

# HiFi-DRAM: Enabling High-fidelity DRAM Research by Uncovering Sense Amplifiers with IC Imaging

Michele Marazzi\*, Tristan Sachsenweger\*, Flavien Solt\*, Peng Zeng\*,  
Kubo Takashi†, Maksym Yarema\* and Kaveh Razavi\*

\*ETH Zürich †Zentel Japan

**Abstract**—DRAM vendors do not disclose the architecture of the sense amplifiers deployed in their chips. Unfortunately, this hinders academic research that focuses on studying or improving DRAM. Without knowing the circuit topology, transistor dimensions, and layout of the sense amplifiers, researchers are forced to rely on best guesses, impairing the fidelity of their studies. We aim to fill this gap between academia and industry for the first time by performing Scanning Electron Microscopy (SEM) with Focused Ion Beam (FIB) on recent commodity DDR4 and DDR5 DRAM chips from the three major vendors. This required us to adequately prepare the samples, identify the sensing area, and align images from the different FIB slices. Using the acquired images, we reverse engineer the circuits, measure transistor dimensions and extract physical layouts of sense amplifiers — all previously unavailable to researchers. Our findings show that the commonly assumed classical sense amplifier topology has been replaced with the more sophisticated offset-cancellation design by two of the three major DRAM vendors. Furthermore, the transistor dimensions of sense amplifiers and their revealed physical layouts are significantly different than what is assumed in existing literature. Given commodity DRAM, our analysis shows that the public DRAM models are up to 9x inaccurate, and existing research has up to 175x error when estimating the impact of the proposed changes. To enable high-fidelity DRAM research in the future, we open source our data, including the reverse engineered circuits and layouts.

## I. INTRODUCTION

DRAM is the target of many research efforts from academia every year with the *sense amplifier* being the fundamental section most commonly modified and simulated [2], [5], [13]–[15], [22], [26], [28], [29], [58], [66], [68], [74], [75], [78], [85], [87], [88], [93], [94], [97], [111]–[113]. Unfortunately, DRAM vendors keep the internal architecture of sense amplifiers in their chips a secret. As a result, researchers are forced to make assumptions and speculate over crucial design factors, impacting the accuracy of their results. We aim to fill this gap by imaging, and subsequently reverse engineering, sense amplifier circuits on modern DDR4 and DDR5 devices.

Existing research commonly assumes commodity DRAM to employ the classical sense amplifier circuit, which we show not to be the case in DRAM chips from two of three major DRAM vendors. Instead, their sense amplifiers include new components and events to perform offset-compensating operations for reliable DRAM operation with smaller technology nodes. Furthermore, we obtain crucial information concerning deployed sense amplifiers, such as transistor dimensions and physical layout. Using this knowledge, we systematically analyze the feasibility and accuracy of DRAM studies spanning a decade

of research. We find that the majority of these studies make inaccurate assumptions about sense amplifiers, resulting in significant errors when estimating the impact of their proposed changes. We formulate recommendations based on our findings to enable high-fidelity DRAM research in the future.

**Research accuracy.** DRAM provides cheap and low-latency memory based on capacitors. The transition between the analogue world of capacitors and the digital world is performed by *sense amplifiers*. Sense amplifiers enhance the extremely weak signals stored in capacitors while adhering to strict timings [38]. Their design and topology must be reliable towards process manufacturing variability and noise, yet strongly optimized to keep high die efficiency [32]. Research based on modifying sense amplifiers depends on three factors for its accuracy and validity: (i) the employed sense amplifier circuit, commonly assumed to be the classical design, (ii) the transistor dimensions, as circuits with overly large transistors will be optimistic towards their reliability, and (iii) complying with existing layouts, given that the sense amplifier region is a highly optimized area, adding new components should be done with care to achieve realistic area overheads. Unfortunately, information on these crucial factors that are necessary for high-fidelity DRAM research is not publicly available to researchers. This paper aims to fill this gap for the first time.

**DRAM reverse engineering.** We perform high-resolution chip imaging to reverse engineer the sense amplifier region in six commodity DDR4 and DDR5 devices from the three major DRAM vendors. For the first time, we report circuit topologies, transistor dimensions, and layouts of sense amplifiers on modern commodity DRAM devices. To this end, we combine Scanning Electron Microscopy (SEM) aided with Focused Ion Beam (FIB) to obtain cross-section images from the samples. Performing SEM/FIB requires adequate sample preparation and the identification of the sensing areas. We then perform a highly sensitive image alignment and noise cancellation on the cross-section images, allowing us to obtain a planar view of the sense amplifier region. Using the planar view images, we reverse engineer the circuit topology by performing a multi-dimension inter- and intra-layer mapping.

**HiFi-DRAM.** We seek to enable high-fidelity DRAM research using our reverse engineered data. We start by comparing our findings to existing DRAM models, discovering that they employ transistors with dimensions up to 9x different than our samples counterpart. Then, we analyze existing

studies that propose to modify the sense amplifier region. We find three major inaccuracies. First, these studies consider inaccurate sense amplifier designs, hence their modifications do not always apply to more modern devices. Second, the addition of new elements assumes free space in existing chips which we find not to be the case in our samples. Third, these proposals do not consider the physical layout of sense amplifiers, underestimating the impact of the proposed changes. Considering these inaccuracies, we find that existing studies can have up to 175x errors in their original estimations when considering modern commodity DRAM. Based on these findings, we formulate a set of recommendations to improve the fidelity of future DRAM research. As an example, future studies must avoid focusing on a single sense amplifier in isolation, since we find that multiple sense amplifiers are interconnected in modern DRAM chips.

**Contributions.** The following summarizes our contributions:

1. Using high-resolution IC imaging, we reverse engineer the sense amplifiers of modern commodity DRAM devices from the three major vendors on both DDR4 and DDR5 devices.
2. We report on important properties of sense amplifiers in our samples such as their circuitry, transistor dimensions and physical layouts.
3. We evaluate 13 papers that aim to modify sense amplifiers using our reverse engineered information to identify inaccuracies and formulate recommendations for high-fidelity DRAM research in the future.

**Open sourcing.** The extracted information including IC images, reversed engineered circuits, transistor dimensions and physical layouts can be reached via <https://comsec.ethz.ch/hifi-dram>.

## II. BACKGROUND

We introduce the DRAM architecture and sense amplifier topologies (II-A) before summarizing research that aims at modifying DRAM sense amplifiers (II-B).

### A. DRAM and Sense Amplifier Topologies

Commodity DRAM is available as chips complying with the DDR protocol, standardized by JEDEC [37], [38]. DRAM used in servers and desktops is usually assembled as multiple identical DRAM chips on dual-inline memory modules (DIMMs) [35], [36].

**Physical organization.** Internally, a DRAM chip has a hierarchical structure made of banks (Fig. 1), each made of multiple MATs that generally contain between half to a million capacitors [25], [51], [68]. Each capacitor in a MAT stores one bit of memory and is identified by the combination of a column, row, and bank address supplied by the memory controller. The MATs are surrounded by row drivers in one direction, and by sense amplifiers (SAs) in the other (Fig. 1). When the memory controller accesses memory, it first activates a specific row of a MAT. With a row *activation*, the capacitors in a MAT are connected via bitlines to the SAs of both sides. The connections are interleaved, resulting in half connections towards each side. Then, to perform reads or writes, the memory

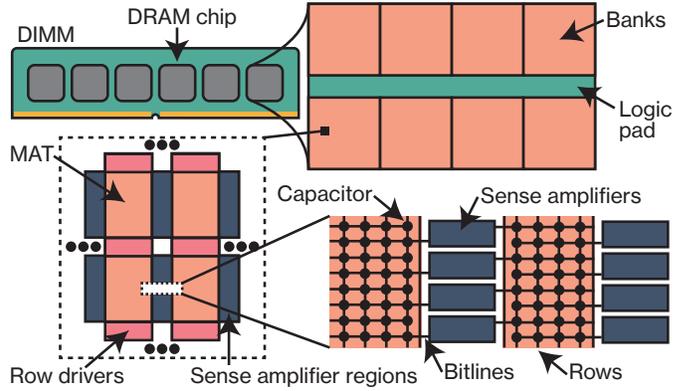


Fig. 1. A DIMM contains several DRAM chips, each made of multiple banks. A bank has many MATs, filled with capacitors. Capacitors are connected to SAs via bitlines after they are selected by rows. SAs are in between MATs.

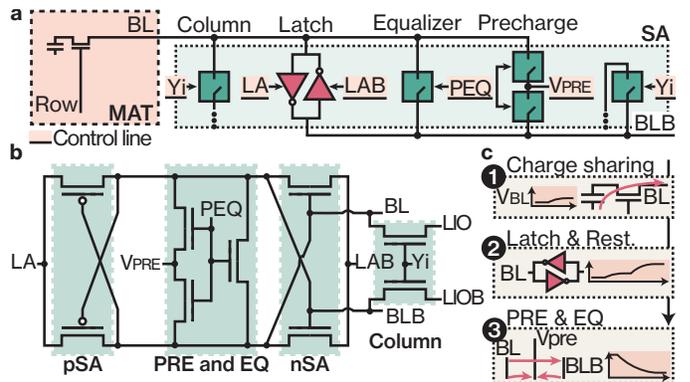


Fig. 2. Latch, equalizer, precharge and column are the main sense amplifier elements (a). In the classic circuit (b), PEQ activates both precharge and equalization. After a row activation, the classic circuit events (c) are charge sharing (1), and latching & restore (2). A precharge makes bitlines connect to  $V_{pre}$  (PRE) and to each others (EQ) (3).

controller specifies a subset of these bits (or capacitors) with the column address. Before activating another row, the memory controller must issue a *precharge* to deactivate the current row.

**Sense amplifiers.** SAs are analog circuits that amplify the weak signals stored in the capacitors. Their design determines the speed of a DRAM chip and must avoid data failures. It is hence a contemporary research topic [8], [32], [41], [44]–[46], [53], [55], [56], [59], [60], [69], [71], [76], [90], [92], [96], [110], [116]. A SA operates by comparing two bitlines, one that is perturbed by the capacitor of the activated row (BL), and one that is the reference (bitline bar or BLB). The reference bitline comes from the MAT opposite to the activated MAT (Fig. 1). This is referred to as an open bitline scheme, currently known to be the most compact scheme and considered the standard [95]. The prime elements of SAs are *latch* circuits (Fig. 2a). Once activated by control lines LA and LAB, the latch circuits amplify and lock the difference between BL and BLB. Then, a *column selector* multiplexes the latched data from a particular SA (selected by control line Yi). Finally, *equalizer* and *precharge* circuits restore the SA

reference voltage ( $V_{pre}$  with the control line PEQ), which is necessary for accessing data from a different row.

**Classic sense amplifier topology and events.** The classic SA is shown in Fig. 2b. The latch element is made of cross-coupled transistors (two pSA and two nSA). Two transistors are used for the precharge, each connecting a different bitline to  $V_{pre}$ . For equalization, one transistor connects BL and BLB. Lastly, the column signal multiplexes both BL and BLB. The classic SA works as follows. After the row activation, each capacitor on the row shares its charge to a specific bitline (BL) perturbing its voltage (*charge sharing*, Fig. 2c). Subsequently, the SA latching elements are activated. This amplification also restores the charge in the capacitor. After the voltage is latched, the memory controller can perform read/write operations. Finally, the memory controller can close the row, and internally, each BL and BLB pair is connected together (*equalized*) and set back to the reference voltage  $V_{pre}$  (*precharged*, Fig. 2c).

Contemporary research systematically assumes that the topology deployed on modern devices is the classic SA [5], [13], [24], [29], [43], [63], [66], [68], [75], [81]–[83], [87], [88], [97], [112].

**New sense amplifiers.** Previous work attempts to enhance DRAM performance with a multitude of new SA topologies [8], [32], [41], [44]–[46], [53], [55], [56], [59], [60], [69], [71], [76], [90], [92], [96], [110], [116]. These proposals change the classic design by adding elements and modifying events. Meanwhile, DRAM designers aim at packing as many rows as possible per MAT, thus increasing the die efficiency. However, having many rows in a MAT reduces the signal strength latched by the SA, increasing the risk of failure (i.e., latching the opposite value) which is exacerbated by smaller technologies. This latching reliability is the result of manufacturing asymmetries in the transistors and bitlines which create an *offset* between BL and BLB. Thus, many of the new SA topologies aim at compensating for these asymmetries, and a subclass of these directly tries to reduce the offset and is known as offset-compensating (or offset-cancellation) SA [32], [41], [44], [45], [59], [60], [69], [76], [96], [110], [116]. In such topologies, the SAs perform additional operations to compensate for these asymmetries.

