A Novel Insight Into the Vulnerability of DDR4 DRAM Cells Across Multiple Hammering Settings

Ranyang Zhou, *Graduate Student Member, IEEE*, Jacqueline Liu, *Student Member, IEEE*,
Nakul Kochar, *Member, IEEE*, Sabbir Ahmed, *Student Member, IEEE*, Adnan Siraj Rakin, *Member, IEEE*,
and Shaahin Angizi, *Senior Member, IEEE*

***Abstract*—RowHammer stands out as a prominent example, potentially the pioneering one, showcasing how a failure mechanism at the circuit level can give rise to a significant and pervasive security vulnerability within systems. Prior research has approached RowHammer attacks within a static threat model framework. Nonetheless, it warrants consideration within a more nuanced and dynamic model. This letter presents a low-overhead DRAM RowHammer vulnerability profiling technique, which utilizes innovative test vectors for categorizing memory cells into distinct security levels. The proposed test vectors intentionally weaken the spatial correlation between the aggressors and victim rows before an attack for evaluation, thus aiding designers in mitigating RowHammer vulnerabilities in the mapping phase. While there has been no previous research showcasing the impact of such profiling to our knowledge, our study methodically assesses 128 commercial DDR4 DRAM products. The results uncover the significant variability among chips from different manufacturers in the type and quantity of RowHammer attacks that can be exploited by adversaries.**

***Index Terms*—DRAM, memory security, RowHammer.**

I. Introduction

RECENT research has demonstrated that adversaries can exploit the RowHammer vulnerability present in DRAM to systematically and precisely manipulate bits across diverse applications, including proficiently trained neural networks, resulting in a notable impact on accuracy [1], [2]. Illustrated in Fig. 1(a), such so-called bit-flip attacks (BFAs) can reduce the accuracy of an 8-bit quantized ResNet-34 on the ImageNet from 73.1% to 0% by targeting only 5 bits. Fig. 1(b) reports that the RowHammer threshold has experienced a notable decline in recent years. For instance, on LPDDR4 (new), the attacker requires ∼4.5 times fewer hammer counts (HCs) compared to DDR3 (new) [3].

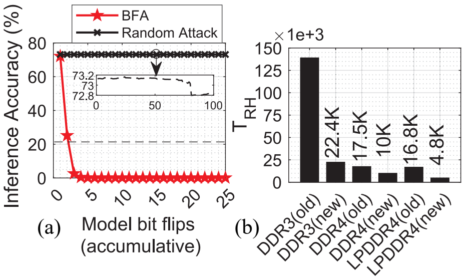

Fig. 1. (a) Targeted versus random bit-flipping for an 8-bit quantized ResNet-34 on ImageNet. (b) RowHammer thresholds [3].

To mitigate RowHammer attacks, comprehensive investigation, and analysis of pertinent influencing factors are imperative. As research progresses, error correction code (ECC) techniques [4], [5] have been developed across various directions to combat RowHammer attacks. Intel's pTRR [6] and various research work propose a proactive strategy involving the monitoring of row activations, termed HC. The memory controller tracks HC and initiates refresh cycles on victim rows once the number of row activations surpasses a predefined maximum activate count (MAC) threshold (T_{MAC}), typically stored on the serial presence detect (SPD) chip within the DRAM module [7]. Previous studies have addressed attacks under a static threat model, emphasizing fixed parameters. Kim et al. [8] were the pioneers in conducting a study on the characteristics of RowHammer bit-flips in DDR3 modules. With the prospect of having a RowHammer-less landscape, DDR4 modules have been introduced. One of the recent works exploring the multisided fault injection model is TRRespass [7]. Multiple software and hardware mitigation mechanisms have been also proposed to reduce the impact of RowHammer-based attacks [8], [9]. The hardware-based research efforts can be classified into two categories, i.e., *victim-focused* mechanism with probabilistic refreshing (e.g., PRA [10] and ProTRR [9]) and *aggressor-focused* mechanism by counting activations (e.g., TRR [11] and Hydra [12]).

Acknowledging the evolving nature of security threats, we advocate for a more sophisticated and adaptable approach in this letter. In contrast to static models, a dynamic framework accommodates the fluidity of attack vectors and defense mechanisms, thus providing a more comprehensive understanding of RowHammer vulnerabilities. By embracing this perspective, researchers can better anticipate emerging threats and devise effective countermeasures to safeguard against RowHammer. In this letter, we introduce a novel technique for profiling DRAM RowHammer vulnerabilities with minimal overhead that employs innovative test vectors to classify memory cells into different security levels. The main contributions of this letter are as follows.

