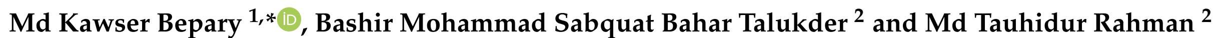
applied sciences

MDPI

Article

DRAM Retention Behavior with Accelerated Aging in Commercial Chips



Md Kawser Bepary [1,*], Bashir Mohammad Sabquat Bahar Talukder [2] and Md Tauhidur Rahman [2]

[1] Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32603, USA
[2] Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174, USA; bbaha007@fiu.edu (B.M.S.B.T.); mdtrahma@fiu.edu (M.T.R.)
* Correspondence: mdkawser.bepary@ufl.edu

Abstract: The cells in dynamic random access memory (DRAM) degrade over time as a result of aging, leading to poor performance and potential security vulnerabilities. With a globalized horizontal supply chain, aged counterfeit DRAMs could end up on the market, posing a significant threat if employed in critical infrastructure. In this work, we look at the retention behavior of commercial DRAM chips from real-time silicon measurements and investigate how the reliability of DRAM cells degrade with accelerated aging. We analyze the retention-based errors at three different aging points to observe the design-induced variations, analyze the pattern dependency, and explore the impacts of accelerated aging for multiple DRAM vendors. We also investigate the DRAM chips' statistical distribution to attribute the vital wear-out effects present in DRAM. We see a continuous increase in retention error as DRAM chips age and therefore infer that the aged retention signatures can be used to differentiate recycled DRAM chips in the supply chain. We also discuss the roles of device signature in DRAM aging and aging-related security implication on DRAM row-hammer error.

Citation: Bepary, M.K.; Talukder, B.M.S.B.; Rahman, M.T. DRAM Retention Behavior with Accelerated Aging in Commercial Chips. *Appl. Sci.* **2022**, *12*, 4332. https://doi.org/10.3390/app12094332

Academic Editor: Justo García Sanz-Calcedo

Received: 10 March 2022
Accepted: 20 April 2022
Published: 25 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Keywords: dynamic random access memory (DRAM); retention error; accelerated aging; bias temperature instability (BTI); time-dependent dielectric breakdown (TDDB)

1. Introduction

With the growth of the worldwide semiconductor market value, the design, manufacturing, and distribution of integrated circuits (ICs) have evolved into a complex globalized supply chain. As electronic parts come from a diverse set of suppliers, the problem of recycled and counterfeit electronic components in the semiconductor supply chain has become an increasing security threat over the past few years [1,2]. The performance of an IC continuously degrades during its lifetime due to different wear-out and aging mechanisms, such as time-dependent dielectric breakdown (TDDB), bias temperature instability (BTI), hot carrier injection (HCI), ionizing radiation, etc. The inclusion of the degraded counterfeit components into the global semiconductor supply chain poses severe security threats to personal and national privacy, critical infrastructures due to their reduced lifespan, inferior quality, and lower performance [1,3]. Using counterfeit and recycled ICs in critical infrastructure such as defense, aerospace, transportation, medical, smart-grid, etc., can result in potential system malfunction and health or safety concerns [4].

The system's reliability and security can be significantly compromised as counterfeit chips may fail any time after being deployed. Moreover, adversaries may exploit counterfeit ICs to bypass security mechanisms and gain remote access or leak sensitive data. Various invasive or non-invasive attacks can endanger the confidentiality, integrity, and availability of a system [5,6]. Recent studies report that modern DRAM memories are vulnerable to row-hammer attacks that can induce disturbance error in a row without accessing it. A row's restricted data are altered by the continuous activation of its neighboring rows [5]

due to the cells' coupling effects, noise, and interference. Data analysis of early failure rate (EFR) leverages the impacts of aging phenomena to distinguish recycled counterfeit chips [1]. Huang et al. [7] discussed a low-cost statistical method based on typical test results from production early failure rate (EFR) analysis, such as V_{min}, F_{max}, and I_{ddq} to train a one-class classifier using brand-new devices only.

It is crucial to use trusted memory chips in order to ensure security in critical applications. Although there are a few works on detecting different recycled memory chips such as SRAM [8] or flash memory [9], no such experiment on DRAM aging is available. Tehranipoor et al. [10] discuss the aging effects that degrade DRAM chips and explore its impact reliability of DRAM-based PUF but do not address its use in recycled chip detection. In [11], a technique to identify DRAM origin is presented that discusses a non-invasive scheme to verify individual DRAM specification by exploiting the DRAM latency variations. In [12,13], effects of Bias Temperature Instability (BTI) and Hot Carrier Injection (HCI) are analyzed with SPICE simulation to show the increase in retention time as the device ages. This research is motivated by the desire to investigate DRAM reliability, name retention behavior with device aging, and develop a practical and effective technique to detect recycled DRAM memory.

The main contributions of this article are as follows.

- We experimentally demonstrate the retention-based error by violating the minimum retention time during the refresh operation on DRAM rows. We assess the reduced refresh rate's effects as increased soft errors on multiple DRAM memory modules for different patterns and timing parameters.
- We utilize the design-induced variation of true cells and anti-cells on different modules to develop a scheme to ensure the maximum stress and wear-out effects during the accelerated aging test.
- We highlight the significant observations in retention behavior with the accelerated aging process.
- We examine the statistical distributions of retention error with aging and analyze its application to detect recycled DRAM chips.

The rest of the paper is organized as follows. In Section 2, we briefly present the DRAM organization and operations as well as common wear-out mechanisms. In Section 3, we discuss the testing methodology for the DRAM retention test and accelerated aging. We describe the experimental setup and present the experimental results in Section 4. Finally, in Section 5, we discuss the roles of latency-based and retention-based signatures in DRAM aging and the impacts of aging on row-hammer error. We conclude our work with future research directions in Section 6.

2. Background and Motivations

In this section, we describe the backgrounds on DRAM organization and operation principles, DRAM retention behaviors, and wear-out mechanisms that causes degradation during the lifetime of a device.

2.1. DRAM Organization and Operation

A modern DRAM system is organized into a hierarchy of channel, rank, bank, DRAM chips, DRAM cells, and memory controller, as illustrated in Figure 1. Memory commands, address space, and data are driven between the memory controller and DRAM modules by the memory channel. Depending on the memory requirement, there might be one or multiple DRAM modules. DRAM modules are divided into one or multiple ranks. A rank consists of several DRAM chips (generally 4/8/16) and provides a wide data-bus (usually 64 bits) together.

To support the parallelism, each DRAM chip consists of multiple banks. In a memory bank, the DRAM cells are placed in a two-dimensional array. The memory address is divided into a row address and column address, which are decoded by a row and a column decoder during a read or write operation to find the corresponding cell. The rows of

the DRAM are known as wordlines (or pages). The columns are known as bitlines, and the chip density determines the total number of rows. The bitlines are connected to the row-buffer, which is a row of sense-amplifier. A DRAM cell consists of two components: an access transistor and capacitor to hold a charge. The charging state of the capacitor determines the stored value. The access transistor is used to access the stored value. A fully charged capacitor represents logic '1', and an empty capacitor represents logic '0'. The access transistor connects the capacitor with a bitline and is controlled by the wordline. The DRAM content (i.e., the state of charge in the capacitor) is read or overwritten by activating a wordline. A V_{dd} to the wordline makes a path between the capacitor and bitline. The bitlines also connect the DRAM cell to the sense amplifier (i.e., row-buffer).