### B. Research on DRAM

Research focusing on commodity DRAM frequently proposes performance enhancements to the SA region. For example, optimizing the precharge event [81], [87] or speeding up the latching mechanism [91] to reduce memory latency. These improvements often require inserting new elements into the SA or MAT area, such as isolation transistors [66]. SAs, by construction, always latch all bits of a given row. In-DRAM Processing-In-Memory (PIM) exploits this parallelism by modifying SAs and MATs, typically relying on dual-contact cells (DCCs) [2], [112]. DCC is a widely-used SA addition, originally described in [88]. It aims to add an extra row in the MAT, in which each capacitor can connect to two different bitlines instead of one, as selected by two wordlines. Generally, one bitline is the standard connection (BL), while the “extra”

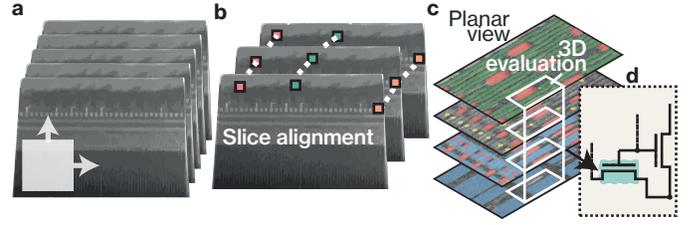


Fig. 3. We filter (a) and align (b) the cross section images. After we obtain a planar view (c), we identify connections between different layers, wires, and transistors. This allows us to reverse engineer the SA circuit (d).

bitline (EBL) is connected to the BLB. Lastly, recent work improves memory integrity by adding elements into the SA area [68].

**Research fidelity.** Multiple aspects critically undermine the accuracy of existing DRAM research that focuses on SAs. First, DRAM MATs and SA regions are highly optimized areas, and they represent the majority of a chip. Therefore, changes in these regions that were intended to be simple, could cause high overheads or require complete re-designs. Second, existing research performs analog simulations either on “best-guess” models or old technologies, and area overheads are based on old values or averages [29], [66], [68], [111]. Lastly, literature assumes that modern DRAM employs the classic SA topology. With HiFi-DRAM, we aim at providing clarity over these aspects.

## III. OVERVIEW AND CHALLENGES

We aim to reverse engineer the SA region in multiple DRAM chips from the three major DRAM vendors. Then, we seek to use the acquired information to extend and improve the accuracy of existing and future research. To these ends, we must overcome challenges that we now describe.

Vendors do not disclose details about the SA region including its location, section dimensions or component feature sizes. Certain physical properties such as the materials as well as the thickness of the Integrated Circuit (IC) layers are further undisclosed. We want to acquire images of this area to then analyze it. Therefore, the first challenge is:

**Challenge (C1).** Acquiring images of the SA region, in a way that the elements of interest and all the layers are visible and identifiable. Then, processing the images to obtain a planar view of the circuit.

We address this challenge in Section IV by employing high-resolution imaging. In particular, we first identify the SA region using a blind approach and proceed to image it by acquiring multiple cross-section slices. Then, we post-process the slices to denoise and align them (Fig. 3a-b), before changing the point-of-view from cross section to top-down (i.e., planar, Fig. 3c).

Once we have obtained a planar view of the different layers of the circuit, we must extract meaningful information. Namely, we must reverse engineer the deployed circuitry and measure

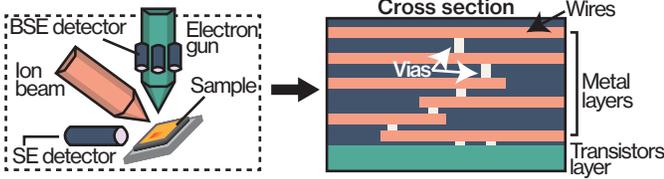


Fig. 4. FIB/SEM imaging requires an ion beam and an electron gun, under which the sample is positioned. The BSE detector is placed on the electron gun, while the SE detector is skewed. FIB/SEM allows to image the cross section of an IC. An IC is made of multiple metal layers, which are interconnected by vias. The transistor layer is placed at the bottom of the IC.

component features, such as transistor widths. This requires analyzing the images on different levels of abstraction. First, identifying different material classes that make up different electrical components and measuring their physical dimensions. Second, mapping the visible components and their connections, which might cross layers and the planar view (Fig. 3c). Third, understanding how these components interact (i.e., the circuit topology, Fig. 3d).

**Challenge (C2).** Starting from the planar view, reverse engineering the circuits considering all the inter-connected layers, and measuring relevant features.

We describe how we address this challenge in Section V. First, we find features that corresponds to gates, wires and vias. After finding the MAT bitlines, we identify different classes of transistors. We then trace their intra- and inter-layer connections and their relation to the MAT bitlines. This way, we associate functionalities to the classes of transistors, which we link to the equivalent circuit block. Lastly, we extensively measure dimensions, including transistor sizes and region areas.

Using this newly acquired data, we aim to understand the accuracy of existing DRAM studies. To accomplish this, we review literature to identify common assumptions that are in conflict with our observations. Further, we must understand if the original estimations can accurately represent overheads on modern devices.

**Challenge (C3).** Evaluating the accuracy of existing research.

We address this challenge in Section VI, where we evaluate 13 different papers. We find that 8 of them result in more than 20x error in the calculation of the overheads and in the extreme case up to 175x when considering the architecture of modern commodity DRAM. We also study the existing available DRAM models, which we found to deviate substantially from the real chips. To enable high-fidelity DRAM research in the future, we formulate a number of recommendations based on our data and on the inaccuracies that we have observed of existing studies.

#### IV. IMAGE ACQUISITION AND POST PROCESSING

Due to the small feature sizes of modern ICs, optical microscopy is not viable for contemporary chip imaging. On

TABLE I  
STUDIED CHIPS. WE STUDY A TOTAL OF SIX CHIPS, FROM THE THREE MAJOR DRAM VENDORS (ANONYMIZED AS A, B AND C). WE REPORT THE CHIP PRODUCTION YEAR, ITS DIMENSION, AND SEM INFORMATION.

ID	Vendor	Storage	Yr.	Size	Det.	MATs	Pixel.Res.
A4	A (DDR4)	8Gb	'17	34 mm <sup>2</sup>	SE	V.	10.4 nm
B4	B (DDR4)	4Gb	'22	48 mm <sup>2</sup>	BSE	N.V.	3.4 nm
C4	C (DDR4)	8Gb	'18	42 mm <sup>2</sup>	BSE	V.	5 nm
A5	A (DDR5)	16Gb	'21	75 mm <sup>2</sup>	SE	N.V.	5.2 nm
B5	B (DDR5)	16Gb	'22	68 mm <sup>2</sup>	BSE	N.V.	4.2 nm
C5	C (DDR5)	16Gb	'22	66 mm <sup>2</sup>	BSE	V.	5 nm

the other hand, Scanning Electron Microscopy (SEM) is an imaging technology that allows resolutions below the optical limit [40]. SEM is based on a system emitting an electron beam on the target sample (Fig. 4). The sample will in turn emit Secondary Electrons (SE) and BackScatter Electrons (BSE), with intensities that depend on its chemical composition.

**SEM parameters.** Many parameters influence the quality of a SEM image [118]. For example, the dwell time represents the time that each spot will receive the beam [107]. A higher dwell time will produce an image with a higher signal-to-noise ratio, but it will require more time which increases the imaging cost since SEM devices are often shared across many different projects. Furthermore, the dwell time is limited by the employed technology and sample stability. The electron beam should be focused and with a high current, while the voltage that accelerates the beam affects brightness. Ultimately, optimal parameters depend on the required resolution, the chip area to image, and the sample under test.

**Detectors.** SEM images are based on either BSE or SE detectors, which have different contrast characteristics. Generally, BSE will enhance the difference in atomic number between the elements of a sample, while SE depends on the conductivity. Depending on the analyzed sample, the image quality might be better with either BSE or SE.

**FIB.** ICs are manufactured as various interconnected layers (Fig. 4), as such, the features of interest are buried inside the chips. Thus, imaging an IC with only SEM would result in merely viewing the upper-most external layer. Focused Ion Beam (FIB) allows milling the sample of interest. By removing material with FIB, the region of interest is exposed and can be imaged via SEM. FIB is usually implemented as GaFIB, where Gallium ion beams are used. Commercially, FIB/SEM are commonly integrated in single machines.

##### A. Sample Preparation

For each of the three major DRAM vendors, we analyze a DDR4 and a DDR5 chip, for a total of 6 chips. We extracted the chips from commodity devices sold as Dual Inline Memory Modules (DIMMs). We purchased the DIMMs from online suppliers and, for each chip, we identified the DRAM vendor using the ID reported on the packaging. The list of chips and production years can be found in Table I (anonymized vendors).

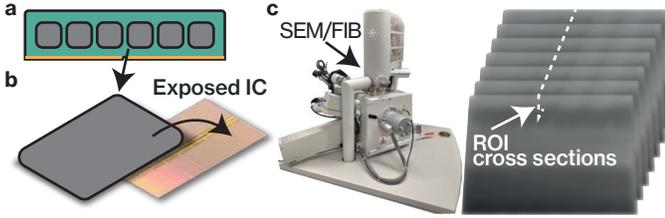


Fig. 5. We extract (a) and decap (b) target chips. Then, we acquire cross sections of the ROI using SEM/FIB (c).

**Die extraction.** We first aim to expose the chip die, which our imaging targets. We desolder the chip from the DIMM by applying a heatgun (400°C). We further use the heatgun to partially remove the epoxy package covering the die. Lastly, we remove the remaining epoxy with a sulfuric acid solution at 140°C (Fig. 5).

**ROI identification.** Given the die dimensions (up to 75 mm<sup>2</sup>, Table I) and the expected features size (tens of nm), imaging the entire chip is not realistic for time, cost and required processing. Hence, we must establish a region of interest (ROI, i.e., the SA region). On the exposed die, we first identify banks and logic pad using an optical microscope (AX10 Imager.M2, ZEISS [10]). In some cases, the die extraction exposed lower layers (Table I). In these chips, we identify the ROI as the largest area surrounding a MAT, as typically row drivers are smaller than SA (Section II) [68].

For the remaining chips, the procedure is more challenging, as optical and electron microscopy only reveal the top layer. Its coarse features solely provide the bank-level organization, leaving the MAT locations unclear. We rely on three properties of DRAM chips to identify the ROI. First, the bank areas are dominated by MATs. Second, the feature lines are either perpendicular or parallel to MATs edges. Third, the area occupied by capacitors visually differs from the analog logic [34], [73]. On this basis, using FIB, we acquire blind cross sections in a bank perpendicularly to the feature lines (Fig. 6). A single image corresponds to less than one millionth of the chip area. We observe the result, and continue acquiring cross sections in the same direction, until we reach a morphological variation in the acquired images. In particular, until we reach an area in which we can identify transistors (Fig. 6). Then, based on the properties, we classify the non-logic area as a MAT. We then perform a perpendicular scan, to obtain the other edge of the MAT (Fig. 6). We identify the ROI as the biggest logic region surrounding the MAT. The identification procedure lasts no more than 2 hours per chip.

### B. ROI Imaging

Once the ROI is identified, we must capture a region large enough to contain complete SAs. We configure SEM/FIB to acquire images of an area of 100  $\mu\text{m}^2$  between two adjacent MATs. We hypothesize, based on existing DRAM models [29], [68], that this is enough to capture SAs. We perform this scan on both DDR4 and DDR5 devices (A4-5), confirming our theory. Each acquisition took more than 24 hours of SEM/FIB

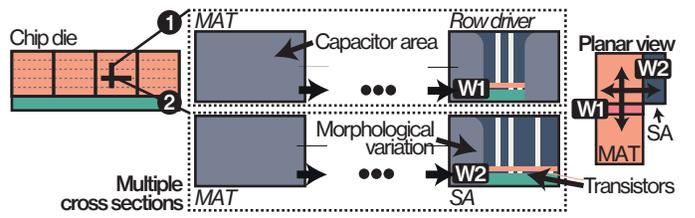


Fig. 6. Starting from one direction (1), we identify a logic region with width  $W_1$  surrounding a MAT. The opposite direction (2) results in a logic region with a width  $W_2$ , bigger than  $W_1$ . We identify this second region as the SAs.

and resulted in imaging many SAs. To reduce the cost of imaging the remaining samples, we perform their acquisitions scanning areas of 30  $\mu\text{m}^2$ , enough to capture multiple SAs.

