node Q and ground. Fig. 8 illustrates the discharge path of node Q with dashed arrows in red. Note that both the NMOS transistors are at off state and provides a sub-threshold leakage current path for the node Q even when power supply is on. The PMOS (P1) replenishes any lost charge

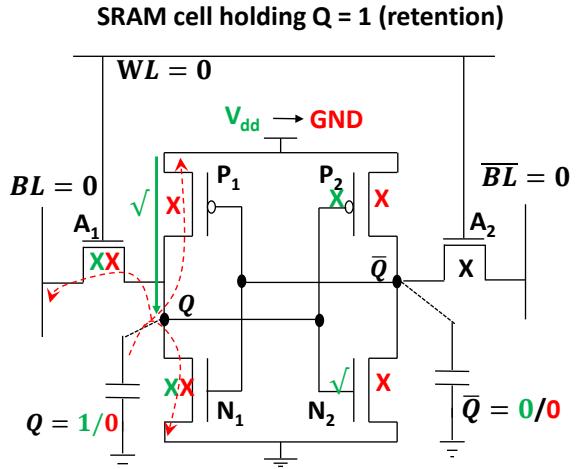


Fig. 8. SRAM Cell showing leakage path. Green path stands for power ON condition, red path stands for power OFF condition

from node Q by providing current to the node (green solid arrow) and maintains high voltage on node Q. However, once power supply is turned OFF, PMOS no longer supply any current and Q starts to discharge through all three transistors (A_1, N_1, P_1) using sub-threshold leakage current (I_{leak}). The discharge process can be described by the following equation:

$$I_{leak} = \frac{\Delta Q}{\Delta t} = C \frac{\Delta V}{\Delta t}, \quad (2)$$

where C represents the effective capacitance at node Q, and $\Delta V/\Delta t$ indicates the rate of voltage change across the capacitor.

The sub-threshold leakage current follows the following relationship,

$$I_{leak} \propto e^{\left(\frac{-V_{th}}{\eta V_T}\right)} \quad (3)$$

where we see that it is exponentially dependent on the threshold voltage (V_{th}) relative to the thermal voltage (V_T), which is a function of the temperature (T) and Boltzmann's constant (k_B). The parameter η is the sub-threshold slope factor which depends on transistors geometrical and material properties.

The period during which data remains recoverable, or DRT depends on the leakage current as follows:

$$DRT \approx \frac{C \Delta V}{I_{leak}} \propto C \times V_{dd} \times e^{\left(\frac{qV_{th}}{\eta k_B T}\right)}, \quad (4)$$

which demonstrates that the remanence time is contingent upon the supply voltage (V_{dd}), the node capacitance of the cell, transistor's geometry (η) and the thermal stability influenced by the temperature.

The above analysis entails the root cause of the vulnerability which is the prolonged retention times at lower temperatures, as depicted in Fig. 6. The long retention occurs due to the temperature variations that influence the electrical properties of memory cells, affecting their ability to retain data. In SRAM, elevated temperatures increase leakage currents, leading to a quicker data loss, lowering the value of DRT. Conversely,

lower temperatures reduce leakage currents, potentially extending the DRT value. The relationship is solidified by the conceptual mathematical model stated in Eq. 4.

D. Cell-by-cell DRT Analysis

In stand-alone or embedded SRAMs, process variations during semiconductor manufacturing can lead to differences in the electrical characteristics of the SRAM cells. To quantify these effects, we examine the bit-by-bit variation in the Data Retention Time (DRT) of the SRAM sector. Fig. 9 shows the DRT for all stable natural bit-cells (those that consistently stabilize to either '0' or '1') among the 2KB memory cells. The distribution closely follows a Gaussian curve, depicted by the solid line. The mean DRT across the SRAM chip is approximately 24ms at room temperature, but there is a wide range of DRT values, spanning from 12ms to over 30ms.

This cell-by-cell DRT variation explains the gradual image distortion observed in Fig. 5, where image degradation began around 14ms. Essentially, the bits in the lower tail of the DRT distribution define the initial onset of data loss in our experiment, while the upper tail marks the point at which complete data loss occurs.

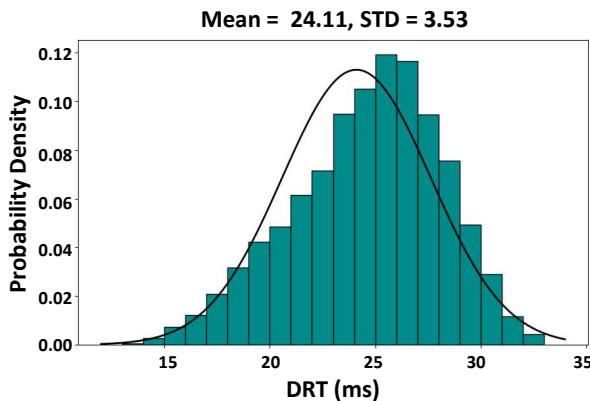


Fig. 9. Data Remanence Time Distribution plot for stable natural bits, at room temperature.

To assess the repeatability of DRT measurements for stable memory cells, we analyzed their properties by conducting the experiment twice. The stable cells are initialized to states opposite to their power-up states, i.e., stable logic '1' cells are initialized to logic '0', and stable logic '0' cells are initialized to logic '1', which creates a petal-shaped correlation plot shown in Fig. 10. We observe a high degree of repeatability in cell behavior, with a 97% to 98% correlation in stability across two experimental sets spaced six months apart. This repeatability confirms the consistent nature of SRAM data remanence, exposing its vulnerabilities to future attacks.