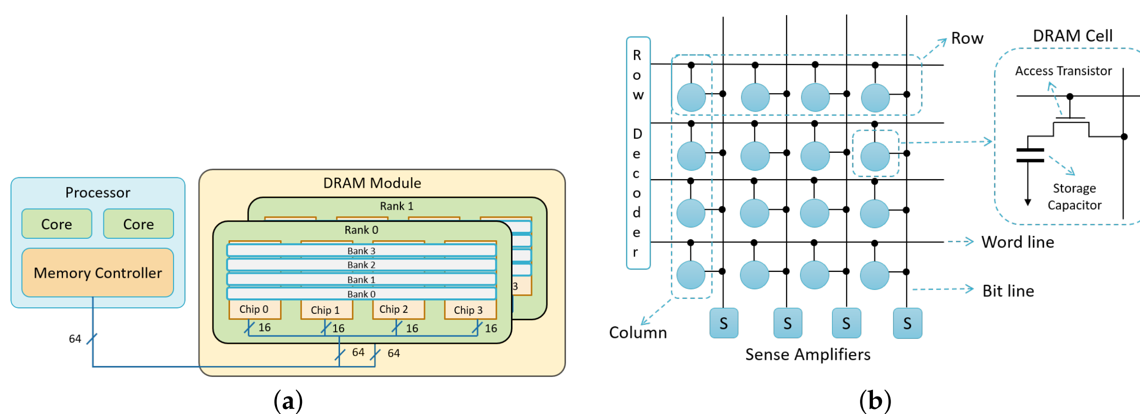


Figure 1. Organization of a modern memory subsystem and array architecture. (**a**) DRAM system hierarchy [14]. (**b**) Array architecture.

Initially, in the precharge state, all bitlines are precharged to $V_{dd}/2$. In the next state, the wordline is activated by raising the value of the wordline to V_{dd}. Once the pass-transistor of the DRAM cell is ON, the charge flows between the capacitor and the respective bitline. The charge moves from the bitline to the capacitor if the stored value is '0' and vice-versa. In the last stage, the sense-amplifier attached to the bitline amplifies the bitline voltage to strong logic '1' (or '0'). Cell data are then read or written by sensing or driving the voltage of the corresponding bitlines. DRAM cells lose data over time as the cell capacitors discharge gradually and need to be replenished periodically. This operation is known as refresh operation, and the time required is the retention time (t_{ret}). The voltage change due to capacitor discharge is known as voltage perturbation. The data stored on the DRAM cell are lost when the sense amplifiers cannot reliably differentiate the perturbation from noise. Traditionally, DRAM cells are required to have a retention time equal to or greater than 64 ms to preserve data integrity. Recall that sense amplifiers drive each bitline fully to V_{dd} or 0 V when a row is activated. Consequently, the activated row's cell capacitor is fully charged to V_{dd} or 0 V.

2.2. Retention-Based Error

A retention error occurs when a DRAM cell loses its data due to charge leakage in the cell capacitor. The leakage current differs between cells depending on the manufacturing variation of the capacitor and the access transistor [15]. Studies show that the retention time in DRAM varies following a bi-modal distribution [16,17]. Over 99% of the cells in a DRAM chip can hold the data for a much higher retention time (main distribution). The remaining 1% of cells have high leakage current and exhibit a very low retention property (tail distribution). In order to ensure the reliability of the DRAM cells across the whole population, the standard retention time is determined by the leaky cells of the tail distribution as it consists of the weakest cells of the device. Prior studies have introduced mechanisms to profile the cells' retention time and refresh DRAM cells intelligently to alleviate substantial energy and performance overhead caused by the refresh operations [15,18,19]. These methods ac-

cumulate cell retention times into different bins based on varying retention times and apply different refresh rates to different bins in order to avoid unnecessary refresh operations.

2.3. Disturbance Error

Disturbance errors are caused by a strong interaction between two circuit components, such as capacitors and transistors. In [5], one specific disturbance error is identified as a row-hammer error when the voltage of the wordline is toggled continuously to cause the nearby rows' weak cell to leak charge at a higher rate. These disturbance errors are attributed to a cumulative interference of different device-level interactions such as electromagnetic coupling, hot-carrier injection, bridge faults, etc. [5].

2.4. True Cells and Anti-Cells

In DRAM cell architecture, a true cell refers to a cell that is a logical 1 with the charged state and an anti-cell represents a logical 1 with the discharged state [5]. Typically, DRAM modules comprise a combination of true cells and anti-cells. Consequently, the retention error probability may be applied to true cells and anti-cell regions independently. For the true cells, retention error will occur only when the stored value is '1' and the capacitor is charged. Thus, the capacitor can be leaked in this state and cause a retention error, but no error will occur when the stored value is '0'. In contrast, the logic values '0' and '1' are treated as the opposite in the inverted anti-cells. The anti-cells will exhibit a retention error when the stored logic value is '0' and the capacitor is in a charged state. Similarly, no error will occur when a logic value of '1' is stored in the anti-cells.

2.5. Wear-Out Mechanisms in DRAM

As process technology scales down to smaller nodes aggressively, integrated circuits suffer from increasing performance variation and concerns about the reliability of operation, leading to a fabricated chip's failure. Failures can occur at various stages in the lifetime of an integrated circuit. ICs can fail during production due to disproportionately high process variations, weak design, or local defects. Electronic devices also suffer from a high failure rate known as infant mortality during the early life in operation [1]. Integrated circuits can also fail during their lifetime because of wear and tear, aging, harsh environments, overuse, etc. Figure 2 presents the typical bathtub curve, which is often used to illustrate electronic device failure characteristics [1].

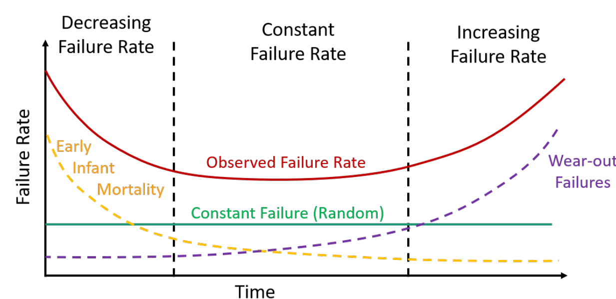


Figure 2. A classical bathtub curve represents common device failure characteristics [1].

Recycled or aged electronic devices often exhibit weaker performance and a shorter time to failure. The fundamental wear-out mechanisms are essential to understand the factors influencing transistor and capacitor degradation with aging. Commonly found wear-out failure mechanisms, such as bias temperature instability (BTI), hot carrier injection (HCI), time-dependent dielectric breakdown (TDDB), etc. are discussed in this section.