**Details of SEM/FIB cross sections.** We perform volumetric reconstruction using the Helios 5 UX (Thermo Fisher Scientific [105]) as follows. We use FIB to repeatedly slice the ROI, by removing perpendicular slices of 20 nm or 10 nm (30 kV Gallium ion beam with 90 pA beam current). For each slice removed, we image the cross section with SEM. The pixel resolution of the SEM images varies across experiments, down to 3.37 nm (Table I). We first acquire the SEM images with SE for A4-5, which provides good quality. For the remaining chips, SE does not provide a good contrast, likely due to manufacturing processes, so we use BSE instead. The space parameter for SEM is large, so it may be possible for SE to produce images with higher quality for vendor B and C under different SEM parameters, such as a much longer dwell time (increasing the cost). For the acquisition, we use an accelerating voltage of 2 kV, and dwell times of 3  $\mu\text{s}$  (A4-5, B4) and 6  $\mu\text{s}$  (B5, C4-5).

### C. Image Post Processing

For each chip, we obtain images representing slices of the SA region. To reverse engineer the circuit, however, we require a planar view. That is, we must align the slices together and change the point of view. This step requires post-processing to address high levels of noise and drift.

**Noise sensitivity.** We measure wire heights in the SA region to be as small as 30 nm (cross-sectional view, B5), while the cross section height is generally 130x this value. This makes the planar view extremely sensitive to slice alignment. We hence need to reduce the slice alignment noise and drift to less than 0.77% (1/130) of the slice. Further, this alignment error must apply across all the acquired slices for consistent planar views.

**Reliable post-processing.** We use the Dragonfly software [16] to perform multiple post processing steps. First, we crop the slices to include only the cross section. Then, we filter the images to reduce noise with edge preserving algorithms (split-Bregman [27] or Chambolle [11] for a total-variation denoising). Once denoised, we align the slices using the mutual-information algorithm of Dragonfly. In particular, each slide is aligned with respect to the previous one. Lastly, we change the point of view from cross section to top-down, and we further rotate the

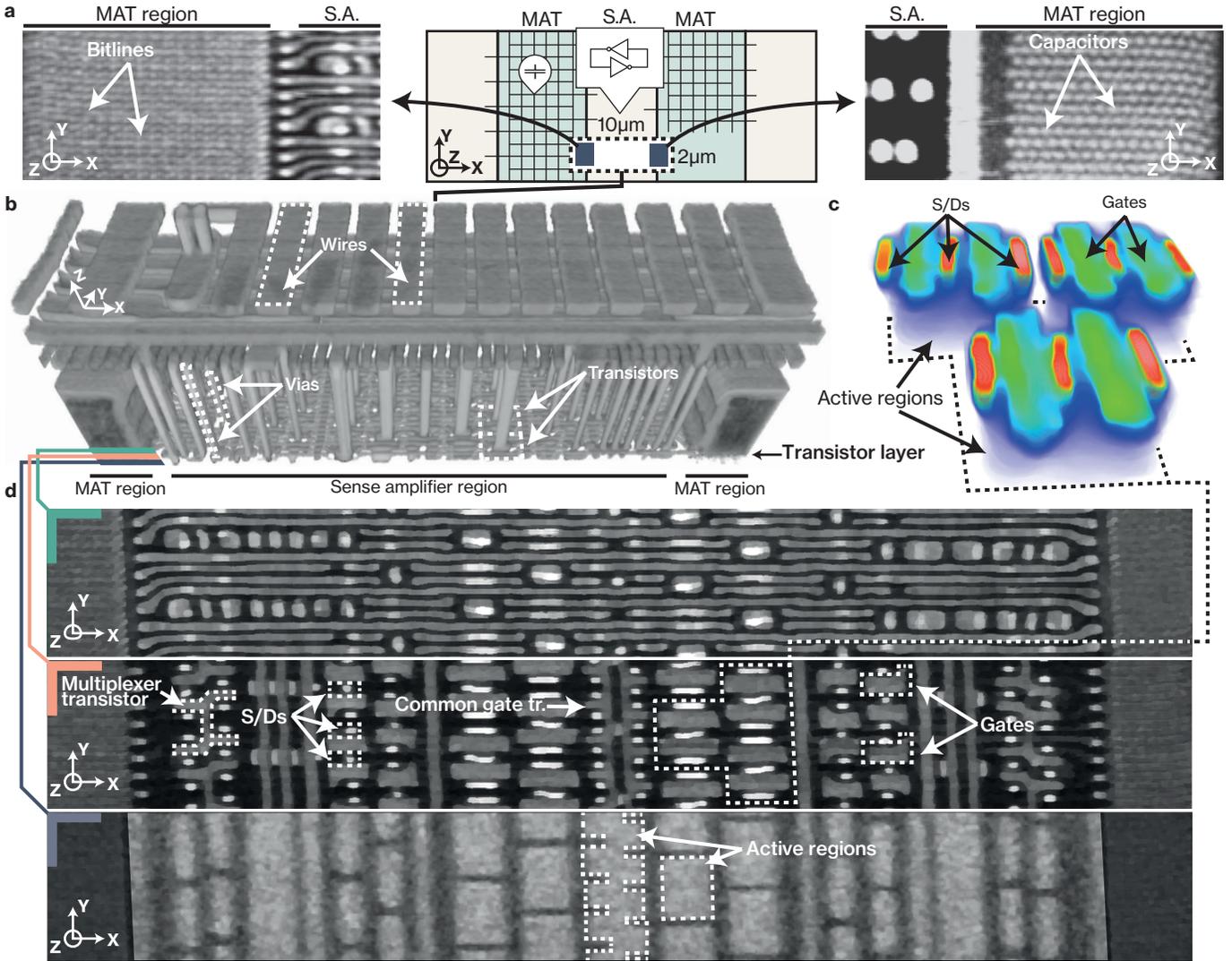


Fig. 7. From the imaged ROI of **C5** (a-center), we can identify bitlines (a-left) and capacitors (a-right). They correspond to two different layers, omitted for simplicity. In the 3D reconstruction of the SA region (b) wires, vias and transistors are visible. An enhanced image (c) shows that different transistors share the same source/drain and active region. From selected planar slices of the 3D reconstruction (d), we can see major elements of the SA circuitry: (1) connections between the different SA elements (i.e., bitlines); (2) gates and source/drains of the transistors; (3) the transistor active regions (enhanced contrast).

volume to correct for possible remaining misalignments. This process is semi-automatic as it requires per-scan tuning, and can be reliably performed in less than 3 hours by an analyst, including the algorithms execution time. Note that the official Thermo Fisher software (Avizo3D [104]), which would include semi-automatic tools for slice alignment and SEM processing, did not produce results matching our requirements.

#### D. Imaging System Capabilities

We focus on the SA region, yet so far, the target component feature size is unclear. As half of the evaluated chips were produced in 2022, it is uncertain if our system can provide a sufficient resolution. In Fig. 7, we demonstrate the reconstruction capabilities of our end-to-end imaging methodology on **C5**. After post-processing, the planar images allow for the identification of transistors, wires and vias in the SA region. For example, Fig. 7a reveals capacitors and bitlines. The capacitors

are arranged in a honeycomb structure and are placed above the bitlines (stacked capacitor [77]). The honeycomb structure has been proposed as a method to increase the capacitance for the same cell size [4], [77]. Fig. 7b shows the 3D reconstruction of the SA region. Fig. 7c depicts two transistors sharing a contact node. Finally, Fig. 7d displays three slices of the SA logic layers, including SA bitlines, transistors and active regions.

### V. CIRCUITS REVERSE ENGINEERING

First, we describe the reverse engineering of the sense amplifier circuit and its challenges (V-A). We then detail the size measurements that we performed (V-B) and analyze the implemented sense amplifier layouts (V-C).

#### A. Reverse Engineering the Circuit Topology

Analog circuits are generally drawn by humans on single-layer schematics, yet this abstraction differs from the final

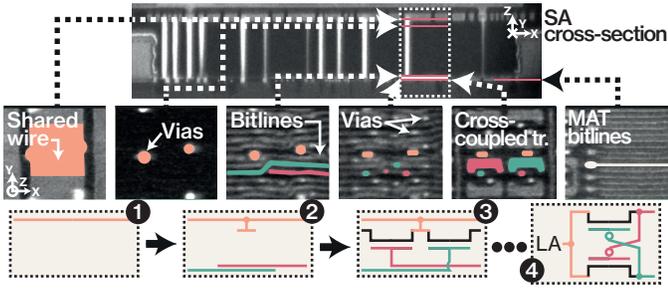


Fig. 8. Cross-coupled transistor reverse engineer on **B5**. From the top slice, we identify a shared line (1-2). Two bitlines are connected with vias to transistors gate and drains (2-3). Once the full circuit is mapped (4), the transistors are identified as the pSA latching element of the SA.

manufactured product, which is implemented as multiple layers stacked vertically (Section IV). Transistors may appear close on the schematic, but they may be deployed far from each other to better utilize the available chip area. Further, some transistor characteristics (e.g., NMOS and PMOS) that impact the circuit topology might be visually indistinguishable, which is the case with our samples as opposed to previous work that targeted different technology [100]. Overall, these factors make reverse engineering analog circuits in DRAM a challenging task, even after a multi-dimensional mapping.

**From images to circuits.** We now describe the methodology that we devised to account for the aforementioned challenges shown in Fig. 7. (i) First, we determine color intensities that correspond to gates, wires and vias (Fig. 7b-d). (ii) Then, we identify the bitlines in the MAT, and their connections in the SA region (Fig. 7a). We use the bitlines as an anchor for inferring the circuit. (iii) We identify the different transistors, the corresponding wires, and the source/drains contacts. To correctly identify transistors, we include active regions in the analysis (Fig. 7d). (iv) We classify three different types of transistor: multiplexer transistors (Fig. 7d), transistors with a common gate spanning the entire region (Fig. 7d), and coupled transistors with a shared source (Fig. 7c). (v) The multiplexer transistors select a group of 4 adjacent bitlines, which are then connected to wires spanning the entire region (not shown). Each of these transistors have a different gate control. Hence, we identify them as column transistors. (vi) We track the bitlines connections to the coupled transistors, which represent a latched connection, and find that the source is shared among all of these transistors. Hence, these represent the latching part of the SA. (vii) In **B4**, **C4** and **C5**, the transistors with a common gate short the bitlines together and with a global value. Therefore, we identify these as precharge/equalizer elements. (viii) Finally, in SAs, PMOS latching transistors are designed with a smaller width than NMOS [45]–[47], [115]. Based on this, we identify the PMOS latching transistors. We considered all other transistors to be of NMOS type. We collaborated with an independent DRAM vendor that confirmed our analysis.

Fig. 8 displays an example of multi-dimensional mapping. We trace inter- and intra-layer connections, and identify cross-coupling transistors. Only after the entire circuit is mapped,

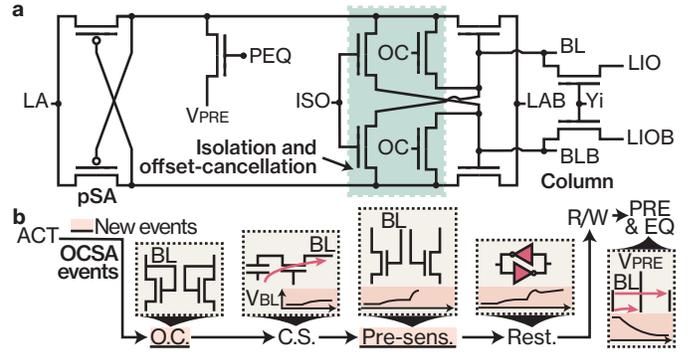


Fig. 9. (a) Offset-cancellation SA (OCSA) circuit, used on **A4**, **A5** and **B5**, and its events (b), which include offset-cancellation and the pre-sensing.

we can identify these elements as the pSA of the latching component. Contrary to existing models and literature about commodity DRAM, we found extra elements in **B5** and **A4-5**.

**Investigating the extra elements.** In chips **B5**, **A4**, and **A5**, we found that the precharge element is stand-alone, and four extra “common-gate” transistors are present (Fig. 10). Moreover, we could not identify the bitlines equalizer. We hypothesized that such a topology might belong to the extensive corpus of past research [8], [32], [41], [44]–[46], [53], [55], [56], [59], [60], [69], [71], [76], [90], [92], [96], [110], [116] and identified similarities with research in *offset cancellation*. Of the many different proposed offset cancellation SAs (OCSAs) [32], [41], [44], [45], [59], [60], [69], [76], [96], [110], [116], we could finally pin-point the reverse-engineered circuits to one design [45].