V. CONCLUSION

In this study, we explored the temperature-induced vulnerabilities of data remanence in embedded SRAM, particularly under cold conditions that significantly prolong data retention times. Our findings highlight the risk of cold boot attacks where sensitive data such as cryptographic keys can

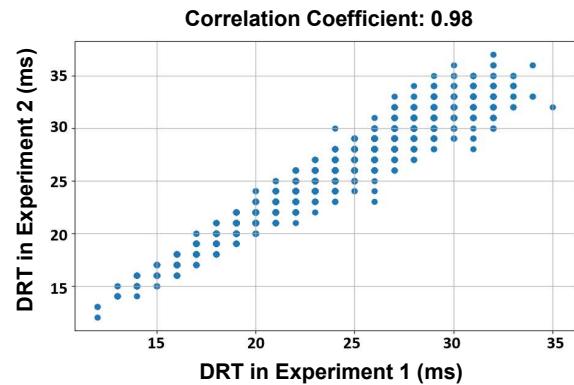


Fig. 10. Correlation between two sets of data from all stable 1s in room temperature.

be compromised due to extended DRT at low temperatures. The research underscores the necessity for more secure SRAM designs, resistant to environmental influences, to safeguard sensitive information in IoT devices. The development of a predictive model based on leakage current variations provides a foundation for future advancements in SRAM security, emphasizing the need for innovations that address these temperature-dependent vulnerabilities.

REFERENCES

- [1] S. P. Skorobogatov, "Low temperature data remanence in static ram," in *Technical Report UCAM-CL-TR-536, University of Cambridge, Computer Laboratory, Cambridge*, 2002. [Online]. Available: <https://api.semanticscholar.org/CorpusID:60957475>
- [2] N. A. Anagnostopoulos, T. Arul, M. Rosensthl, A. Schaller, S. Gabmeyer, and S. Katzenbeisser, "Low-temperature data remanence attacks against intrinsic sram pufs," in *2018 21st Euromicro Conference on Digital System Design (DSD)*, Prague, Czech Republic, 2018, pp. 581–585. [Online]. Available: <https://doi.org/10.1109/DSD.2018.00102>
- [3] B. M. S. Bahar Talukder, F. Ferdaus, and M. T. Rahman, "Memory-based pufs are vulnerable as well: A non-invasive attack against sram pufs," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4035–4049, 2021.
- [4] J. Hovanes, Y. Zhong, and U. Guin, "Beware of discarding used srams: Information is stored permanently," in *2022 IEEE Physical Assurance and Inspection of Electronics (PAINE)*, 2022, pp. 1–7.
- [5] J. Mahmud and M. Hicks, "Sram has no chill: exploiting power domain separation to steal on-chip secrets," in *Proceedings of the 27th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '22)*. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1043–1055. [Online]. Available: <https://doi.org/10.1145/3503222.3507710>
- [6] U. Surendranathan, H. Wilson, L. R. Cao, A. Milenkovic, and B. Ray, "Analysis of sram puf integrity under ionizing radiation: Effects of stored data and technology node," *IEEE Transactions on Nuclear Science*, vol. 71, no. 4, pp. 485–491, 2024.
- [7] S. Skorobogatov, "Hardware security implications of reliability, remanence, and recovery in embedded memory," *Journal of Hardware and Systems Security*, vol. 2, December 2018.
- [8] Texas Instruments, *MSP430x5xx and MSP430x6xx Family User's Guide (Rev. N)*, n.d., online available: <https://www.ti.com/lit/pdf/slau208>. Accessed on: May 19, 2024. [Online]. Available: <https://www.ti.com/lit/pdf/slau208>
- [9] U. Surendranathan, H. Wilson, M. Wasiolek, K. Hattar, A. Milenkovic, and B. Ray, "Total ionizing dose effects on the power-up state of static random-access memory," *IEEE Transactions on Nuclear Science*, vol. 70, no. 4, pp. 641–647, April 2023.

- [10] W. Wang, A. Singh, U. Guin, and A. Chatterjee, “Exploiting power supply ramp rate for calibrating cell strength in sram pufs,” in *2018 IEEE 19th Latin-American Test Symposium (LATS)*, March 2018, pp. 1–6.
- [11] J. Guajardo, S. S. Kumar, G.-J. Schrijen, and P. Tuyls, “Fpga intrinsic pufs and their use for ip protection,” in *Cryptographic Hardware and Embedded Systems - CHES 2007*, P. Paillier and I. Verbauwhede, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 63–80.
- [12] U. Guin, W. Wang, C. Harper, and A. D. Singh, “Detecting recycled socs by exploiting aging induced biases in memory cells,” in *2019 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, May 2019, pp. 72–80.
- [13] K. Xiao, M. T. Rahman, D. Forte, Y. Huang, M. Su, and M. Tehranipoor, “Bit selection algorithm suitable for high-volume production of sram-puf,” in *2014 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, 2014, pp. 101–106.
- [14] J. A. Halderman, S. D. Schoen, N. Heninger, W. Clarkson, W. Paul, J. A. Calandriño, A. J. Feldman, J. Appelbaum, and E. W. Felten, “Lest we remember: Cold-boot attacks on encryption keys,” in *Proceedings of the 17th USENIX Security Symposium*. San Jose, CA: USENIX Association, 2008.