2.5.1. Bias Temperature Instability (BTI)

Bias Temperature Instability (BTI) is the phenomenon in which the transistor threshold voltage shifts because of the trapped charges at the Si–SiO_2 interface. Dangling bonds at

the substrate–oxide interface act as trapped charges, increasing the donor concentration and threshold voltage [20]. NMOS cell transistors are affected by Positive Bias Temperature Instability (PBTI) when their gate bias voltages are positive. Generally, they experience less BTI impact than PMOS transistors [10]. The long-term impact of BTI depends on the transistor's duty factor as the cell transistors are stressed out only during the active mode of DRAM operations.

2.5.2. Hot Carrier Injection (HCI)

HCI is the adverse effect where a charge carrier gains enough kinetic energy and a potential barrier to break from an interface state. The hot carriers create trap states in the gate oxide by breaking the silicon–hydrogen bonds. Most of the degradation occurs near the end of the channel as the charge carriers have high kinetic energy in this region [21]. The switching characteristics may be permanently altered because of the trapped charge carriers at the gate dielectric. NMOS circuits suffer more from the HCI effect than PMOS circuits as the electrons in the n-type device have a smaller energy barrier than holes in p-type devices.

2.5.3. Time-Dependent Dielectric Breakdown (TDDB)

In the TDDB failure mechanism, the gate oxide breaks down due to a relatively low electric field's long-time application. When MOSFETs are operated at a voltage close to or beyond their specified operating voltages, trapped charges may form a conducting path through the dielectric gate oxide to the substrate. Defects occupied by electrons or holes will generate a tunneling current through this conductive path once the trap density reaches a threshold point. Thus, it will result in a large current flow through the gate, heating the oxide layer and finally breaking down the dielectric completely [22].

2.5.4. Radiation Failure

Ionizing radiation from natural or artificial sources may create hundreds of electron–hole pairs in the gate oxide, which can degrade the transistor and even disrupt performance [23]. The electrons get swept away out of the gate oxide due to high mobility, leaving trapped positive charge at the Si–SiO_2 interface. Consequently, it results in a shift of threshold voltage and increases the channel leakage current [20,24]. Over an extended period of time, trapped charges at the interface and oxide may cause the TDDB effect and a complete breakdown of the dielectric oxide.

An integrated circuit's performance slowly but gradually degrades over time due to such wear-out phenomena. The degradation rate might depend on the switching activity, the gate voltage, the content of the DRAM cell, and several other factors [25,26].

3. Testing Methodology

This section presents a brief description of the testing methodology and the steps involved for the experimental study. The testing methodology can be divided into two parts: DRAM retention test and accelerated aging. In a retention test, the cell capacitors in a DRAM module are allowed to leak charge for an extended retention period to induce error in the written data. In accelerated aging, the module is taken through a burn-in test at a high temperature and voltage for a long time. At the stress condition, the wear-out mechanisms affect the cells.

Initially, a DRAM retention test is carried out to analyze the induced errors in the cells at a lower refresh rate for fresh modules. Then, the DRAM module is taken through an accelerated aging process for 8 h. The retention test and accelerated aging process is repeated several times to observe the effects on retention error. In our experiment, the retention error measurement is performed for each bank at three different time points: $t_0 = 0$ h, $t_1 = 8$ h, $t_2 = 16$ h. The tests for the same devices are performed under the same condition to mimic the effects of aging degradation with time. While devices at $t = t_0$ are referred to as fresh or brand new, devices at $t > t_0$ are presented as aged or recycled

or counterfeit devices. Now, the induced retention errors can be compared for the fresh chips and aged chips at two different aging points. Figure 3 illustrates an overview of the testing methodology and steps required to find the retention-induced error and the effects of accelerated aging.

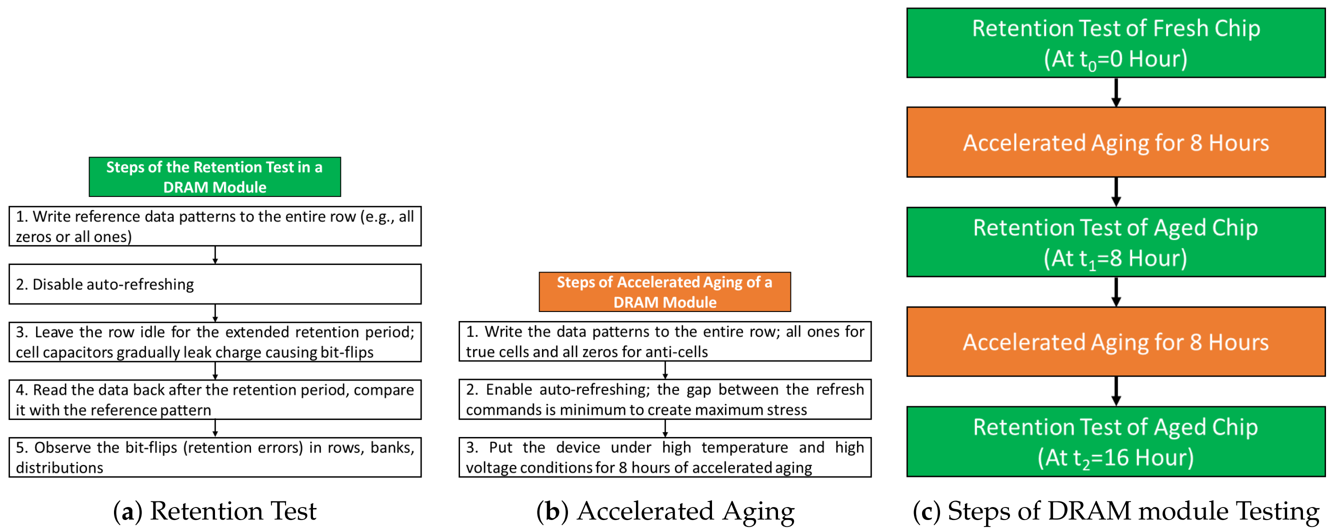


Figure 3. Overview of the testing methodology of a DRAM module.

3.1. DRAM Retention Test

A DRAM retention test is performed by increasing the time to refresh the DRAM cells and observing the errors induced due to cell capacitor leakage. In our experiment, retention-based errors are induced by refreshing the rows with a much higher retention time than the standard retention period of 64 ms. This allows the cell capacitors to leak for an extended period of time, causing bit flips to a significant number of cells. The retention test involves the following steps:

1. A reference data pattern is written to a complete row. We use all zeros (0x00) and all ones (0xFF) patterns in our experiments.
2. Auto-refreshing is disabled because we want the cells to leak charge for an extended retention period without being refreshed.
3. The rows are left idle for the specified retention time (10 s, 20 s, etc.). The cell capacitors gradually leak charge, causing bit flips in the cell data.
4. The data are read back from the rows after the retention period and sent back to the host machine. The read data are compared with the written data to detect the cells that could not hold the data for that duration as the bits are flipped.
5. The raw data are analyzed to observe the bit flips/retention errors of a DRAM module in the rows, banks, and their distributions.