**Deployed OCSAs.** Half of the devices (**B4**, **C4**, **C5**) use the classic SA circuit [42]. Instead, chips **A4**, **A5**, and **B5** implement an OCSA with the circuit shown in Fig. 9a [45]. Our paper is the first to publicly report that OCSA topology is being used in modern commodity DRAM. The deployment of OCSA circuits is likely due to the need for reliable sensing of capacitor charge in smaller technology nodes. Hence, it is very likely that manufacturers that currently use the classic SA circuits will move to OCSA in the future as well.

The OCSA topology differs substantially from the conventional circuit and adds four transistors and two control signals (Fig. 9a). Two of these transistors perform isolation (ISO) and the other two offset cancellation (OC). The OCSA adds two operations to the classic row activation of a SA (Fig. 9b). First, charge sharing is anteceded by an offset cancellation operation. Second, the restoring operation is preceded by a pre-sensing event. The pre-sensing operation latches the capacitor value without the bitline load and without recharging the capacitor.

**Isolation and equalization in OCSAs.** DRAM research often proposes adding isolation transistors to the SAs [66], [68], [87], [99]. Typically, this allows decoupling the bitlines from the latching circuit. The isolation transistors used in OCSAs differ from these proposals, as the bitlines are decoupled from the latch amplifier drains but not from the gates. Furthermore and as previously explained, precharge and equalizer circuits are

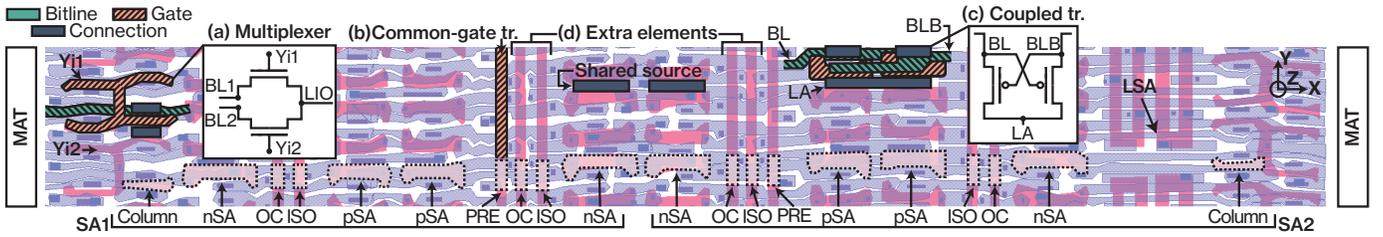


Fig. 10. **Reverse engineered layout of the A5 chip.** To reverse engineer the chips, we identified different transistor classes. Multiplexer (a), common-gate (b) and coupled (c) transistors. In chips **A4-5** and **B5** we discovered extra elements (d). In all the chips, two SA were stacked between two MATs (i.e., along X).

necessary to operate DRAM. Normally, this is achieved by a three-transistor setup (Section II). In the reverse engineered circuits that employ OCSAs, the equalizer transistor is absent. Instead, this functionality is achieved by the simultaneous activation of both the isolation and offset-cancellation elements.

### B. Measurement of the DRAM Elements

DRAM research that performs analog simulations relies on the transistor widths and lengths. The ratios between width and length ( $W/L$ ) of the various elements strongly affect the stability and speed of DRAM operations. Therefore, we measure the width and length of each transistor employed by the SA. To measure their length, we consider the gate pitch between source and drain. To measure their width, we consider the overlapping area between the gate and the active region [6]. We perform multiple distinct measurements for each dimension and for each transistor. In total, we make 835 size measurements using Dragonfly [16]. In Section VI, we provide more details about these measurements and their impact on existing DRAM research.

**Effective sizes.** Adding elements to the SA region must take IC design rules into account. As a proxy for important design rules for DRAM research, we measure the effective spacing dimensions required for each element. That is, we measure the element size including the full gate dimension and the element distance from the other components. These dimensions are higher than the width and length of transistors, as they must include safety margins. We measure the dimensions of the SA and MAT regions, and of each die. We make all our measurements for each of the samples available online.

### C. Layout Design Analysis

Given the absence of information, previous studies that modify SA regions have mostly ignored the physical layout of modern DRAM devices. Unfortunately, this has repercussions on the accuracy and overhead of research as we will show in Section VI. To bridge this gap, we re-created the physical layout of the SA regions in all of our samples, which we make available in the standard GDSII format. As an example, Fig. 10 shows the layout of the **A5** chip. We find that all the samples employ an open-bitline architecture, considered the standard since many years [95]. We now describe the immediate differences that we found when comparing our findings to the assumptions made on the DRAM layout in existing work.

**SA arrangement.** In all studied chips, SA elements are always arranged horizontally (i.e., along the X axis, Fig. 10). All chips have two stacked SAs (side by side) between each MAT (Fig. 10, “SA1” and “SA2”), with transistors positioned symmetrically. Both sets of SAs connect to bitlines from each MAT. This is different from the usual description of SAs [5], [13], [22], [29], [48], [63], [66], [68], [112], where only one SA is placed between two MATs. Consequently, the overhead of elements shared among all bitlines (e.g., isolation transistors) is lower than what previous research has assumed. The column transistors are always the first elements connected by MAT bitlines in the SA region which results in inaccuracies in previous studies as we outline in Section VI. Finally, the SA region further contains latching elements connected to the selected column (i.e., to LIO). They represent the next data-path step and are not part of the SA circuit (Fig. 10a, LSA). However, because they are part of the SA region, their presence reduces the relative overhead of new elements.

**Transition from MAT to SA.** The transition of a bitline from MAT to planar logic requires an overhead on average of 318 nm (DDR4) and 275 nm (DDR5) in the bitline direction. This information is useful for research that proposes to add transistors in between the MAT (i.e., creating a new logic region), and to the best of our knowledge, it has not been reported in literature so far. For example, [58] proposes to place isolation transistors in a MAT to create shorter bitlines. On top of the overhead due to a single isolation transistor, two transitions are required, as the MAT is split in two. On average this represents 1.6% and 1.1% of a MAT in DDR4 and DDR5 respectively.

**Transistor characteristics.** Previous work that adds transistors to the SA region often calculates overheads as an increase in SA height (i.e., along X) related to the transistors width ( $W$ ) [68], [81]. This is correct for the latching components, as their width is parallel to the SA height. Instead, we find that precharge, isolation and offset-cancellation transistors are designed with a common gate spanning the entire SA region (i.e., along Y). As result, the width of these transistors is perpendicular to the width of the other elements (Fig. 10). Therefore, the addition of these elements causes a SA height overhead that depends on their lengths ( $L$ ). Finally, we find that the access transistors used in the MAT region have a bitline/wordline layout typical of Buried Channel Array Transistors (BCAT) across all vendors. This is consistent with literature [77].

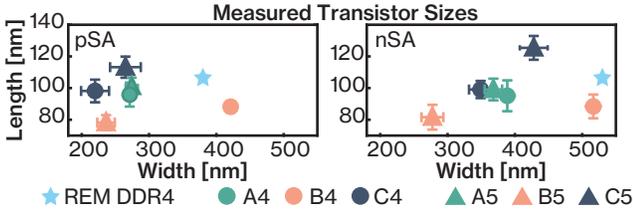


Fig. 11. Measured transistor sizes of the pSA and nSA for all the chips and the values used in REM. CROW values are omitted as severely out the range.

## VI. EVALUATION OF EXISTING DRAM RESEARCH

We first analyze analog DRAM models based on our reverse engineered data (Section VI-A). Then, we evaluate previous studies that modify the sense amplifier region to identify the sources of inaccuracies (Section VI-B), and provide a more accurate measurement of area overhead in these studies (Section VI-C). Finally, we comment on the reliability of physical experiments on DRAM (Section VI-D) and conclude with a set of recommendations for accurate DRAM research in the future (Section VI-E).

### A. Inaccuracies of Existing Analog Models

Analog simulations of DRAM SAs are widely used in research papers [2], [5], [13], [15], [22], [26], [28], [29], [58], [66], [68], [75], [78], [85]–[88], [93], [94], [97], [111]–[113]. However, no DDR5 model exists, and only two public models exist for DDR4. In particular, DDR4 SAs are simulated with CROW (2019) [29] or with REM (2022) [68] models. Neither of these two models are based on commodity DRAM devices from the major manufacturers. CROW is employed in multiple recent work [29], [66], [111], however its transistor dimensions are based on best guesses and it does not include column transistors. REM is based on real DDR4 transistor dimensions of a smaller vendor (Zentel Japan [117]) that uses 25nm technology. This technology, however, is one generation older than current commodity DDR4 device from the three major DRAM vendors [101]. Neither models include the OCSA design. In Fig. 11, we report the dimensions of the latching transistors (nSA and pSA) for all the chips that we reverse engineered and for REM. We omit CROW in this figure as its values are vastly out of range. To understand if existing DRAM models provide an accurate representation of commodity DRAM devices, we analyzed the width-to-length ratio ( $W/L$ ) of their transistors. Generally, higher width-to-length ratios correspond to more optimistic simulations [68]. In particular, we compared each model element to each ratio obtained for that element in each chip. We included a comparison to DDR5 technology to determine whether these models can provide a reasonable approximate description of these chips as well.

**Results.** We report a summary of our analysis in Fig. 12 and discuss the inaccuracies when compared to DDR4 chips. On average, CROW has the higher inaccuracy between the two models (236%). The precharge of CROW has the highest  $W/L$  inaccuracy (562% when compared with the measured values of C4). We further analyzed the individual widths and lengths.

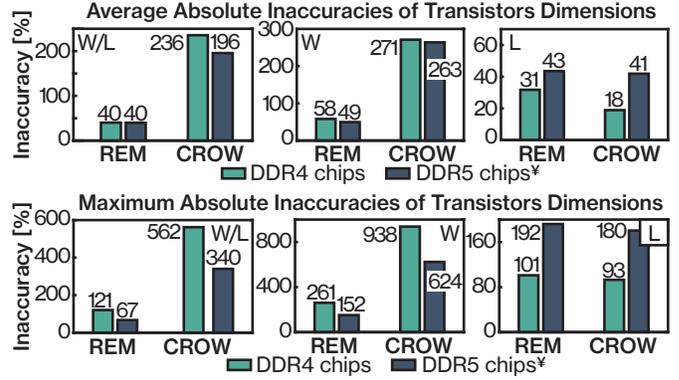


Fig. 12. Average and maximum absolute inaccuracies of REM and CROW compared to the measured transistors in all the chips, as  $W/L$  ratios, and separately width and length. (¥) Portability to DDR5.

CROW gives the most inaccurate widths on average (271%), with an inaccuracy of 938% when compared to the C4’s precharge transistors. REM has the most inaccurate lengths on average (31%), with an inaccuracy of 101% when compared to C4’s equalizer transistors. The models follow a similar trend when considering the DDR5 technology.

### B. Deriving Research Inaccuracies

We now analyze research that proposes to modify the SA regions of commodity DRAM. Correctly estimating overheads and feasibility of proposed modifications to DRAM is a complex task. Even for variations that would appear minor, the absence of public information about modern devices is an obstacle that forces researchers to make blind assumptions and estimate overheads based on outdated ranges. Once a new technology is released, such as DDR5, it is further challenging to understand if existing proposals remain feasible. Unfortunately, this is a de-facto accepted status in the field. With HiFi-DRAM, we provide researchers with a realistic basis upon which they can evaluate their work.

**Papers summary and analysis.** We study 13 papers, crossing technologies (DDR3-DDR4) and spanning a decade (2013-2023). Among these, [66], [81], [83], [87], [94], [99] seek to improve performance by modifying DRAM. [68] modifies DRAM to improve data integrity. The remaining studies aim to implement in-DRAM PIM [1], [2], [21], [28], [88], [112]. Studying these papers, we enumerate the sources of research inaccuracy when compared to commodity devices. As research on DRAM has been based on similar repeated assumptions throughout the years, we identify common inaccuracies across most papers under study, which we describe as **I1** to **I5**.