1) We demonstrate that the bit-flip induced by RowHammer attacks is intricate and variable, necessitating varied analyses associated with different patterns applied in the RowHammer attack model.

Manuscript received 20 August 2024; accepted 20 August 2024. Date of current version 26 November 2024. This work was supported in part by the National Science Foundation under Grant 2228028. This manuscript was recommended for publication by A. Shrivastava. *(Corresponding author: Shaahin Angizi.)*

Ranyang Zhou, Nakul Kochar, and Shaahin Angizi are with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: rz26@njit.edu; nk548@njit.edu; shaahin.angizi@njit.edu).

Jacqueline Liu, Sabbir Ahmed, and Adnan Siraj Rakin are with the Department of Computer Science, State University of New York at Binghamton, Binghamton, NY 13790 USA (e-mail: jliu28@binghamton.edu; sahmed9@binghamton.edu; arakin@binghamton.edu).

Digital Object Identifier 10.1109/LES.2024.3449232

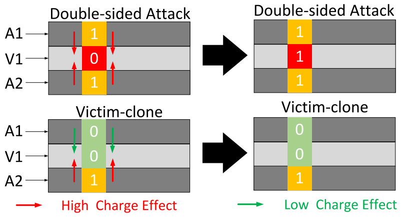

Fig. 2. DB and VC models.

2) We propose a comprehensive classification of DRAM cells referred to as *cell's security level* within the chip to enhance the visibility of the impact of RowHammer attacks.
3) We experimentally reveal substantial variability in the robustness of cells across 128 chips sourced from seven major DRAM manufacturers.

II. DRAM Security Level: Novel Insight

A bit-flip occurs exclusively when there is a disparity in the bit values of adjacent rows. This raises the query regarding the differentiation of data among DRAM rows, a consequence of manufacturers' topology techniques. Consequently, the likelihood of adjacent rows differing from the target row on every bit is exceedingly low, resulting in numerous bits within the victim row sharing identical values with those in the adjacent row. According to this hypothesis, certain bits remain immune to flipping when adversaries employ a single-sided (SG) attack strategy. Nonetheless, in the double-side (DB) attack model, the scenario becomes intricate. Ideally, the two assailant rows would exhibit diversity, each contrasting with the victim row on every bit. However, in specific instances, the sheer abundance of distinct bits complicates this ideal scenario.

Previous studies [5], [7], [12] have overlooked comparable specifics, and their assessment of RowHammer relies on analyzing the subsequent conditions: 1) complete dissimilarity between all bits of the attack row and the victim row and 2) conducting experiments using real DRAM storage data. However, owing to technical disparities among various manufacturers, this data pattern can be perceived as random. To enhance comprehension of the factors contributing to bit-flipping in RowHammer attacks, we decided to create a new research model. As shown in Fig. 2, we assume the DB attack is based on the ideal case where each bit of the aggressor rows (A1 and A2) differs from the victim (V1). We hypothesize that both cells in the attacking row exert a significant charging effect on the cells in the victim row. Victim-Clone (VC) is our proposed attack model to make the victim row suffer less when the DRAM is under the DB attack. Leveraging this model, we can focus on a more detailed study of the effects of leakage between cells and prolong the stability of cells within the victim's row, preventing them from experiencing bit-flips for an extended period. The VC model essentially copies the victim row to one of the aggressor rows to ensure that each bit of the victim row is only affected by one adjacent flipped bit. As shown in Fig. 2, in this model, the cell in A1 has a low charge effect, and that in A2 has a high charge effect.