These steps are performed for all the rows in a DRAM bank and for all the banks in a module at time. To ensure that only the refresh or retention time affects the cell leakage in this test, we perform the read and write operations with the standard timing parameters. In our experiment, we perform the retention test for three different retention times: at t_{ret} = 10 s, 20 s, 50 s.

3.2. True and Anti-Cell Detection

After collecting and analyzing the new DRAM modules' retention errors for different patterns and retention times, the errors are attributed to process and architectural variation. From the retention test results, the phenomenon of design-induced variation is observed in two forms: (i) true cells and anti-cells due to architectural variation of the cells and (ii) variation across local parts of the bank due to relative distance from the peripherals [27]. True cells induce error only for logic '1' as the cell capacitor discharges from HIGH to LOW

state, and they do not induce any error for logic ‘0’ as the capacitor remains at the LOW state. On the other hand, anti-cells discharge for logic value ’0’ and induce error and show no error for logic value ‘1’.

In order to identify the true cells and anti-cells, the follwing steps are required:

1. The retention test is performed separately with all-zeros (0x00) and all-ones (0xFF) patterns in a module.
2. The retention errors in bank are aggregated for each row.
3. The rows that induce error only for the 0xFF pattern and show no error for the 0x00 pattern are identified as true cells. The rows inducing error only for the 0x00 pattern are identified as anti-cells. The rows that show errors for both patterns have a mixture of true and anti-cell configuration.

Identifying true cells and anti-cells is important for the aging process. In the aging process, our goal is ensure maximum stress to the storage capacitors. In order to achieve that, data pattern 0xFF is written to the true cells and pattern 0x00 is written to the anti-cells to ensure uniform aging across the DRAM array. In the modules with a mixture of both true cells and anti-cells, the 0xFF pattern is written to all the rows during the aging process.

3.3. Accelerated Aging of DRAM Modules

In order to observe the wear-out effects on the DRAM chips, a test similar to the burn-in test is performed on the DRAM modules. The module is exposed at an elevated temperate and DC voltage for a prolonged period of time to simulate a stress condition. This stress condition accelerates the aging process in the cell transistors. When DC stress is applied to the DRAM node, the leakage current increases due to trapped emission electrons, and the DRAM cell capacitors suffer from asymmetric damage, degrading the retention behavior [28]. Evaluating the retention error at this accelerated aging condition can be used to identify the impacts of wear-out mechanisms (e.g., BTI, HCI, electromigration) and detect the aged device as well as the infant mortality failures and unexpected failures [2].

The following steps are involved in the accelerated aging process:

1. Initially, a data pattern is written to the entire row using standard timing parameters. In order to ensure maximum stress, all ones (0xFF) are written to the true cells, and all zeros (0x00) are written to the anti-cells.
2. Auto-refresh is enabled. The timing parameter t_{REFI} (the gap between refresh command) is set to 0, a minimum value (one clock cycle), so that DRAM cells are refreshed more frequently and create maximum stress on the storage capacitor.
3. The DRAM module is placed under the thermostream device at a high temperature of 70 °C and a high-voltage condition ($V_{dd} = 1.55$ V instead of standard 1.5 V. At this stressed condition, wear-out effects are accelerated in the DRAM cells. The module is left in this condition for 8 h for wear-out mechanisms to affect the DRAM cells.

After the aging process in stress conditions, the module is cooled down to room temperature, and the retention test is performed to evaluate the bit flips. The retention errors are compared with the fresh chips and the additional errors are attributed to various aging mechanisms.

4. Experimental Results and Analysis

We describe the experimental framework and retention test results for the fresh and aged DRAM modules in this section. We observe the retention test results for new modules and determine how the number of retention errors in DRAM modules varies depending on the data pattern, timing parameter, and architectural variation. Then, we observe the retention results of the aged modules to characterize the effects of wear-out mechanisms.

4.1. Experimental Framework

A DRAM testing framework based on Xilinx ML605 FPGA, and a DDR3 memory controller called SoftMC [29] is used to test the DDR3 modules in our experiments. Retention

error data are collected from nine commercially available off-the-shelf single rank DDR3 SODIMMs modules (small outline dual in-line memory module) covering all three major DRAM vendors. We used a Micron 1 GB DDR3 module, Hynix 2 GB module, and Samsung 2 GB modules in our tests. Each memory module contains four or eight ICs with eight banks spreading across the whole module, and each bank consists of 16 k or 32 k rows (1 GB or 2 GB).

SoftMC is an FPGA-based open-source experimental memory testing infrastructure that implements all low-level DRAM commands in a programmable memory controller. This programmable memory controller is introduced to the user with an easy-to-use and flexible interface. This interface can be programmed in C++ with customized library files to generate bit files. The user can generate a sequence of DDR3 operations by calling the SoftMC API functions [29]. DRAM cell data can be written or read from each memory bank using the testing platform implemented on a Xilinx Virtex ML605 evaluation board with SoftMC.

The experimental framework provides the ability to:

1. Generate different timing parameters and test patterns for various latency-based tests of the DRAM module.
2. Facilitate with an interface to connect the host machine with the Xilinx FPGA board.
3. Enable precise and accurate control of the temperature of the DRAM module during the test.

In order to sustain a stable operating temperature for the DIMMs during the DRAM aging process, the modules are placed on the FPGA board under an ATS Series Thermostream device [30]. The Thermostream device allows for fast and precise thermal conditioning with a temperature range of −20 to +225 °C.

4.2. Retention Error in Fresh DRAM Modules

For standard DRAM operations, each row is refreshed periodically after $t_{ret} = 64$ ms. If the DRAM cell capacitors are refreshed after a longer interval, the cells lose charges and cause a bit flip. The number of retention errors for a fresh DRAM module depends on the retention time, data pattern, and architectural variations.