A major inaccuracy arises from the implementation of dual-contact cells (DCCs), discussed in Section II-B. Their overhead is estimated to be approximately two wordlines, i.e., negligible [88]. In all the chips that we studied, MATs do not have available space for the extra bitlines (**I1**, Fig. 13a). Because of this, implementing a DCC requires doubling the MAT area. As MATs represent the majority of a DRAM chip, implementing even a single DCC results in a significant

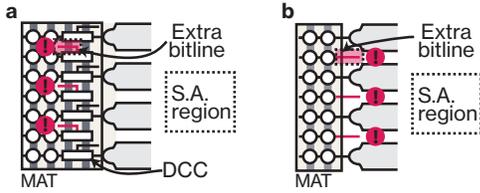


Fig. 13. (a) No free space to add new bitlines in the MAT (I1) and (b) SA region (I2).

overhead. Further, making the MATs larger will increase the wordline length. Due to this increase, extending the MATs will also require placing new row drivers to correctly drive the new load. Row drivers areas are comparable to the sense amplifier areas [68]. As measured from our data, all the papers affected by I1 require on average 57% chip overhead, solely for the MAT extension.

**Array size and implications related to I1.** DRAM vendors are constantly trying to reduce the size of their memory arrays to increase chip yield. The current standard is the open bitline design, which has an area consumption per cell of  $6F^2$  ( $F$  represents the feature size). If MAT bitlines could be made closer, this would effectively reduce the cell size to under  $6F^2$ . Prior work describes the same type of cell structure as a DCC, resulting in an area of  $12F^2$  when reverting to a folded-bitline architecture [33], confirming the aforementioned overhead.

**Inaccuracy (I1).** No free space for bitlines in the MAT area.

The main DCC application is an inverter PIM operation exploiting row parallelism, so the EBL is connected to the reference bitline (BLB). There is, however, no extra space for bitlines crossing the SA region (I2, Fig. 13b). The same inaccuracy arises in [66], which aims to connect all bitlines in the MAT to the same SA region. In [68], additional wires are required for routing purposes of the new circuitry. All these papers do not consider extra overhead due to the new required wiring.

**Inaccuracy (I2).** No free space for bitlines in the SA area.

In standard IC processes (i.e., non-DRAM), designers could try to resolve I1-2 by exploiting the many available IC layers. However, as confirmed by our observations and literature, this is not possible in the DRAM SA regions and MATs, where the number of IC layers is limited [49], [87], [98]. One possibility is shrinking the bitlines. However, MAT bitlines are the main contributor of timing and signal level. Changing their dimensions would severely affect the functioning of the SA.

The feasibility of changing the SA bitlines depends on the fabrication process. These changes would affect resistance and parasitics, and must respect the design rules, such as the minimum distance between bitlines. These considerations are not addressed in the aforementioned studies. We refer the reader to Appendix A for further explanations.

TABLE II  
RESEARCH INACCURACIES, AVERAGE OVERHEAD ERROR AND PORTABILITY COST. THE OVERHEAD ERROR IS EVALUATED ON THE ORIGINAL TECHNOLOGY IF POSSIBLE. PORTABILITY COST REPRESENTS THE OVERHEAD VARIATION OF DDR3 TO DDR4/5 AND DDR4 TO DDR5.

Research	Inacc.	Error	Port. Cost	DDR	Yr.
CHARM [94]	I5	N/A	0.29x	3	'13
R.B. DEC. [87]	I4,5	N/A	-0.25x	3	'14
AMBIT [88]	I1,2,5	N/A	68x	3	'17
DrACC [21]	I1,2,5	35x	34x	4	'18
Graphide [2]	I1,2,5	54x	52x	4	'19
In-Mem.Lowcost. [1]	I1,2,5	70x	67x	4	'19
ELP2IM [112]	I2,3,5	N/A	90x	3	'20
CLR-DRAM [66]	I2,5	22x	21x	4	'20
SIMDRAM [28]	I1,2,5	70x	67x	4	'21
Nov. DRAM [99]	I4,5	0.49x	0.001x	4	'21
PF-DRAM [81]	I5	0.35x	-0.01x	4	'21
REGA [68]	I2,4,5	8x	7x	4	'23
CoolDRAM [83]	I1,2,3,5	175x	168x	4	'23

Some studies rely on SA circuitries that differ from the ones that are currently deployed. This affects [112], in which the authors consider the precharge and equalization gates to be independent for each SA. In reality, the gates of these elements are spanning the entire SA region and are shared across all the SAs. In [83], the authors assume that isolation transistors are available in the design. As described earlier (Section V), these isolation transistors differ from the one employed in OCSAs. These inaccuracies could be resolved by adding new elements to the SA region or by duplicating existing ones, introducing additional area overhead.

**Inaccuracy (I3).** Assuming a SA circuitry that is not deployed in practice.

Additionally, [68], [87], [99] modify SAs by adding isolation transistors and assuming that column transistors are physically placed after the SAs. In reality, column transistors are the first elements after the MATs. Therefore, these modifications require a reorganization of the SA elements.

**Inaccuracy (I4).** Assuming a SA physical layout that does not correspond to the ones deployed.

Finally, no paper includes the OCSA topology in their studies, in contrast to chips A4-5, B5. This affects the overheads of additional components and the timings of the new events as well as the reliability of analog simulations, impacting the performance, energy and power overheads of the affected operations.

**Inaccuracy (I5).** Not considering offset-cancellation designs as the deployed SA topologies.

In Table II, we provide a summary of our findings, and evaluate the overhead inaccuracies which we detail next.

### C. Evaluation of Research Inaccuracies

The area overhead of DRAM research is a main factor influencing its feasibility. However, previous estimations have been performed by referencing outdated values or average ranges [1], [21], [64], [66]–[68], [81], [83], [87], [112]. It is unclear if the reported results are realistic when considering commodity DRAM and how they would change if applied to a newer technology such as DDR5. We now study these aspects for the papers under analysis. When the paper is evaluated on its original technology, we describe the variation of overhead as overhead error. In case the original technology is older than DDR4, the analysis is not applicable (N/A). Instead, when the paper is compared to a different technology, we describe the variation in overhead as porting cost. To this end, we use the transistors effective sizes, the region areas measured in Section V, and we include the effects of the inaccuracies discussed in Section VI-B. For our calculations, we follow the description of the original document as closely as possible, and calculate the overheads for each chip. For papers requiring isolation transistors that gave no indication about their sizing, we used dimensions from the existing isolation transistors if any isolation transistor is present in the chip, else we scaled their average dimensions to the chip values.

**Effects of I1-2.** Inaccuracies I1-2 result in an extension of MAT and SA regions. For example, if for every existing bitline a new bitline must be added, this effectively means doubling the width of the region. We contacted the authors of the papers affected by I1 or I2, as the quantitative effect of these inaccuracies is severe. While many authors replied ([1], [28], [66], [68], [83], [88], [112]), none provided clarifying details that would resolve the inaccuracies given the content of the original papers.

However, the authors of [112] suggested a feasible approach to implementing the NOT operation (not evaluated in the original paper). The authors of [28], [66], [88] explained that a detailed implementation was outside the scope of their paper, in line with prior work. They also suggested that if adding a new metal layer were possible, alternative implementations (not evaluated in the original paper) could reduce the overhead. The authors of [68] reported that their collaborating (smaller) DRAM vendor did not report I2 to be an issue on their technology and are exempted from I2 in chips A4-5 as discussed in Appendix A. Finally, the authors of [83] explained that their evaluations were based on the original paper on the topic ([88]) and limited by not having access to proprietary details of DRAM circuitry. We believe these communications show that HiFi-DRAM is highlighting inaccuracies in previous work, and will provide value to researchers focusing on DRAM.

**Results.** The average overhead errors and porting costs are reported in Table II. Papers affected by I1 or I2 have consistently large errors and porting costs (e.g., up to 175x) across all vendors. Such large errors occur due to the (often) very small overheads reported by the papers (e.g., 0.4% [83]). In Fig. 14, we report the inaccuracies and porting costs individually per DRAM vendor. We omit proposals where these are always higher than 10x. The formulas used to calculate

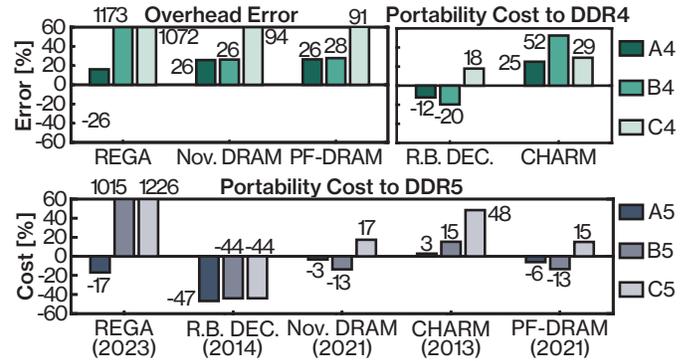


Fig. 14. Research portability cost and overhead error divided per DRAM vendor. Papers where the cost/error is always higher than 10x are omitted.

the errors are reported in Appendix B.

**Observation 1.** The overheads of papers can vary significantly across vendors. For example, [94] has a variation of 0.45x when passing from Vendor A to Vendor C on DDR5 chips.

**Observation 2.** Porting modifications that are originally intended for older DRAM devices results in much lower overhead in DDR5 due to smaller technology nodes. The biggest variation is for [87] (-0.47x on A5). This analysis shows that, in newer technologies, researchers can generally afford more complex circuits.

### D. Out-of-spec DRAM Experiments and OCSA

Issuing commands to DRAM without complying with the DDR standard is used for reverse engineering [72], transistor speed evaluation [68], and DRAM characterization [12], [54], [57]. Furthermore, operating DRAM out of specification is used to exploit the interactions of SAs with rows to perform logic operations [24]. To these ends, researchers expect commodity DRAM to deploy classic SAs. However, chips employing OCSAs have key differences in timings and functionalities that could impact similar experiments. First, charge sharing is usually assumed to occur immediately upon a row activation with the classical SA design [24]. Instead, in chips with OCSAs, charge sharing is delayed and happens after the offset cancellation. This could impact, for example, studies that intend to perform majority-based row operations [24], where multiple rows perform charge sharing without starting the latch operation. Second, bitlines have only two states in the classic circuit, either being latched or precharged and equalized. Instead, OCSAs briefly connect bitlines to diode-connected transistors as a way to improve the sensing margin (Section V). This could impact studies that skip the precharge command to avoid any effect on bitlines [24].

### E. Recommendations

Based on our findings, we now formulate a number of recommendations for future DRAM research. First, simple changes might result in non-negligible overheads when applied to commodity devices (I1-2).

**Recommendation (R1).** Overheads should be estimated including all additions to MATs or SAs, such as wires connections.

Second, research usually focuses on a single SA element, which can lead to wrong assumptions such as considering SA control lines being independent of other SAs (**I3**).

**Recommendation (R2).** Research modifying SAs should consider the impact on all the interconnected SAs.

Third, differences between the abstract SA circuit schematic and its physical layout can create inaccuracies, for example when adding isolation elements (**I4**).

**Recommendation (R3).** Research should consider the physical layout and organization of SAs blocks.

Finally, the deployed SA topology impacts analog simulations, overheads and timings of research proposals (**I5**). Moreover, it can impact DRAM experiments that operate the devices outside of the standard.

**Recommendation (R4).** Research should consider OCSA in the evaluation.

**On existing and future work.** Our results discussed in Section VI-C show that some previous work incur high overhead when considering current commodity devices. We would like to clarify that our evaluation does not reduce the value of these proposals, some of which have led to a high sprout of subsequent work. HiFi-DRAM’s aim is to increase the fidelity of DRAM research, and we sincerely hope that it is not indiscriminately used to stop novel future work due to potentially higher (but more accurate) reported overheads.

## VII. RELATED WORK

### A. DRAM Reverse Engineering

To the best of our knowledge, HiFi-DRAM is the first public research that reverse engineers the sense amplifier topologies, transistor sizes and physical layout of DRAM. In [30], the authors directly issue DRAM commands to devices with an FPGA. Their aim is to reverse engineer internal digital control mechanisms that protect against memory corruption. This results in estimating the existence of row activation counters and their sizes and is unrelated to sense amplifiers. Recent work [72] tries to obtain the number of rows in MATs by exploiting data corruption.

Techinsights [102] is a company that sells access to reverse engineered chips, including DRAM. Unfortunately, the price for accessing DRAM information equivalent to this paper is prohibitive for academic researchers (in the order of \$100 ks). Moreover, the corresponding license [103] prohibits sharing the data publicly and may prevent the resulting publication. This makes it impossible to reproduce and verify research based on such data.

### B. IC Imaging

IC imaging is commonly performed by device manufacturers for identifying manufacturing failures in the produced chips [7], [50], [62], [84] or possible malicious modifications [61], [65], [100]. It is further used by private organizations to verify IP infringement [9], [23], [106], and to perform offensive research [19], [20], [80], [108] such as breaking protection systems [40] or extracting private keys [52].