Based on the findings reported in our preceding study [13], a direct correlation exists between the increment of HC and the observed rise in bit-flip occurrences in DRAM This phenomenon signifies an escalating number of cells susceptible to charge leakage as HC values increase. Alternatively, a granular examination of individual HC values unveils distinct patterns in cell presence across different levels. Some cells demonstrate consistent presence across multiple HC levels, while others exhibit sporadic or negligible presence. Utilizing a color-coded scheme to represent DRAM cell frequencies at varying HC levels allows for the visualization of these patterns, facilitating a more comprehensive understanding of DRAM vulnerability to RowHammer attacks. As shown in Fig. 3, assume we collect samples from the same chip subjected to RowHammer attacks at various HC levels. In every tier, we emphasize the cells where bit-flips occur. As the HC escalates from upper to lower tiers, the number of these cells generating bit-flips rises. We note that cells causing bit-flips at lower HC levels persist in generating bit-flips at higher HC levels, implying their consistency across varying HC levels. Therefore, the more frequently a flipped cell appears at all levels, the more vulnerable it is. In other words, if some cells appear in different HC levels simultaneously, the highlighted color will be darker. Thus, the color bars in Fig. 3 can represent the vulnerability of cells. The four colors from left to right (from bright yellow to dark red) represent the degree of vulnerability from low to high. This model can be exploited for the following reasons.

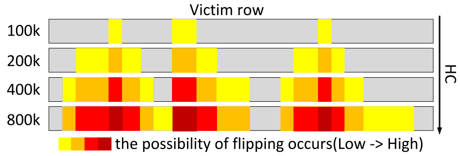

Fig. 3. Vulnerability of cells associated with the HC.

1) To empirically analyze significant variability among chips from different manufacturers. Consequently, we aim to investigate whether this discrepancy correlates with the quantity of highly vulnerable cells within the chip.
2) To investigate variations in the rate at which the number of bit-flips increases with rising HC levels. Therefore, this model facilitates a detailed examination of the differences between cells from manufacturers.
3) To explore RowHammer attack modes yield outcomes. This model enables us to evaluate the resilience of cells to various attack modes.

III. Experiments

Framework Setup and Testing Infrastructure: We test the DRAM chips with DRAM-Bender [14]. The testing infrastructure in Fig. 4 consists of the Alveo U200 Data Center Accelerator Card [15] as the FPGA that accepts DDR4 modules and runs the test programs by sending DDR4 command traces generated by the host machine. The idea is to take control of memory modules for DDR4 interfaces with straightforward high-level programming to test and run the generated programs on the host machine. Besides, to have a fair comparison among various under-test DRAM chips, the temperature is kept below 30 °C with INKBIRDPLUS 1800-W temperature controller.

Minimizing Interference: Before implementing the attack, DRAM refresh [16] and rank-level ECC are disabled to minimize interference with RowHammer bit-flips following [14]. However, proprietary RowHammer protection techniques (e.g., Target Row Refresh [7], [11]) are in place.

Chips Tested: To profile DRAM cell vulnerabilities, the experiments are conducted on a range of 128 commercialized DRAM chips from seven different manufacturers (mf.) as listed in Table I with various die densities and die revisions.

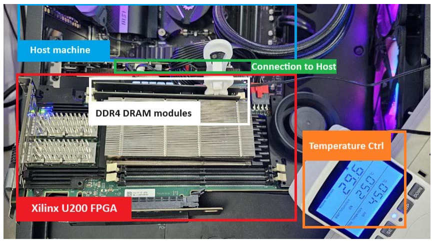

Fig. 4. Our testing infrastructure for DDR4 modules.

TABLE I
UNDER-TEST DRAM CHIPS

Vendor	#Chips	Freq (MHz)	Die rev.	Org.	Date
mf-A (Crucial 16 GB)	16	3200	C	x8	N/A
mf-B (Kingston 16GB)	16	2666	G	x8	2152
mf-C (Micron 16GB)	16	2133	B	x4	2126
mf-D (NEMIX 16GB)	16	2133	B	x4	1733
mf-E (SK Hynix 16GB)	16	2400	A	x8	1817
mf-F (Patriot Viper 16GB)	16	3600	C	x8	N/A
mf-G (Samsung 16GB)	16	2400	B	x8	2053

Our findings stem from a detailed examination of a single row. Before selecting this row, we conducted thorough testing for every chip and observed varying numbers of bit-flips, ranging from 0 (indicating minimal activations) to 200 (indicating maximal activations). Following initial tests, we selected one of the patterns tested for our experiment. To ensure precision and minimize variability, we repeated each activation count 100 times. This approach mitigates fluctuations in our data.