4.2.1. Dependence on Retention Time

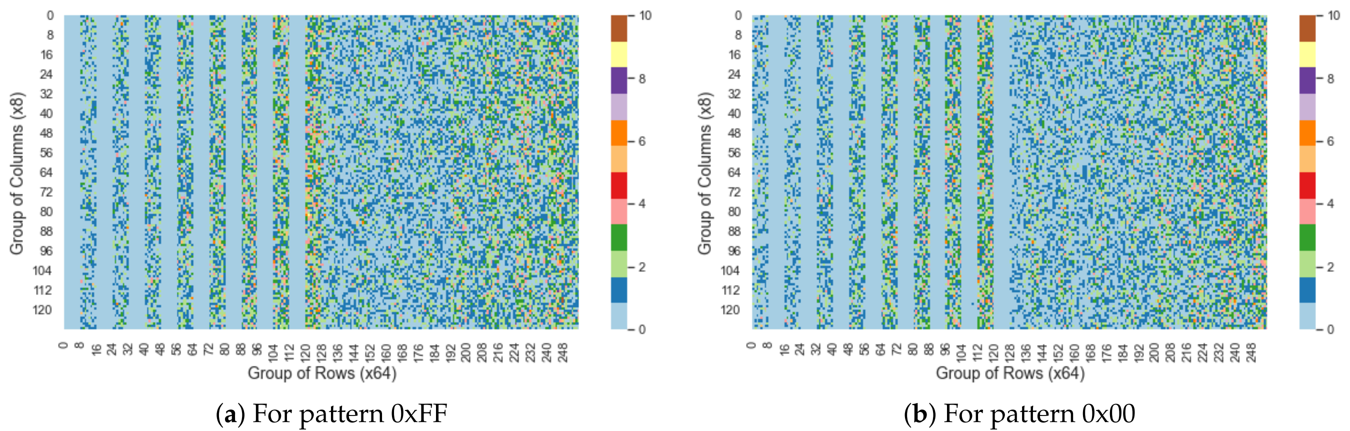
Since the standard retention time is chosen with a high reliability margin [16], the number of retention errors is nonexistent or very few in a retention test with standard retention time. In our experiment, we increase the retention time to t_{ret} = 10 s, 20 s, and 50 s to induce a significant number of bit flips in the cells. We observe that the number of erroneous cells increases by over 100 times when the t_{ret} is increased from 10 to 20 s and by over 1000 times when the retention time is increased to 50 s, as shown in Table 1. It is the extreme case scenario, as t_{ret} values are picked carefully from hands-on experience to result in the maximum transition in the number of retention errors. However, it is expected to see a sharp increase in error at higher retention times as more and more cells lose their charge due to increased leakage.

4.2.2. Dependence on Data Patterns

We collected retention data for the fixed patterns 0x00 and 0xFF, as previous work suggests that this testing is sufficient [31]. Consecutive alternate bits of 1 and 0 (checkerboard patterns) increase the coupling noise with a voltage difference between the neighboring bitlines, which may induce the worst-case data pattern [31].

From the retention error results for 0x00 and 0xFF, the true cells and anti-cells can be identified. Figure 4a,b present the heatmap of retention errors across the rows and columns in a Micron module for patterns 0xFF and 0x00. The x-axis represents the 16,384 rows divided into 256 groups and the y-axis represents 1024 columns divided into 128 groups. Aggregated number of errors for 64 consecutive rows and 8 columns is represented by each color-coded block. We observe the number of errors increase from left to right, and there

are intermittent gaps after every 512 rows. This can be explained by the true and anti-cell architecture of the module. Some rows show a retention error for the 0xFF pattern but not for the 0x00 pattern, which are identified as the true cells. Similarly, the rows that induce error only for pattern 0x00 are identified as anti-cells. The rows that induce error for both cases have a mixture of true-cell and anti-cell architecture.



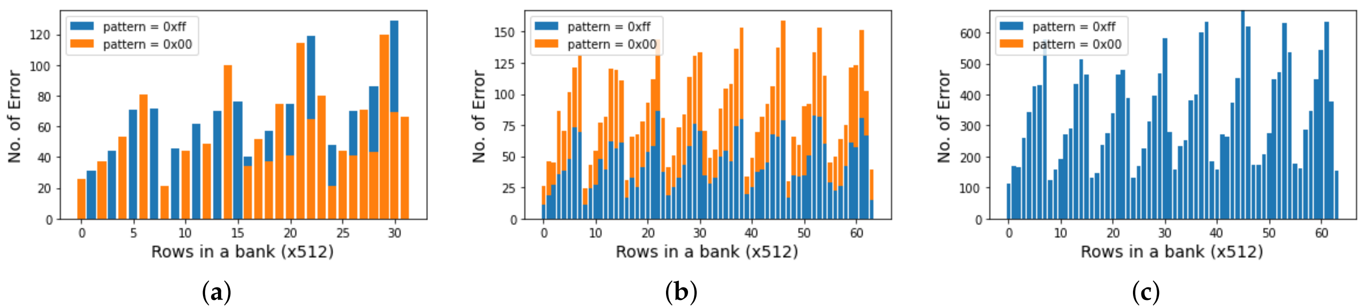
(**a**) For pattern 0xFF (**b**) For pattern 0x00

Figure 4. Distribution of retention errors across rows and columns for pattern 0xFF and 0x00 in a Micron module.

4.2.3. Design-Induced Variations

There exists variation in access latency among different regions in a DRAM chip due to the relative distances of different components from the peripheral structures. Cells located closer to the sense amplifiers can be accessed faster as they experience smaller RC delays [27]. Design-induced variation is static and depends on the physical organization, i.e., it is always determined by the cell's location. DRAMs with the same internal design demonstrate a similar pattern in failure distribution [27].

In Figure 5, we observe the design-induced variation for different DRAM vendors. The x-axis represents the bins of row numbers where every 512 consecutive rows are grouped together. The y-axis represents the aggregated number of bit flips in a group of 512 consecutive rows for patterns 0xFF and 0x00. We observe that the number of failures in the rows gradually increases with increased distance from the peripherals, and this pattern periodically repeats. In addition, the design-induced variation pattern repeats every 4096 pages (eight groups of 512 pages). The repetition is caused by the introduction of a new global sense amplifier and completely dependent on a specific array architecture. We observe two different regions in the Micron design, which is explained in detail in Section 4.2.4.



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

Figure 5. Retention error in a module increases periodically across the row due to design variations. (**a**) Retention error (t_{ret} = 20 s) in a Micron 1 GB DRAM module. (**b**) Retention error (t_{ret} = 20 s) in a Hynix 2 GB DRAM module. (**c**) Retention error (t_{ret} = 20 s) in a Samsung 2 GB DRAM module.

4.2.4. Impact of True Cells and Anti-Cells

We observe the impact of previously discussed architectural variation as true cell and anti-cell for Micron and Hynix DRAM modules in Figure 5a,b. For the Micron 1 GB module, we observe two different architectural variations inside a bank. Each bank consists of 16,000 rows or pages, and each page holds 1000 columns, each with a 64 bit length. We observe a spatial locality of 512 pages in the first half of the bank. In this portion, the first 512 pages induce error for pattern 0x00 and are characterized as anti-cells, and the next 512 consecutive pages are characterized as true cells. On the second half of the pages, we could not find any large-scale granularity, which means that either one or very few consecutive pages are true cells or the next page or a few are anti-cells. On a different note, one or more individual bits in a column could be true cells, and the next one or more could be anti-cells, which is the most likely case to maintain uniformity across the 3D stack of a DDR3 memory module. The Hynix module follows the same architecture as the second half of the Micron module across the whole module (Figure 5b). Figure 5c illustrates the retention error across rows in the case of the Samsung modules. It shows that all the cells are designed with true cells, which explains the errors found only for data pattern 0xFF.

4.3. Aging Effects on Retention Error

From the silicon results, we find that the retention errors are vulnerable to aging; as the devices age, retention errors increase. We present the summary of total retention error and their pattern dependency, retention time dependency, and the impact of accelerated aging for sample DRAM modules from three different vendors in Table 1. The noteworthy observations from the experimental results are highlighted below.

Table 1. Summary of retention error for sample DRAM modules.