Previous work reports imaging of proprietary ICs on different targets, from EEPROMs [20], baseband chips [106], lockout chips [40], microcontrollers [52], processors [31], to system-on-chips [17]. A substantial orthogonal research direction is based on trying to automatize standard cell recognition and netlist extraction [3], [18], [50], [79], [109].

## VIII. CONCLUSIONS

We reverse engineered the sense amplifier region in DDR4 and DDR5 devices from the major DRAM vendors. We discovered that half of the chips employ an offset-cancellation sense amplifier, instead of the commonly assumed classical design. We measured transistor dimensions and further reverse engineered the physical layouts of the sense amplifiers. With this acquired knowledge, we validated the overhead and accuracy of existing research in the past decade that modify DRAM sense amplifiers. Considering commodity modern DRAM, our analysis shows that the public DRAM models are up to 9x inaccurate, and existing research has up to 175x error when estimating the impact of the proposed changes. We hope this paper and the reverse engineered information to enable high-fidelity DRAM research in the future and foster new research directions.

## ACKNOWLEDGEMENTS

We thank our anonymous reviewers and shepherd for their valuable feedback. We also like to thank Onur Mutlu for his valuable feedback on an earlier version of this paper. This work was supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40 180545, the Swiss State Secretariat for Education, Research and Innovation under contract number MB22.00057 (ERC-StG PROMISE), and a Microsoft Swiss JRC grant.

## APPENDIX A EFFECTS OF CHANGING BITLINES

Among many other things, IC design rules describe the minimum width of wires and their safety distance from other elements. Bitlines are the narrowest wires placed on the lowest metal layer (M1) of the SA region, which then extend in the MAT region. We now briefly discuss on the feasibility of shrinking bitlines, which is related to **I1-2**.

**Process feasibility.** From a pure manufacturing point of view, shrinking wires that are already narrow can cause the interruption of their conductivity (i.e., creating disjointed wires), while reducing the distance with adjacent wires can create short circuits [39].

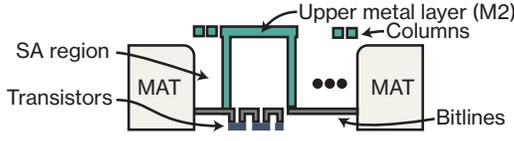


Fig. 15. Connection on M2 of existing bitlines (A4-5). Columns connect to LIO/LIOB (Fig. 2b).

**Electrical impact.** Shrinking wires increase their electrical resistance (R), while making wires closer increases crosstalk [6]. The increase in R will reduce the speed of transmission of bitlines. In DRAM applications, this further translates to the time required to: precharge bitlines, charge sharing, latching the stored data and recharge the capacitor. Crosstalk is modeled as capacitive coupling between parallel wires. Effectively, crosstalk means that a variation in one wire will affect its adjacent wires. This is a particularly well known problem in DRAM applications and can cause read failure [70], [89], [114].

DRAM manufacturing processes use the smallest bitlines width possible to keep the design as compact and reliable as possible. However, even in the case that shrinking bitlines would be possible and not performed in the original layout, doubling the number of bitlines would still result in large overhead. As an example, if halving B5 SA-region bitlines were possible, the result would still add 21% chip area overhead on top of the existing overheads. As the safe distance ( $d$ ) is kept, and the bitline width ( $B_w$ ) is  $B_w \simeq d \times 2$ , the SA extension in the Y direction would be:

$$\begin{aligned} Ext &= \frac{T_B \times 2 \times (d + B_w/2)}{T_B \times (d + B_w)} - 1 \\ &= \frac{2 \times (B_w/2 + B_w/2)}{(B_w/2 + B_w)} - 1 = \frac{4}{3} - 1 \simeq 33\% \end{aligned} \quad (1)$$

where  $T_B$  is the initial number of bitlines in the region. Due to layout requirements, this extension is required on the MAT as well (or alternatively, it introduces empty spaces), resulting in 21% chip overhead.

**Metal layer 2 in A4-5.** In chips A4-5, bitlines that use the second set of SAs are connected via the metal layer 2 (M2). This differs from the other chips where these bitlines are directly connected on M1, and M2 is fully dedicated to other connections. These bitlines are *already* present in the SA region, and the translation is performed after the column transistors (Fig. 15). M2 wires are around 8x bigger than bitlines on M1, are not packed closely, and the layer presents empty spaces. We evaluated that [68] would require reducing these wires by 0.25x to accommodate new connections, and thus we consider this possible. Papers that require adding *new* bitlines to the SA region are not impacted by this because they still require the new bitlines to enter the sense amplifier region as shown in Fig. 13b.

## APPENDIX B OVERHEAD CALCULATIONS

We now briefly discuss the mathematical formulas we used to derive overhead errors. For each paper, we estimate its overhead ( $P_{chip}$ ) for each imaged chip (*chip*). Given the paper original overhead estimation ( $P_{oe}$ ), we report the average of  $(P_{chip}/P_{oe} - 1)$  in Table II. For simplicity, we define  $P_{chip} = P_{extra}/Chiparea$  and now describe  $P_{extra}$ .

Papers that suffer **I1** or **I2** will require a severe extension of the SA or MAT region. Due to layout requirements, extending only the MAT or SA alone will require its counterpart to be extended as well (or alternatively, it introduces equivalent empty spaces). Generally, calculating the effect of **I1** and/or **I2** for a paper that doubles the bitline can be approximated as:

$$P_{extra} = MAT_{area} + SA_{area}$$

REGA [68] requires adding one new bitline every three in chips B4-5 and C4-5:

$$P_{extra} = (MAT_{area} + SA_{area})/3$$

We now report the formulas of  $P_{extra}$  for the remaining papers and for REGA on chips A4-5. In the following,  $MATs$  represents the number of MATs in a chip,  $SA_w$  the SA region width. Instead  $iso_{ls}$ ,  $san_{ws}$ ,  $sap_{ws}$ ,  $col_{ws}$  are the horizontal (i.e., X direction) sizes of isolation, nSA, pSA and column transistors, respectively. Most of the calculations represent a horizontal extension multiplied by the width, replicated for the total number of SA regions in a chip. As all the chips implement two stacked sense amplifiers, original papers that requires adding a new SA, will actually require adding 2 SAs to connect all bitlines. Further, isolation transistors are shared among multiple rows.

REGA [68] requires new isolation transistors and SAs (A4-5):

$$P_{extra} = MATs \times SA_w \times (2 \times iso_{ls} + 8 \times \frac{san_{ws} + sap_{ws}}{6})$$

NR.B. DEC. [87] requires new isolation transistors:

$$P_{extra} = MATs \times SA_w \times 2 \times iso_{ls}$$

Nov. DRAM [99] adds isolation, column and SA transistors:

$$P_{extra} = MATs \times SA_w \times (2 \times iso_{ls} + 2 \times col_{ws} + 8 \times (san_{ws} + sap_{ws}))$$

CHARM [94] changes the aspect ratio of MATs, where 1% is an overhead due to layout reorganization (we use the configuration [x2,/4] from the original paper):

$$P_{extra} = MATs \times SA_w \times SA_h/4 + 0.01 \times Chiparea$$

PF-DRAM [81] adds independent isolation transistors and an SA imbalancer, similarly to an SA:

$$P_{extra} = MATs \times SA_w \times (4 \times iso_{ls} + 8 \times (san_{ws} + sap_{ws}))$$

We invite the reader to refer to the original content of the papers and to <https://comsec.ethz.ch/hifi-dram> for more details.

## REFERENCES

- [1] M. F. Ali, A. Jaiswal, and K. Roy, "In-memory low-cost bit-serial addition using commodity dram technology," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 1, pp. 155–165, 2019.
- [2] S. Angizi and D. Fan, "Graphide: A graph processing accelerator leveraging in-dram-computing," in *Proceedings of the 2019 on Great Lakes Symposium on VLSI*, 2019, pp. 45–50.
- [3] L. Azriel, J. Speith, N. Albartus, R. Ginosar, A. Mendelson, and C. Paar, "A survey of algorithmic methods in ic reverse engineering," *Journal of Cryptographic Engineering*, vol. 11, no. 3, pp. 299–315, 2021.
- [4] N. Bae, S. Thibaut, T. Wada, A. Metz, A. Ko, and P. Biolsi, "Advanced multiple patterning technologies for high density hexagonal hole arrays," in *Advanced Etch Technology and Process Integration for Nanopatterning X*, vol. 11615. SPIE, 2021, pp. 24–31.
- [5] F. Bai, S. Wang, X. Jia, Y. Guo, B. Yu, H. Wang, C. Lai, Q. Ren, and H. Sun, "A low-cost reduced-latency dram architecture with dynamic reconfiguration of row decoder," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 31, no. 1, pp. 128–141, 2022.
- [6] R. J. Baker, *CMOS: circuit design, layout, and simulation*. John Wiley & Sons, 2019.
- [7] A.-C. Bette, P. Brus, G. Balazs, M. Ludwig, and A. Knoll, "Automated defect inspection in reverse engineering of integrated circuits," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 1596–1605.
- [8] T. N. Blalock and R. Jaeger, "A subnanosecond clamped-bit-line sense amplifier for 1t dynamic rams," in *1991 International Symposium on VLSI Technology, Systems, and Applications*. IEEE Computer Society, 1991, pp. 82–83.
- [9] U. J. Botero, R. Wilson, H. Lu, M. T. Rahman, M. A. Mallaiyan, F. Ganji, N. Asadizanjani, M. M. Tehranipoor, D. L. Woodard, and D. Forte, "Hardware trust and assurance through reverse engineering: A tutorial and outlook from image analysis and machine learning perspectives," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 17, no. 4, pp. 1–53, 2021.
- [10] Carl Zeiss AB, "Carl zeiss ab," <https://www.zeiss.com/>, 2023.
- [11] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, pp. 89–97, 2004.
- [12] K. K. Chang, A. Kashyap, H. Hassan, S. Ghose, K. Hsieh, D. Lee, T. Li, G. Pekhimenko, S. Khan, and O. Mutlu, "Understanding latency variation in modern dram chips: Experimental characterization, analysis, and optimization," in *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*, 2016, pp. 323–336.
- [13] K. K. Chang, P. J. Nair, D. Lee, S. Ghose, M. K. Qureshi, and O. Mutlu, "Low-cost inter-linked subarrays (lisa): Enabling fast inter-subarray data movement in dram," in *2016 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2016, pp. 568–580.
- [14] K. K. Chang, A. G. Yağlıkcı, S. Ghose, A. Agrawal, N. Chatterjee, A. Kashyap, D. Lee, M. O'Connor, H. Hassan, and O. Mutlu, "Understanding reduced-voltage operation in modern dram devices: Experimental characterization, analysis, and mechanisms," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 1, no. 1, pp. 1–42, 2017.
- [15] J. Choi, W. Shin, J. Jang, J. Suh, Y. Kwon, Y. Moon, and L.-S. Kim, "Multiple clone row dram: A low latency and area optimized dram," *ACM SIGARCH Computer Architecture News*, vol. 43, no. 3S, pp. 223–234, 2015.
- [16] Comet Technologies Canada Inc., "Dragonfly 2022.2," <https://www.theobjects.com/dragonfly>, 2023.
- [17] D. F. Courbon, "In-house transistors' layer reverse engineering characterization of a 45nm soc," in *ISTFA 2018*. ASM International, 2018, pp. 272–279.
- [18] F. Courbon, "Practical partial hardware reverse engineering analysis: For local fault injection and authenticity verification," *Journal of Hardware and Systems Security*, vol. 4, no. 1, pp. 1–10, 2020.
- [19] F. Courbon, J. J. Fournier, P. Loubet-Moundi, and A. Tria, "Combining image processing and laser fault injections for characterizing a hardware aes," *IEEE transactions on computer-aided design of integrated circuits and systems*, vol. 34, no. 6, pp. 928–936, 2015.
- [20] F. Courbon, S. Skorobogatov, and C. Woods, "Reverse engineering flash eeprom memories using scanning electron microscopy," in *International Conference on Smart Card Research and Advanced Applications*. Springer, 2016, pp. 57–72.
- [21] Q. Deng, L. Jiang, Y. Zhang, M. Zhang, and J. Yang, "Dracc: A dram based accelerator for accurate cnn inference," in *Proceedings of the 55th annual design automation conference*, 2018, pp. 1–6.
- [22] J. D. Ferreira, G. Falcao, J. Gómez-Luna, M. Alser, L. Orosa, M. Sadrosadati, J. S. Kim, G. F. Oliveira, T. Shahroodi, A. Nori, and O. Mutlu, "pluto: Enabling massively parallel computation in dram via lookup tables," in *2022 55th IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2022, pp. 900–919.
- [23] M. Fyrbiak, S. Strauß, C. Kison, S. Wallat, M. Elson, N. Rummel, and C. Paar, "Hardware reverse engineering: Overview and open challenges," in *2017 IEEE 2nd International Verification and Security Workshop (IVSW)*. IEEE, 2017, pp. 88–94.
- [24] F. Gao, G. Tziatzoulis, and D. Wentzlaff, "Computedram: In-memory compute using off-the-shelf drams," in *Proceedings of the 52nd annual IEEE/ACM international symposium on microarchitecture*, 2019, pp. 100–113.
- [25] M. Gao, C. Delimitrou, D. Niu, K. T. Malladi, H. Zheng, B. Brennan, and C. Kozyrakis, "Draf: A low-power dram-based reconfigurable acceleration fabric," *ACM SIGARCH Computer Architecture News*, vol. 44, no. 3, pp. 506–518, 2016.
- [26] M. M. Ghaffar, C. Sudarshan, C. Weis, M. Jung, and N. Wehn, "A low power in-dram architecture for quantized cnns using fast winograd convolutions," in *The International Symposium on Memory Systems*, 2020, pp. 158–168.
- [27] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM journal on imaging sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [28] N. Hajinazar, G. F. Oliveira, S. Gregorio, J. D. Ferreira, N. M. Ghiasi, M. Patel, M. Alser, S. Ghose, J. Gómez-Luna, and O. Mutlu, "Simdram: A framework for bit-serial simd processing using dram," in *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, 2021, pp. 329–345.
- [29] H. Hassan, M. Patel, J. S. Kim, A. G. Yaglikci, N. Vijaykumar, N. M. Ghiasi, S. Ghose, and O. Mutlu, "Crow: A low-cost substrate for improving dram performance, energy efficiency, and reliability," in *Proceedings of the 46th International Symposium on Computer Architecture*, 2019, pp. 129–142.
- [30] 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 *MICRO-54: 54th Annual IEEE/ACM International Symposium on Microarchitecture*, 2021, pp. 1198–1213.
- [31] M. Holler, M. Guizar-Sicairos, E. H. Tsai, R. Dinapoli, E. Müller, O. Bunk, J. Raabe, and G. Aepli, "High-resolution non-destructive three-dimensional imaging of integrated circuits," *Nature*, vol. 543, no. 7645, pp. 402–406, 2017.
- [32] S. Hong, S. Kim, J.-K. Wee, and S. Lee, "Low-voltage dram sensing scheme with offset-cancellation sense amplifier," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 10, pp. 1356–1360, 2002.
- [33] L. L. Hsu, R. V. Joshi, and R. Carl, "Compact dual-port dram architecture system and method for making same," Jan. 7 2003, uS Patent 6,504,204.
- [34] D. James, "Recent innovations in dram manufacturing," in *2010 IEEE/SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*. IEEE, 2010, pp. 264–269.
- [35] JEDEC, "DDR4 SDRAM SODIMM Design Specification," 2019.
- [36] JEDEC, "DDR4 SDRAM UDIMM Design Specification," 2019.
- [37] JEDEC Solid State Technology Association, "JESD79-4B, DDR4 Specification," 2017.
- [38] JEDEC Solid State Technology Association, "JESD79-5, DDR5 Specification," 2020.
- [39] H. Kaeslin, *Top-down digital VLSI design: from architectures to gate-level circuits and FPGAs*. Morgan Kaufmann, 2014.
- [40] M. Kammerstetter, M. Muellner, D. Burian, C. Platzer, and W. Kastner, "Breaking integrated circuit device security through test mode silicon reverse engineering," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 549–557.
- [41] H.-B. Kang, S.-K. Hong, H.-Y. Chang, H.-C. Park, N.-K. Park, M. Y. Sung, J.-H. Ahn, and S.-J. Hong, "A sense amplifier scheme with offset