Analysis of the Results: Fig. 5 represents the comprehensive analysis results of the security levels of DRAM cells. In every plot, there are three curves for different RowHammer attack models, i.e., DB, SG, and VC. The x-axis denotes HC, and the y-axis represents the number of cells at which bit-flip occurred. The typical t_{RAS} values for DDR4 memory modules can range from approximately 36 to 48 t_{CK} [17], although these values may vary depending on the module's speed rating (e.g., DDR4-2133, DDR4-2400, DDR4-3200, etc.). For example, the duration of a clock cycle for DDR4-2400 memory can be calculated as $t_{\mathrm{CK}} = (1/2400\mathrm{MT/s})$. In our design, each t_{RAS} comprises three components: `ACT`, `Sleep`, and `PRE`, where `Sleep` is set to $5t_{\mathrm{CK}}$. In order to more accurately emulate real-world scenarios, we set a maximum limit of 1M for the HC.

Given that we suspended the DRAM refresh command, it became necessary to manually account for retention time. So in a refresh window (t_{REF}) the maximum number of HC must be less than $(t_{\mathrm{REF}}/t_{\mathrm{RAS}}) = 1.37\mathrm{M}$. Practically, the application cannot be composed entirely of activations, so we limit the number of activations used for RowHammer to 1M. Here, we list our key observations regarding the under-test chips.

Obs.#1: Compared with DB model, VC model cannot effectively reduce the number of bit-flips.

As discussed, the VC attack model is a way to make the victim row less vulnerable by copying the victim row to one of the aggressor rows. However, based on the empirical findings, it is evident that only chips from three manufacturers exhibit improvement following the implementation of the VC. As depicted in Fig. 5(b), (e), and (f), upon reaching the HC limit (1M), the bit-flips induced by VC decreased by approximately 25% compared to DB yet remained over four times more than those induced by SG. This observation contradicts our initial hypothesis: within the DB model, replicating the victim row onto one of the aggressor rows does not significantly decrease the frequency of bit-flips. We can draw a new conclusion from this: when the attacker ensures that one bit in the aggressor rows differs from the victim row, they can efficiently flip the one in the victim row. Confirming the prior reports, the DB attack is more likely to produce a bit-flip than the SG attack. As shown in Fig. 5(a), (d), and (g), the results of VG and DB almost overlap, meaning that these cells can produce bit-flip as long as at least one cell in the adjacent row differs from it.

Obs.#2: Various cells demonstrate diverse levels of resistance to various attack models.

Within the same chip, certain cells are susceptible to RowHammer attacks, whereas others remain unaffected. However, determining the susceptibility of a cell poses a challenge. To address this, we employ a visual approach for classification. We have opted to use a four-level scale, ranging from level 1 to level 4, to denote the extent of cell vulnerability. Lower levels indicate a lower likelihood of bit-flips. Take Fig. 5(b), (e), and (f) as examples, among these chips, we posit that if a cell succumbs to SG, it can be deemed the most vulnerable to attack. Consequently, when HC is 1M, we classify all cells that induce bit-flips as level 4. Next, we consider cells that do not induce bit-flips in SG but exhibit them in VC. We categorize these cells as level 3. If cells with high vulnerability manifest bit-flips in low-threat attacks, they are also likely to experience bit-flips in high-threat attacks. Consequently, when HC is 1M, we derive the level 3 count by subtracting the total number of bit-flips in VC from the total number of bit-flips in SG. Applying the same principle, we classify cells exhibiting behaviors between DB and VC as level 2. Finally, if cells withstand even the DB attack, we classify them as level 1. Excluding the chips from these three manufacturers, as shown in Fig. 5(a), (d), and (g), due to the scarcity of cells between DB and VC, we delete level 2 and keep level 3.

Obs.#3: Tailored DRAM protection mechanisms, designed according to specific chip topologies, will be necessary and more efficient.

From our experiments, we discovered significant variations in RowHammer attacks across chips from different manufacturers, likely due to distinct manufacturing processes. Consequently, we contend that designing tailored defense mechanisms based on the specific characteristics of individual chips may yield greater effectiveness. For example, considering Fig. 5(c), (d), and (g), which has a significant proportion of level 1 cells, it may opt to employ a defense strategy targeting levels 3, 2, and 1. In Fig. 5(b), (e), and (f), the primary characteristic is the exceedingly low number of cells in level 4, coupled with a larger number of cells in levels 3 and 2. As such, implementing a defense mechanism against DB attacks could be appropriate. Finally, in Fig. 5(a), all the levels are average, so the counter-based defense mechanisms can be recommended. While there are multiple approaches to defending against RowHammer attacks, the most straightforward method involves identifying the factors that render cells deferentially vulnerable to attack. Enhancing these influencing factors will represent the most effective defense against RowHammer attacks.