Vendor	Chip Number	Retention Time (t_{ret})	Pattern	No. of Errors ($t_0 = 0$ h)	No. of Errors ($t_1 = 8$ h)	% Change ($t_1 = 8$ h)	No. of Errors ($t_2 = 16$ h)	% Change ($t_2 = 16$ h)
Micron	MT8JSF12864HZ-1G4F1	10 s	0x00	802	898	11.97%	953	18.83%
			0xFF	978	1040	6.34%	1144	16.97%
		20 s	0x00	11,130	12,076	8.50%	13,029	17.06%
			0xFF	12,306	12,759	3.68%	14,240	15.72%
		50 s	0x00	202,659	222,322	9.70%	232,886	14.92%
			0xFF	222,006	226,026	1.81%	243,442	9.66%
Hynix	HMT325S6BFR8C-H9	10 s	0x00	531	587	10.55%	743	39.92%
			0xFF	557	610	9.52%	649	16.52%
		20 s	0x00	12,234	14,279	16.72%	16,275	33.03%
			0xFF	12,488	13,541	8.43%	16,075	28.72%
		50 s	0x00	338,707	404,833	19.52%	455,105	34.37%
			0xFF	364,282	391,104	7.36%	458,367	25.83%
Samsung	M471B5773DH0-CH9	10 s	0x00	87	86	−1.15%	88	1.15%
			0xFF	11,759	10,052	−14.52%	10,085	−14.24%
		20 s	0x00	232	230	−0.86%	234	0.86%
			0xFF	209,714	178,239	−15.01%	188,041	−10.33%
		50 s	0x00	653	648	−0.77%	648	−0.77%
			0xFF	5,953,567	5,228,952	−12.17%	5,455,663	−8.36%

4.3.1. Increase in Retention Error with Aging

The number of total erroneous cells (retention errors) for all three different retention times ($t_{ret} = 10$ s, 20 s, 50 s) increases drastically as the devices age (Table 1). We observe about a 5–10% increase in retention error after 8 h of aging and about 15–30% after 16 h

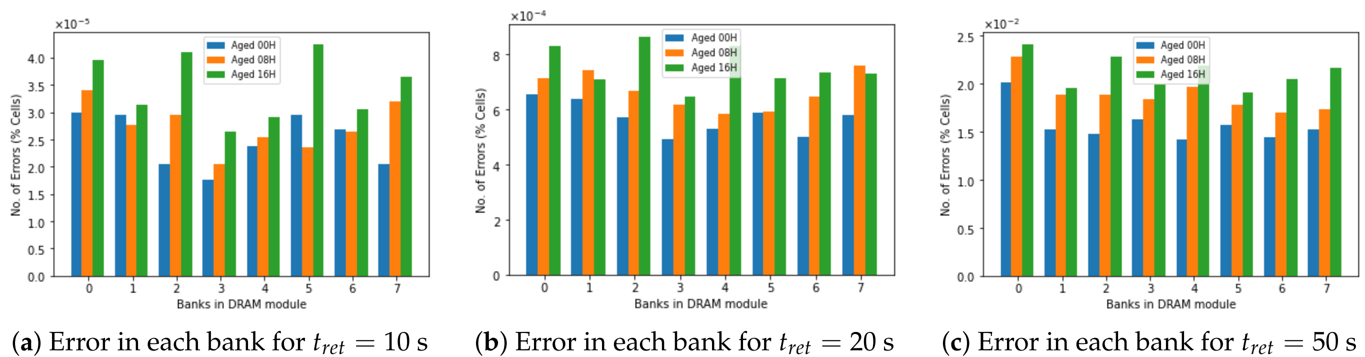
of aging for Micron and Hynix DDR3 modules. When deployed in the field for several months or years, the change in retention error will be significant enough to distinguish an aged module from a new one. In the case of the Samsung modules, we observe a deviation of aging behavior as very few errors are induced for data pattern 0x00 because of design variation. The errors induced for data pattern 0xFF are clustered in a few rows and decrease with aging. In this case, the design-induced variations dominate over aging impacts.

Moreover, for our chosen data patterns, we observe that the 0x00 pattern induces a slightly higher retention error than the 0xFF pattern, as shown in Table 1. Even though there are some variations in retention error due to input patterns, they are in the same range, which illustrates that the number of true cells and anti-cells is almost equal in Micron and Hynix manufactured DRAM chips. On the other hand, the Samsung modules do not show any error for the input pattern 0x00, which indicates that the Samsung chips do not have the architecture of true cells and anti-cells; instead, all the cells are designed following the same true-cell architecture.

4.3.2. Error Distribution with Aging

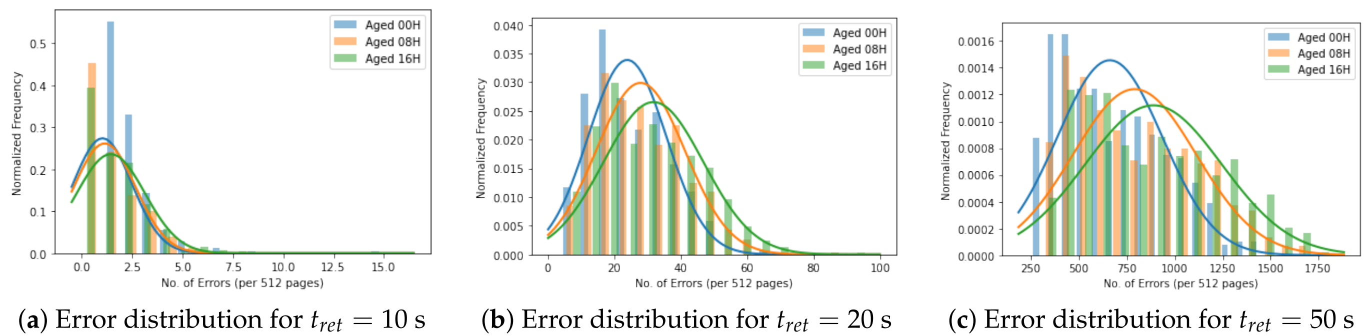
In order to illustrate the aging effects on the retention behavior of DRAM modules, we explore the distribution of error across the banks and group of rows. Figure 6 illustrates the retention errors for pattern 0x00 in each bank of a Hynix 2 GB DRAM module and their corresponding distribution curve. In Figure 6a–c, the number of total retention errors across the banks is plotted at three aging points. We observe that the total retention errors in each bank increases consistently with aging as the wear-out effects occur. This phenomena is consistent for different retention times, as shown in Figure 6a–c.

Retention errors for consecutive 512 rows are aggregated in a group in order to observe the changes across the rows. Figure 7a–c show the histogram and the approximate normal distribution curve of errors in a group of rows. Additionally, Table 2 presents the corresponding mean and standard deviation of the retention error distribution curves. We observe that as the device ages, the mean of the distribution curve shifts rightward. Moreover, the standard deviation of the distribution increases as the variation in total errors in a row across the bank is higher as the device ages. In these figures, we notice that the error distribution of higher retention times is more useful as they induce a high number of errors, and the difference between the new and aged/recycled modules is very evident. These retention-based signatures can be utilized to observe the evidence of device aging and classify the recycled devices by comparing them with the signatures of a new device.