- cancellation for giga-bit dram,” *Journal of Semiconductor Technology and Science*, vol. 7, no. 2, pp. 67–75, 2007.
- [42] B. Keeth, R. J. Baker, B. Johnson, and F. Lin, *DRAM circuit design: fundamental and high-speed topics*. John Wiley & Sons, 2007, vol. 13.
- [43] J. S. Kim, M. Patel, H. Hassan, L. Orosa, and O. Mutlu, “D-range: Using commodity dram devices to generate true random numbers with low latency and high throughput,” in *2019 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2019, pp. 582–595.
- [44] S. M. Kim, T. W. Oh, and S.-O. Jung, “Sensing voltage compensation circuit for low-power dram bit-line sense amplifier,” in *2018 International Conference on Electronics, Information, and Communication (ICEIC)*. IEEE, 2018, pp. 1–4.
- [45] S. M. Kim, B. Song, and S.-O. Jung, “Sensing margin enhancement technique utilizing boosted reference voltage for low-voltage and high-density dram,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 10, pp. 2413–2422, 2019.
- [46] S. M. Kim, B. Song, and S.-O. Jung, “Imbalance-tolerant bit-line sense amplifier for dummy-less open bit-line scheme in dram,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 6, pp. 2546–2554, 2021.
- [47] S. M. Kim, B. Song, T. W. Oh, and S.-O. Jung, “Analysis on sensing yield of voltage latched sense amplifier for low power dram,” in *2018 14th Conference on Ph. D. Research in Microelectronics and Electronics (PRIME)*. IEEE, 2018, pp. 65–68.
- [48] Y. Kim, V. Seshadri, D. Lee, J. Liu, and O. Mutlu, “A case for exploiting subarray-level parallelism (salp) in dram,” *ACM SIGARCH Computer Architecture News*, vol. 40, no. 3, pp. 368–379, 2012.
- [49] Y.-B. Kim and T. W. Chen, “Assessing merged dram/logic technology,” *Integration*, vol. 27, no. 2, pp. 179–194, 1999.
- [50] A. Kimura, J. Scholl, J. Schaffranek, M. Sutter, A. Elliott, M. Strizich, and G. D. Via, “A decomposition workflow for integrated circuit verification and validation,” *Journal of Hardware and Systems Security*, vol. 4, pp. 34–43, 2020.
- [51] T. Kirihata, G. Mueller, M. Clinton, S. Loeffler, B. Ji, H. Terletzki, D. Hanson, C.-L. Hwang, G. Lehmann, D. Storaska, G. Daniel, L. Hsu, O. Weinfurter, T. Boehler, J. Schnell, G. Frankowsky, D. Netis, J. Ross, A. Reith, O. Kiehl, and M. Wordeman, “A 113 mm/sup 2/600 mb/s/pin 512 mb ddr2 sdram with vertically-folded bitline architecture,” in *2001 IEEE International Solid-State Circuits Conference. Digest of Technical Papers. ISSCC (Cat. No. 01CH37177)*. IEEE, 2001, pp. 382–383.
- [52] C. Kison, J. Frinken, and C. Paar, “Finding the aes bits in the haystack: Reverse engineering and sca using voltage contrast,” in *Cryptographic Hardware and Embedded Systems—CHES 2015: 17th International Workshop, Saint-Malo, France, September 13-16, 2015, Proceedings 17*. Springer, 2015, pp. 641–660.
- [53] A. Kotabe, Y. Yanagawa, S. Akiyama, and T. Sekiguchi, “0.5-v low-vt cmos preamplifier for low-power and high-speed gigabit-dram arrays,” *IEEE journal of solid-state circuits*, vol. 45, no. 11, pp. 2348–2355, 2010.
- [54] Z. Lang, P. Jattke, M. Marazzi, and K. Razavi, “Blaster: Characterizing the blast radius of rowhammer,” in *3rd Workshop on DRAM Security (DRAMSec) co-located with ISCA 2023*. ETH Zurich, 2023.
- [55] C. Lee, T. Yim, and H. Yoon, “Bit-line sense amplifier using pmos charge transfer pre-amplifier for low-voltage dram,” in *TENCON 2018-2018 IEEE Region 10 Conference*. IEEE, 2018, pp. 1357–1361.
- [56] C. Lee and H. Yoon, “Highly robust and sensitive charge transfer sense amplifier for ultra-low voltage drams,” in *Fifth Asia Symposium on Quality Electronic Design (ASQED 2013)*. IEEE, 2013, pp. 227–232.
- [57] D. Lee, Y. Kim, G. Pekhimenko, S. Khan, V. Seshadri, K. Chang, and O. Mutlu, “Adaptive-latency dram: Optimizing dram timing for the common-case,” in *2015 IEEE 21st International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2015, pp. 489–501.
- [58] D. Lee, Y. Kim, V. Seshadri, J. Liu, L. Subramanian, and O. Mutlu, “Tiered-latency dram: A low latency and low cost dram architecture,” in *2013 IEEE 19th International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2013, pp. 615–626.
- [59] M. J. Lee, “A sensing noise compensation bit line sense amplifier for low voltage applications,” *IEEE journal of solid-state circuits*, vol. 46, no. 3, pp. 690–694, 2011.
- [60] K.-N. Lim, W.-J. Jang, H.-S. Won, K.-Y. Lee, H. Kim, D.-W. Kim, M.-H. Cho, S.-L. Kim, J.-H. Kang, K.-W. Park, and B.-T. Jeong, “A 1.2 v 23nm 6f 2 4gb ddr3 sdram with local-bitline sense amplifier, hybrid lio sense amplifier and dummy-less array architecture,” in *2012 IEEE International Solid-State Circuits Conference*. IEEE, 2012, pp. 42–44.
- [61] B. Lippmann, N. Unverricht, A. Singla, M. Ludwig, M. Werner, P. Egger, A. Duebotzky, H. Graeb, H. Gieser, M. Rasche, and O. Kellermann, “Verification of physical designs using an integrated reverse engineering flow for nanoscale technologies,” *Integration*, vol. 71, pp. 11–29, 2020.
- [62] B. Lippmann, M. Werner, N. Unverricht, A. Singla, P. Egger, A. Dübotzky, H. Gieser, M. Rasche, O. Kellermann, and H. Graeb, “Integrated flow for reverse engineering of nanoscale technologies,” in *Proceedings of the 24th Asia and South Pacific Design Automation Conference*, 2019, pp. 82–89.
- [63] J. Liu, B. Jaiyen, Y. Kim, C. Wilkerson, and O. Mutlu, “An experimental study of data retention behavior in modern dram devices: Implications for retention time profiling mechanisms,” *ACM SIGARCH Computer Architecture News*, vol. 41, no. 3, pp. 60–71, 2013.
- [64] S.-L. Lu, Y.-C. Lin, and C.-L. Yang, “Improving dram latency with dynamic asymmetric subarray,” in *Proceedings of the 48th International Symposium on Microarchitecture*, 2015, pp. 255–266.
- [65] M. Ludwig, A.-C. Bette, and B. Lippmann, “Vital: Verifying trojan-free physical layouts through hardware reverse engineering,” in *2021 IEEE Physical Assurance and Inspection of Electronics (PAINE)*. IEEE, 2021, pp. 1–8.
- [66] H. Luo, T. Shahroodi, H. Hassan, M. Patel, A. G. Yağlıkçı, L. Orosa, J. Park, and O. Mutlu, “Clr-dram: A low-cost dram architecture enabling dynamic capacity-latency trade-off,” in *2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2020, pp. 666–679.
- [67] M. Marazzi, P. Jattke, F. Solt, and K. Razavi, “Protrr: Principled yet optimal in-dram target row refresh,” in *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2022, pp. 735–753.
- [68] M. Marazzi, F. Solt, P. Jattke, K. Takashi, and K. Razavi, “Rega: Scalable rowhammer mitigation with refresh-generating activations,” in *44rd IEEE Symposium on Security and Privacy (SP 2023)*. IEEE, 2023.
- [69] J. Moon and B. Chung, “Sense amplifier with offset mismatch calibration for sub 1-v dram core operation,” in *2010 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2010, pp. 3501–3504.
- [70] Y. Nakagome, M. Aoki, S. Ikenaga, M. Horiguchi, S. Kimura, Y. Kawamoto, and K. Itoh, “The impact of data-line interference noise on dram scaling,” *IEEE Journal of Solid-state circuits*, vol. 23, no. 5, pp. 1120–1127, 1988.
- [71] M. Nakamura, T. Takahashi, T. Akiba, G. Kitsukawa, M. Morino, T. Sekiguchi, I. Asano, K. Komatsuzaki, Y. Tadaki, S. Cho, K. Kajigaya, T. Tachibana, and K. Sato, “A 29-ns 64-mb dram with hierarchical array architecture,” *IEEE Journal of Solid-State Circuits*, vol. 31, no. 9, pp. 1302–1307, 1996.
- [72] H. Nam, S. Baek, M. Wi, M. J. Kim, J. Park, C. Song, N. S. Kim, and J. H. Ahn, “X-ray: Discovering dram internal structure and error characteristics by issuing memory commands,” *IEEE Computer Architecture Letters*, 2023.
- [73] H. Oh, J. Kim, J. Kim, S. Park, D. Kim, S. Kim, D. Woo, Y. Lee, G. Ha, J. Park, N. Kang, H. Kim, J. Hwang, B. Kim, D. Kim, Y. Cho, J. Choi, B. Lee, S. Kim, M. Cho, Y. Kim, J. Choi, D. Shin, M. Shim, W. Choi, G. Lee, Y. Park, W. Lee, and B. Ryu, “High-density low-power-operating dram device adopting 6f/sup 2/cell scheme with novel s-rcat structure on 80nm feature size and beyond,” in *Proceedings of 35th European Solid-State Device Research Conference, 2005. ESSDERC 2005*. IEEE, 2005, pp. 177–180.
- [74] O. Okobiah, S. P. Mohanty, E. Kougiianos, and M. Poolakkaparambil, “Towards robust nano-cmos sense amplifier design: a dual-threshold versus dual-oxide perspective,” in *Proceedings of the 21st edition of the great lakes symposium on Great lakes symposium on VLSI*, 2011, pp. 145–150.
- [75] L. Orosa, Y. Wang, M. Sadrosadati, J. S. Kim, M. Patel, I. Puddu, H. Luo, K. Razavi, J. Gómez-Luna, H. Hassan, N. Mansouri-Ghiasi, S. Ghose, and O. Mutlu, “Codic: A low-cost substrate for enabling custom in-dram functionalities and optimizations,” in *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2021, pp. 484–497.
- [76] J. Park, D.-H. Shin, Y.-H. Cho, and K.-W. Kwon, “Inverted bit-line sense amplifier with offset-cancellation capability,” *Electronics Letters*, vol. 52, no. 9, pp. 692–694, 2016.
- [77] J. M. Park, Y. S. Hwang, S.-W. Kim, S. Y. Han, J. S. Park, J. Kim, J. W. Seo, B. S. Kim, S. H. Shin, C. H. Cho, S. W. Nam, H. S. Hong,