Obs.#4: The stability of cells in chips varies among different manufacturers.

Here, we introduced a cell classification method, yet the stability of cells influences our classification to some extent. Stability refers to the fluctuation range in the number of cells that induce bit-flips when HC is at a specific value. A broader range indicates lower stability. For instance, in real-world scenarios, cells may occasionally trigger bit-flips once HC reaches a particular value due to interference from various

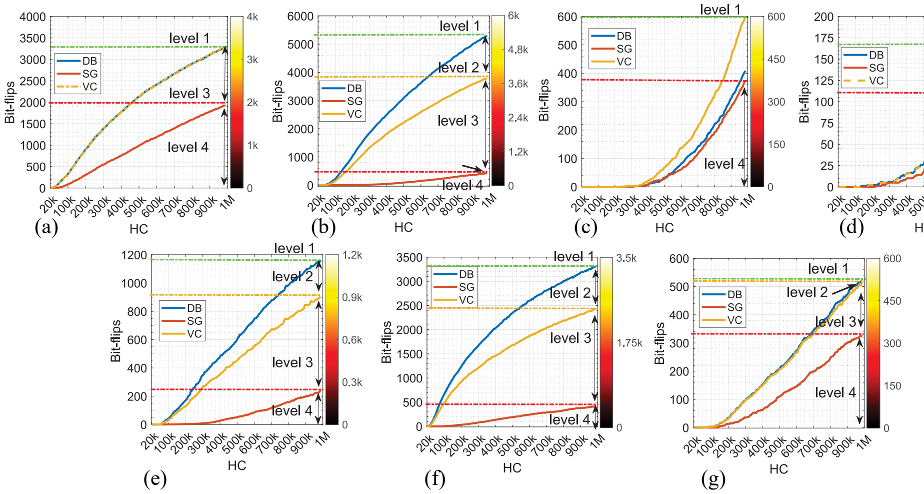

Fig. 5. Analysis of the security levels of cells on (a) mf-A, (b) mf-B, (c) mf-C, (d) mf-D, (e) mf-E, (f) mf-F, and (g) mf-G.

TABLE II
NUMBER OF REQUIRED ITERATIONS FOR BFA ATTACK [2] TO DEGRADE ACCURACY TO A RANDOM GUESS LEVEL

Vendor	Single-sided attack	Victim-Clone attack	Double-sided attack
mf-A	18	15	15
mf-B	50	17	14
mf-C	20	46	55
mf-D	40	27	27
mf-E	74	34	20
mf-F	49	14	15
mf-G	28	19	19

factors. However, in experimental settings, no bit-flips occur. Fig. 5(d) and (g) serves as examples of low stability. We observe that the curves for chips of these two manufacturers exhibit irregular fluctuations, indicating significant variability in the number of bit-flips at certain HC values. In comparison, other chips are relatively stable.

DNN Weight Attack: To further analyze the effectiveness of the conducted study in DNN application, we incorporate the three different attack models/levels, i.e., SG, VC, and DB, into the popular BFA attack framework [2], [18] via only targeting cells that will succumb to the corresponding attack levels, and conduct the adjusted BFA attack on a quantized ResNet-20 trained on CIFAR-10 [19]. Table II displays the number of iterations needed to degrade model accuracy to a random guess level (i.e., 10%) under the three distinct attack strategies across all under-test DRAM chips. We observe that the numbers of required iterations vary extraordinarily across different chips. Echoing the observations from Fig. 5(a), (d), and (g), where the curves of VC and DB overlap, the number of required iterations are identical (15, 27, and 19, respectively). Generally, an SG row hammer requires more bit-flips to achieve the attacker's objective on most chips.

IV. CONCLUSION

This letter introduces a mechanism for experimental DRAM RowHammer vulnerability profiling. This mechanism is proposed to make the analysis of the RowHammer attack model more comprehensive and visible. We explore various RowHammer models to reintroduce a more authentic setting, addressing a previously overlooked aspect in prior research. The revised model provides a more nuanced understanding of performance variations across different manufacturers' chips, highlighting the necessity for a dynamic, rather than static, approach to the RowHammer problem.