(**a**) Error in each bank for $t_{ret} = 10$ s (**b**) Error in each bank for $t_{ret} = 20$ s (**c**) Error in each bank for $t_{ret} = 50$ s

Figure 6. Retention error for 0x00 pattern in each bank with aging in a Hynix 2GB DRAM module.



(**a**) Error distribution for $t_{ret} = 10$ s (**b**) Error distribution for $t_{ret} = 20$ s (**c**) Error distribution for $t_{ret} = 50$ s

Figure 7. Retention error distribution curve across groups of rows for 0x00 pattern in a Hynix 2 GB DRAM module.

Table 2. Mean and standard deviation of the retention error distribution curve for 0x00 pattern in a Hynix 2 GB DRAM module.

	t_{ret} = 10 s			t_{ret} = 20 s			t_{ret} = 50 s		
	t_0 = 0 h	t_1 = 8 h	t_2 = 16 h	t_0 = 0 h	t_1 = 8 h	t_2 = 16 h	t_0 = 0 h	t_1 = 8 h	t_2 = 16 h
Mean	1.04	1.15	1.45	23.89	27.89	31.78	661.54	790.69	888.87
Std	1.46	1.53	1.69	11.77	13.79	15.03	274.02	321.71	356.43

5. Discussion

5.1. Roles of Signatures in Recycled DRAM Detection

Although we observe an increasing trend in retention error and distribution in DRAM with aging, the aged modules cannot be clearly distinguished from fresh ones with an acceptable level of confidence. This will require us to age the device for an extensive period of time (few months). We also tested for activation latency-based analysis for a few modules and found that retention-based signatures are more useful to observe the wear-out effects. This phenomenon can be explained by the fact that the endurance of DRAM is very high (over 10^{16} write operations) [32] compared to other memory technologies. For example, SRAM cells may experience degradation the entire time the chip is on because the feedback loop keeps the transistors on. On the other hand, DRAM cells are activated only when the wordlines are activated. So, the cells age very slowly, and it takes a long time to experience detectable degradation from the wear-out mechanisms. In case of an activation test, the cells are activated before they are latched properly with the bitline. For this reason, the errors are dominated by the timing violation rather than aging impacts. On the other hand, in the case of the retention test, the cell capacitors are allowed to leak for a long time. As the wear-out mechanisms directly increase the leakage current, retention signatures provide better observability of aging impacts.

5.2. Aging Impacts on Row Hammer

The aging degradation of DRAM cells may increase disturbance errors, such as row-hammer error in a recycled DRAM (if it is already vulnerable) or may become vulnerable (if it was row-hammer-resistant before). The disturbance errors can also be used to detect whether a DRAM is used or fresh, since the total number of cells affected by the disturbance errors changes with time. Increasing the periodic refresh rate can prevent the DRAM cells from leaking. A greater reduction of the refresh period might be required to prevent the system from row-hammer attack due to device aging, since the DRAM cells become leakier with the usage. However, such types of solutions will suffer more energy and performance overhead.

6. Conclusions

This paper presents an experimental study to characterize DRAM aging behavior by analyzing the retention error. From real-time silicon measurements of DRAM devices from multiple vendors, it is concluded that retention-based error can be used as an effective device signature to observe aging effects in DRAM cells. The architectural variation results in true-cell and anti-cell configurations among cells in a DRAM module. These variations can be utilized to speed up the aging process by a factor of two. There are other design-induced variations, such as local variations among pages in a bank that depend on the distance of the cells from the peripheral structures. A considerable increase in retention error and a shift in the distribution curve are found with accelerated aging. It is concluded that an extensive study of the error statistics can be used to classify recycled chips with a threshold value of the average increase in retention error. However, the recycled DRAM detection is challenging because of the slow aging process of DRAM cells and a lack of available resources on the specifications of the modules available. The DRAM memory controller is also limited in many aspects because DDR4, and higher generations of modules, cannot be tested with the platform. Moreover, the testing process is very time consuming, and many memory modules working in the field do not behave as expected from the literature. In the future, the findings of this research can be utilized to develop an independent, non-invasive method to detect recycled DRAM chips without any additional circuitry or cost. The impacts of individual aging mechanisms can be evaluated, and the inconsistent behavior of a few DRAM modules can be explained with the reliability prediction model.

Author Contributions: Conceptualization, B.M.S.B.T. and M.T.R.; Data curation, M.K.B.; Formal analysis, M.K.B.; Funding acquisition, M.T.R.; Investigation, M.K.B. and B.M.S.B.T.; Methodology, M.K.B. and B.M.S.B.T.; Project administration, M.T.R.; Resources, M.T.R.; Software, M.K.B. and B.M.S.B.T.; Supervision, B.M.S.B.T. and M.T.R.; Validation, B.M.S.B.T. and M.T.R.; Visualization, M.K.B.; Writing—original draft, M.K.B.; Writing—review & editing, B.M.S.B.T. and M.T.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Science Foundation under Grant Number CNS-1850241.

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

References

1. Guin, U.; Huang, K.; DiMase, D.; Carulli, J.M.; Tehranipoor, M.; Makris, Y. Counterfeit integrated circuits: A rising threat in the global semiconductor supply chain. *Proc. IEEE* **2014**, *102*, 1207–1228. [CrossRef]
2. Guin, U.; DiMase, D.; Tehranipoor, M. Counterfeit integrated circuits: Detection, avoidance, and the challenges ahead. *J. Electron. Test.* **2014**, *30*, 9–23. [CrossRef]
3. Guin, U.; Asadizanjani, N.; Tehranipoor, M. Standards for hardware security. *Getmobile: Mob. Comput. Commun.* **2019**, *23*, 5–9. [CrossRef]
4. SIA Anti-Counterfeiting Task Force. *Winning the Battle Against Counterfeit Semiconductor Products*; Semiconductor Industry Association: Washington, DC, USA, 2013; pp. 4–8.
5. Kim, Y.; Daly, R.; Kim, J.; Fallin, C.; Lee, J.H.; Lee, D.; Wilkerson, C.; Lai, K.; Mutlu, O. Flipping bits in memory without accessing them: An experimental study of DRAM disturbance errors. *ACM Sigarch Comput. Archit. News* **2014**, *42*, 361–372. [CrossRef]
6. Schaller, A.; Xiong, W.; Anagnostopoulos, N.A.; Saleem, M.U.; Gabmeyer, S.; Katzenbeisser, S.; Szefer, J. Intrinsic rowhammer PUFs: Leveraging the rowhammer effect for improved security. In Proceedings of the 2017 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), Mclean, VA, USA, 1–5 May 2017; pp. 1–7.
7. Huang, K.; Carulli, J.M.; Makris, Y. Parametric counterfeit IC detection via support vector machines. In Proceedings of the 2012 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT), Austin, TX, USA, 3–5 October 2012; pp. 7–12.
8. Guo, Z.; Rahman, M.T.; Tehranipoor, M.M.; Forte, D. A zero-cost approach to detect recycled SoC chips using embedded SRAM. In Proceedings of the 2016 IEEE international symposium on hardware oriented security and trust (HOST), McLean, VA, USA, 3–5 May 2016; pp. 191–196.
9. Kumari, P.; Talukder, B.B.; Sakib, S.; Ray, B.; Rahman, M.T. Independent detection of recycled flash memory: Challenges and solutions. In Proceedings of the 2018 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), Washington, DC, USA, 30 April–4 May 2018; pp. 89–95.