- K. P. Lee, G. Y. Jin, and E. S. Jung, "20nm dram: A new beginning of another revolution," in *2015 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2015, pp. 26–5.
- [78] R. V. W. Putra, M. A. Hanif, and M. Shafique, "Sparkxd: A framework for resilient and energy-efficient spiking neural network inference using approximate dram," in *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2021, pp. 379–384.
- [79] R. Quijada, R. Dura, J. Pallares, X. Formatje, S. Hidalgo, and F. Serra-Graells, "Large-area automated layout extraction methodology for full-ic reverse engineering," *Journal of Hardware and Systems Security*, vol. 2, no. 4, pp. 322–332, 2018.
- [80] M. T. Rahman, Q. Shi, S. Tajik, H. Shen, D. L. Woodard, M. Tehranipoor, and N. Asadizanjani, "Physical inspection & attacks: New frontier in hardware security," in *2018 IEEE 3rd International Verification and Security Workshop (IVSW)*. IEEE, 2018, pp. 93–102.
- [81] N. Rohbani, S. Darabi, and H. Sarbazi-Azad, "Pf-dram: a precharge-free dram structure," in *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2021, pp. 126–138.
- [82] N. Rohbani, M. A. Soleimani, and H. Sarbazi-Azad, "Pipf-dram: processing in precharge-free dram," in *Proceedings of the 59th ACM/IEEE Design Automation Conference*, 2022, pp. 1075–1080.
- [83] N. Rohbani, M. A. Soleimani, and H. Sarbazi-Azad, "Cooldram: An energy-efficient and robust dram," in *2023 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*. IEEE, 2023, pp. 1–6.
- [84] R. Rosenkranz, "Failure localization with active and passive voltage contrast in fib and sem," *Journal of Materials Science: Materials in Electronics*, vol. 22, pp. 1523–1535, 2011.
- [85] S. Roy, M. Ali, and A. Raghunathan, "Pim-dram: Accelerating machine learning workloads using processing in commodity dram," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 11, no. 4, pp. 701–710, 2021.
- [86] H. Seol, W. Shin, J. Jang, J. Choi, J. Suh, and L.-S. Kim, "In-dram data initialization," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 11, pp. 3251–3254, 2017.
- [87] O. Seongil, Y. H. Son, N. S. Kim, and J. H. Ahn, "Row-buffer decoupling: A case for low-latency dram microarchitecture," in *2014 ACM/IEEE 41st International Symposium on Computer Architecture (ISCA)*. IEEE, 2014, pp. 337–348.
- [88] V. Seshadri, D. Lee, T. Mullins, H. Hassan, A. Boroumand, J. Kim, M. A. Kozuch, O. Mutlu, P. B. Gibbons, and T. C. Mowry, "Ambit: In-memory accelerator for bulk bitwise operations using commodity dram technology," in *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture*, 2017, pp. 273–287.
- [89] S. M. Seyedzadeh, D. Kline Jr, A. K. Jones, and R. Melhem, "Mitigating bitline crosstalk noise in dram memories," in *Proceedings of the International Symposium on Memory Systems*, 2017, pp. 205–216.
- [90] S. M. Sharroush, "A precharged bitline 1t-1c dram readout scheme," *Microelectronics Journal*, vol. 83, pp. 168–184, 2019.
- [91] W. Shin, J. Choi, J. Jang, J. Suh, Y. Kwon, Y. Moon, H. Kim, and L.-S. Kim, "Q-dram: Quick-access dram with decoupled restoring from row-activation," *IEEE Transactions on Computers*, vol. 65, no. 7, pp. 2213–2227, 2015.
- [92] J.-Y. Sim, "Circuit design of dram for mobile generation," *Journal of Semiconductor Technology and Science*, vol. 7, no. 1, pp. 1–10, 2007.
- [93] G. Singh, A. Wagle, S. Khatri, and S. Vrudhula, "Cidan-xe: Computing in dram with artificial neurons," *Frontiers in Electronics*, vol. 3, p. 834146, 2022.
- [94] Y. H. Son, O. Seongil, Y. Ro, J. W. Lee, and J. H. Ahn, "Reducing memory access latency with asymmetric dram bank organizations," in *Proceedings of the 40th annual international symposium on computer architecture*, 2013, pp. 380–391.
- [95] A. Spessot and H. Oh, "1t-1c dynamic random access memory status, challenges, and prospects," *IEEE Transactions on Electron Devices*, vol. 67, no. 4, pp. 1382–1393, 2020.
- [96] J.-W. Sub, K.-M. Rho, C.-K. Park, and Y.-H. Koh, "Offset-trimming bit-line sensing scheme for gigabit-scale dram's," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 7, pp. 1025–1028, 1996.
- [97] L. Subramanian, K. Vaidyanathan, A. Nori, S. Subramoney, T. Karnik, and H. Wang, "Closed yet open dram: achieving low latency and high performance in dram memory systems," in *Proceedings of the 55th Annual Design Automation Conference*, 2018, pp. 1–6.
- [98] C. Sudarshan, J. Lappas, M. M. Ghaffar, V. Rybalkin, C. Weis, M. Jung, and N. Wehn, "An in-dram neural network processing engine," in *2019 IEEE international symposium on circuits and systems (ISCAS)*. IEEE, 2019, pp. 1–5.
- [99] C. Sudarshan, L. Steiner, M. Jung, J. Lappas, C. Weis, and N. Wehn, "A novel dram architecture for improved bandwidth utilization and latency reduction using dual-page operation," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 5, pp. 1615–1619, 2021.
- [100] T. Sugawara, D. Suzuki, R. Fujii, S. Tawa, R. Hori, M. Shiozaki, and T. Fujino, "Reversing stealthy dopant-level circuits," in *Cryptographic Hardware and Embedded Systems—CHES 2014: 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings 16*. Springer, 2014, pp. 112–126.
- [101] TechInsights Inc., "Micron 1a dram technology," <https://www.techinsights.com/blog/memory/micron-1a-dram-technology>, 2023.
- [102] TechInsights Inc., "The semiconductor information platform," <https://www.techinsights.com>, 2023.
- [103] TechInsights Inc., "Terms and conditions – content licensing," <https://www.techinsights.com/sites/default/files/2023-09/Ts%26Cs%20-%20Content%20Licensing%202023%20-%20v1.1.pdf>, 2023.
- [104] Thermo Fisher Scientific Inc., "Avizo software," <https://www.thermofisher.com/ch/en/home/electron-microscopy/products/software-em-3d-vis/avizo-software.html>, 2023.
- [105] Thermo Fisher Scientific Inc., "Helios 5 ux dualbeam for materials science," <https://www.thermofisher.com/order/catalog/product/HELIOS5UX>, 2023.
- [106] R. Torrance and D. James, "The state-of-the-art in ic reverse engineering," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2009, pp. 363–381.
- [107] P. Trampert, F. Bourghorbel, P. Potocek, M. Peemen, C. Schlinkmann, T. Dahmen, and P. Slusallek, "How should a fixed budget of dwell time be spent in scanning electron microscopy to optimize image quality?" *Ultramicroscopy*, vol. 191, pp. 11–17, 2018.
- [108] A. Vijayakumar, V. C. Patil, D. E. Holcomb, C. Paar, and S. Kundu, "Physical design obfuscation of hardware: A comprehensive investigation of device and logic-level techniques," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 64–77, 2016.
- [109] S. Wallat, N. Albartus, S. Becker, M. Hoffmann, M. Ender, M. Fyrbiak, A. Drees, S. Maaßen, and C. Paar, "Highway to hal: open-sourcing the first extendable gate-level netlist reverse engineering framework," in *Proceedings of the 16th ACM International Conference on Computing Frontiers*, 2019, pp. 392–397.
- [110] Y. Watanabe, N. Nakamura, and S. Watanabe, "Offset compensating bit-line sensing scheme for high density dram's," *IEEE journal of solid-state circuits*, vol. 29, no. 1, pp. 9–13, 1994.
- [111] M. Wi, J. Park, S. Ko, M. J. Kim, N. S. Kim, E. Lee, and J. H. Ahn, "Shadow: Preventing row hammer in dram with intra-subarray row shuffling," in *2023 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. IEEE, 2023, pp. 333–346.
- [112] X. Xin, Y. Zhang, and J. Yang, "Elp2im: Efficient and low power bitwise operation processing in dram," in *2020 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2020, pp. 303–314.
- [113] A. G. Yağlıkçı, H. Luo, G. F. De Oliviera, A. Olgun, M. Patel, J. Park, H. Hassan, J. S. Kim, L. Orosa, and O. Mutlu, "Understanding rowhammer under reduced wordline voltage: An experimental study using real dram devices," in *2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 2022, pp. 475–487.
- [114] Z. Yang and S. Mourad, "Crosstalk induced fault analysis and test in drams," *Journal of Electronic Testing*, vol. 22, pp. 173–187, 2006.
- [115] J. Yeung and H. Mahmoodi, "Robust sense amplifier design under random dopant fluctuations in nano-scale cmos technologies," in *2006 IEEE International SOC Conference*. IEEE, 2006, pp. 261–264.
- [116] H. Yoon, J. Y. Sim, H. S. Lee, K. N. Lim, J. Y. Lee, N. J. Kim, K. Y. Kim, S. M. Byun, W. S. Yang, C. H. Choi, H. S. Jeong, J. H. Yoo, D. I. Seo, K. Kim, B. I. Ryu, and C. G. Hwang, "A 4 gb ddr sdram with gain-controlled pre-sensing and reference bitline calibration schemes in the twisted open bitline architecture," in *2001 IEEE International Solid-State Circuits Conference. Digest of Technical Papers. ISSCC (Cat. No. 01CH37177)*. IEEE, 2001, pp. 378–379.
- [117] Zentel Japan Corp., "Zentel home," <https://www.zentel-japan.com/>, 2023.

- [118] W. Zhou, R. Apkarian, Z. L. Wang, and D. Joy, "Fundamentals of scanning electron microscopy (sem)," *Scanning Microscopy for Nanotechnology: Techniques and Applications*, pp. 1–40, 2007.