REFERENCES

[1] S. Hong, P. Frigo, Y. Kaya, C. Giuffrida, and T. Dumitras, "Terminal brain damage: Exposing the graceless degradation in deep neural networks under hardware fault attacks," in *Proc. 28th USENIX Conf. Secur. Symp.*, 2019, pp. 497–514.

[2] A. S. Rakin, Z. He, and D. Fan, "Bit-flip attack: Crushing neural network with progressive bit search," in *Proc. ICCV*, 2019, pp. 1211–1220.

[3] J. Woo, G. Saileshwar, and P. J. Nair, "Scalable and secure row-swap: Efficient and safe row hammer mitigation in memory systems," 2022, *arXiv:2212.12613*.

[4] O. Mutlu and J. S. Kim, "Rowhammer: A retrospective," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 8, pp. 1555–1571, Aug. 2020.

[5] E. Lee, I. Kang, S. Lee, G. E. Suh, and J. H. Ahn, "TWiCe: Preventing row-hammering by exploiting time window counters," in *Proc. ISCA*, 2019, pp. 385–396.

[6] M. Kaczmarski, *Thoughts on Intel Xeon e5-2600 v2 Product Family Performance Optimisation–Component Selection Guidelines*, Intel, Santa Clara, CA, USA, 2014.

[7] P. Frigo et al., "TRRespass: Exploiting the many sides of target row refresh," in *Proc. IEEE Symp. Security Privacy (SP)*, 2020, pp. 747–762.

[8] Y. Kim et al., "Flipping bits in memory without accessing them: An experimental study of DRAM disturbance errors," *ACM SIGARCH Comput. Archit. News*, vol. 42, no. 3, pp. 361–372, 2014.

[9] M. Marazzi, P. Jattke, F. Solt, and K. Razavi, "ProTRR: Principled yet optimal in-DRAM target row refresh," in *Proc. IEEE Symp. Security Privacy (SP)*, 2022, pp. 735–753.

[10] D.-H. Kim, P. J. Nair, and M. K. Qureshi, "Architectural support for mitigating row hammering in DRAM memories," *IEEE Comput. Archit. Lett.*, vol. 14, no. 1, pp. 9–12, Jan.–Jun. 2014.

[11] H. Hassan, Y. C. Tugrul, J. S. Kim, V. Van der Veen, K. Razavi, and O. Mutlu, "Uncovering in-DRAM rowhammer protection mechanisms: A new methodology, custom rowhammer patterns, and implications," in *Proc. MICRO*, 2021, pp. 1198–1213.

[12] M. Qureshi, A. Rohan, G. Saileshwar, and P. J. Nair, "Hydra: Enabling low-overhead mitigation of row-hammer at ultra-low thresholds via hybrid tracking," in *Proc. ISCA*, 2022, pp. 699–710.

[13] R. Zhou, J. Liu, S. Ahmed, N. Kochar, A. S. Rakin, and S. Angizi, "Threshold breaker: Can counter-based RowHammer prevention mechanisms truly safeguard DRAM?" 2023, *arXiv:2311.16460*.

[14] A. Olgun et al., "DRAM bender: An extensible and versatile FPGA-based infrastructure to easily test state-of-the-art DRAM chips," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 42, no. 12, pp. 5098–5112, Dec. 2023.

[15] (Xilinx inc., San Jose, CA, USA). *Xilinx Alveo u200 FPGA Board*. 2021. [Online]. Available: https://www.xilinx.com/products/boards-and-kits/alveo.html

[16] "JESD79-4C: DDR4 SDRAM standard." 2020. [Online]. Available: https://www.xilinx.com/products/boards-and-kits/alveo.html

[17] H. Choi, D. Hong, J. Lee, and S. Yoo, "Reducing DRAM refresh power consumption by runtime profiling of retention time and dual-row activation," *Microprocess. Microsyst.*, vol. 72, Feb. 2020, Art. no. 102942.

[18] F. Yao, A. S. Rakin, and D. Fan, "DeepHammer: Depleting the intelligence of deep neural networks through targeted chain of bit flips," in *Proc. 29th USENIX Conf. Secur. Symp.*, 2020, pp. 1463–1480.

[19] A. Krizhevsky, V. Nair, and G. Hinton. "The CIFAR-10 dataset." 2014. [Online]. Available: http://www.cs.toronto.edu/kriz/cifar.html