10. Tehranipoor, F.; Karimian, N.; Yan, W.; Chandy, J.A. Investigation of DRAM PUFs reliability under device accelerated aging effects. In Proceedings of the 2017 IEEE International Symposium on Circuits and Systems (ISCAS), Baltimore, MD, USA, 28–31 May 2017; pp. 1–4.
11. Talukder, B.B.; Menon, V.; Ray, B.; Neal, T.; Rahman, M.T. Towards the Avoidance of Counterfeit Memory: Identifying the DRAM Origin. In Proceedings of the 2020 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), San Jose, CA, USA, 7–11 December 2020; pp. 111–121.
12. Najdi, A.; Rossi, D.; Tenentes, V. Analysis on Retention Time and Adaptive Refresh in Embedded DRAMs with Aging Benefits. In Proceedings of the 2019 IEEE 25th International Symposium on On-Line Testing and Robust System Design (IOLTS), Rhodes, Greece, 1–3 July 2019; pp. 281–286.
13. Fieback, M. Dram Reliability: Aging Analysis and Reliability Prediction Model. Master's Thesis, Delft University of Technology, Delft, The Netherlands, 2017.
14. Talukder, B.B.; Ray, B.; Forte, D.; Rahman, M.T. PreLatPUF: Exploiting DRAM latency variations for generating robust device signatures. *IEEE Access* **2019**, *7*, 81106–81120. [CrossRef]
15. Liu, J.; Jaiyen, B.; Veras, R.; Mutlu, O. RAIDR: Retention-aware intelligent DRAM refresh. *ACM SIGARCH Comput. Archit. News* **2012**, *40*, 1–12. [CrossRef]
16. Hamamoto, T.; Sugiura, S.; Sawada, S. On the retention time distribution of dynamic random access memory (DRAM). *IEEE Trans. Electron Devices* **1998**, *45*, 1300–1309. [CrossRef]
17. Swift, G.; Padgett, D.; Johnston, A. A new class of single event hard errors [DRAM cells]. *IEEE Trans. Nucl. Sci.* **1994**, *41*, 2043–2048. [CrossRef]
18. Venkatesan, R.K.; Herr, S.; Rotenberg, E. Retention-aware placement in DRAM (RAPID): Software methods for quasi-non-volatile DRAM. In Proceedings of the The Twelfth International Symposium on High-Performance Computer Architecture, 2006, Austin, TX, USA, 11–15 February 2006; pp. 155–165.
19. Wang, S.; Bojnordi, M.N.; Guo, X.; Ipek, E. Content aware refresh: Exploiting the asymmetry of DRAM retention errors to reduce the refresh frequency of less vulnerable data. *IEEE Trans. Comput.* **2018**, *68*, 362–374. [CrossRef]
20. Neamen, D.A. *Semiconductor Physics and Devices: Basic Principles*; McGraw-Hill: New York, NY, USA, 2012.
21. Quader, K.N.; Li, C.; Tu, R.; Rosenbaum, E.; Ko, P.; Hu, C. A new approach for simulation of circuit degradation due to hot-electron damage in NMOSFETs. In Proceedings of the International Electron Devices Meeting 1991 [Technical Digest], Washington, DC, USA, 8–11 December 1991; pp. 337–340.
22. Gielen, G.; De Wit, P.; Maricau, E.; Loeckx, J.; Martin-Martinez, J.; Kaczer, B.; Groeseneken, G.; Rodriguez, R.; Nafria, M. Emerging yield and reliability challenges in nanometer CMOS technologies. In Proceedings of the Conference on Design, Automation and Test in Europe, 2008, Munich, Germany, 10–14 March, 2008; pp. 1322–1327.
23. Galloway, K.; Schrimpf, R. MOS device degradation due to total dose ionizing radiation in the natural space environment: A review. *Microelectron. J.* **1990**, *21*, 67–81. [CrossRef]
24. Bezhenova, V.; Michalowska-Forsyth, A.M. Effects of ionizing radiation on integrated circuits. *e & i Elektrotechnik Und Informationstechnik* **2016**, *133*, 39–42.
25. Keane, J.; Wang, X.; Persaud, D.; Kim, C.H. An all-in-one silicon odometer for separately monitoring HCI, BTI, and TDDB. *IEEE J. Solid-State Circ.* **2010**, *45*, 817–829. [CrossRef]
26. Schroeder, B.; Pinheiro, E.; Weber, W.D. DRAM errors in the wild: A large-scale field study. *ACM SIGMETRICS Perform. Eval. Rev.* **2009**, *37*, 193–204. [CrossRef]
27. Lee, D.; Khan, S.; Subramanian, L.; Ghose, S.; Ausavarungnirun, R.; Pekhimenko, G.; Seshadri, V.; Mutlu, O. Design-induced latency variation in modern DRAM chips: Characterization, analysis, and latency reduction mechanisms. *Proc. Acm Meas. Anal. Comput. Syst.* **2017**, *1*, 1–36. [CrossRef]
28. Kim, G.J.; Lee, N.H.; Kim, J.; Seok, J.E.; Lee, Y. Effect of DC/AC stress on the reliability of cell capacitor in DRAM. *Microelectron. Reliab.* **2018**, *88*, 179–182. [CrossRef]
29. Hassan, H.; Vijaykumar, N.; Khan, S.; Ghose, S.; Chang, K.; Pekhimenko, G.; Lee, D.; Ergin, O.; Mutlu, O. SoftMC: A flexible and practical open-source infrastructure for enabling experimental DRAM studies. In Proceedings of the 2017 IEEE International Symposium on High Performance Computer Architecture (HPCA), Austin, TX, USA, 4–8 February 2017; pp. 241–252.
30. inTEST Thermal, Solutions. ATS-605 Thermostream Datasheet and Specifications. Available online: https://www.intestthermal.com/temptronic/thermostream (accessed on 5 March 2022).
31. Liu, J.; Jaiyen, B.; Kim, Y.; Wilkerson, C.; Mutlu, O. An experimental study of data retention behavior in modern DRAM devices: Implications for retention time profiling mechanisms. *ACM SIGARCH Comput. Archit. News* **2013**, *41*, 60–71. [CrossRef]
32. Moyer, B. What Is DRAM's Future? *Semiconductor Engineering*, 9 April 2020. Available online: https://semiengineering.com/what-is-drams-future/#:~:text=DRAM%20does%20have%20infinite%20endurance,in%20something%20over%20300%20years (accessed on 5 March 2022).