



HAL
open science

Smart write algorithm to enhance performances and reliability of a RRAM macro

Bastien Giraud, Sebastien Ricavy, Cyrille Laffond, Ilan Sever, Valentin Gherman, Florent Lepin, Mariam Diallo, Khadija Zenati, Sylvain Dumas, Olivier Guille, et al.

► To cite this version:

Bastien Giraud, Sebastien Ricavy, Cyrille Laffond, Ilan Sever, Valentin Gherman, et al.. Smart write algorithm to enhance performances and reliability of a RRAM macro. IEEE Journal of Solid-State Circuits, 2023, 59 (9), pp.3045 - 3057. 10.1109/JSSC.2024.3386429 . cea-04771574

HAL Id: cea-04771574

<https://cea.hal.science/cea-04771574v1>

Submitted on 18 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Smart Write Algorithm to Enhance Performances and Reliability of a RRAM Macro

B. Giraud, S. Ricavy, C. Laffond, I. Sever, V. Gherman, F. Lepin, M. Diallo, K. Zenati, S. Dumas, O. Guille, M. Vershkov, A. Bricalli, G. Piccolboni, J.-P. Noel, A. Samir, G. Pillonnet, *Senior Member, IEEE*, Y. Thonnart, *Member, IEEE*, G. Molas, *Senior Member, IEEE*

Abstract— This paper presents a comprehensive assessment of the impact of various design assist techniques on the inherent performance and reliability of native RRAM on silicon. The collaborative optimization of design and technology plays a crucial role in replacing conventional flash memory as the leading solution. We showcase that employing read-before-write, current-limitation, and write-termination techniques result in reductions of power consumption during programming operations by 47%, 56%, and 13% respectively. Through their combination with write verification and error correction code mechanisms, these enhancements collectively achieve 83% reduction in energy consumption and a notable 55% decrease in access time. These advancements are made possible by the introduction of a novel smart write algorithm (SWA). Leveraging a representative 128kb RRAM macro implemented in 130nm CMOS technology, this study significantly contributes to the feasible integration of RRAM for embedded applications. Experimental assessments on silicon validate the increased reliability, with a demonstrated 28.1 μ A read margin after undergoing 1 million cycles without encountering any read errors, maintaining a Bit Error Rate (BER) below 10^{-7} .

Index Terms—Non-volatile memory, ECC, smart algorithm, Adaptive and Reconfigurable Systems, Variant-tolerant

I. INTRODUCTION

IN sub-40nm CMOS technologies, the dominance of Resistive RAM (RRAM) over NOR eFlash is being contested due to the prohibitive cost and manufacturing complexities associated with the latter [1-2]. Even for older technology nodes, RRAM is challenging the supremacy of eFlash, particularly for applications that demand high-energy efficiency and robust security (attributed to its lowered vulnerability to side-channel attacks as a result of reduced power consumption). This technology boasts shorter programming times, lower operating voltages, and compatibility with standard CMOS processes (requiring only 2 additional masks). Nevertheless, the stochastic nature of RRAM programming, coupled with issues in reduced read margins and data retention, collectively limit the overall memory macro performance [3-6].

Leading foundries have recently embraced RRAM technology through substantial efforts in circuit design [7-8]. This design-technology co-optimization has yielded improvements in read-margin (RM) and energy consumption [9-10]. Besides, the endeavor to build systems based on RRAM devices encounters

technical hurdles like increasing cycle-to-cycle variability, leading to programming complexities, as well as reductions in endurance and read margin. Prior researches [11-16] have introduced several programming design assist techniques to ensure consistent RRAM functionality across all cycles or to enhance the performance of memory macros. However, the divergence in RRAM processes and design methods in previous studies has led to incomplete reporting of the individual and synergistic advantages in the trade-off between RM, endurance, power consumption, and speed.

This paper builds upon our recently published work [17], offering an in-depth exploration of these interdependencies in pursuit of performance and endurance enhancement. The manuscript delves into a comprehensive analysis of programming strategies, including read-before-write (RBW), current limiter (CL), write termination (WT), write verification (WV), and error correction code (ECC), orchestrated by a versatile smart write algorithm (SWA). Furthermore, the paper investigates architectural and circuit aspects, along with silicon-based data, while applying our methodology for fine-tuning a wide array of key parameters. The resultant performance enhancements achieved through the combination of these design-aided techniques are validated through silicon-based experimentation, with comparisons drawn against the original RRAM macro performance and other relevant publications.

The subsequent sections of this paper are organized as follows: Section II presents the integration of design assist techniques within the proposed SWA; Section III describes the testchip, RRAM macro architecture, and its interface with the RISC-V sub-system leveraging SWA; Section IV finally presents silicon results.

II. DESIGN ASSIST TECHNIQUES AND SWA

A. Design Techniques

Column-wise bitcell programming is achieved by employing a distinct programming current and biasing voltage. The programming current and voltage can be independently fine-tuned for both Set and Reset operations, offering substantial flexibility in their adjustment. This adaptability is essential to guarantee the operational integrity of the entire array, effectively addressing variations in cell-to-cell attributes,

(Corresponding author: Bastien Giraud).

B. Giraud, C. Laffond, V. Gherman, F. Lepin, M. Diallo, K. Zenati, J.-P. Noel and Y. Thonnart are with Univ. Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France (e-mail: bastien.giraud@cea.fr).

S. Ricavy, S. Dumas, O. Guille, A. Samir and G. Pillonnet are with Univ. Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France.

I. Sever, M. Vershkov, A. Bricalli and G. Piccolboni are with WeeBit Nano Ltd.

G. Molas was with Univ. Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France. He is now with WeeBit Nano Ltd.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

filament growth mechanisms, and the dynamic IR drops stemming from the application of high programming currents. These intrinsic variabilities degrade the read margin and detrimentally affect the attainable bit error rate (BER).

In this study, the aim is to enhance the inherent performance and energy efficiency of the RRAM macro, achieved by incorporating the following design assist techniques.

Read-Before-Write (RBW) involves validating the bit value before initiating the write process. The primary objective is to save write energy in cases where the bitcell already holds the intended state. This strategy is particularly effective due to the notably lower read energy when compared to write energy, along with the indefinite read endurance. Furthermore, RBW contributes to enhanced aging characteristics and prevents stuck bits arising from consecutive Set/Reset operations.

Current Limiter (CL) involves restricting the current during the Set operation (transitioning from high to low resistive state) to prevent the oxide breakdown of the RRAM device caused by high currents. Figure 1 illustrates our schematic for the CL, positioned on the source line (SL) side opposite to the current injection facilitated by the LDO. Consequently, the current limiter operates autonomously from the LDO's operation and triggers when the current in the bitcell exceeds the reference value predefined by the current source (I_{LIMREF}). We define the current limitation factor as $CL_{RATIO} = I_{LIMREF} / I_{LIMREFmax}$. This mechanism ensures swift and effective current control, inhibiting the occurrences of excessive current overshoots.

The **Write Termination (WT)** approach, whether sudden or gradual, serves to halt a write process once the filament formation for Set or filament disruption for Reset is achieved, optionally incorporating a controlled time delay. The primary purpose of WT is to ensure uniform programming energies across various written bitcells, irrespective of their transient or spatial variability. The underlying principle of WT is to discontinue the writing operation as soon as the programming current, denoted as I_{Sense} , reaches an adjustable threshold current, I_{WT_REF} . This is accomplished through a current comparator, as depicted in Figure 1. I_{Sense} encompasses both the effective current traversing the selected bitcell (I_{RRAM}) and the idle current in the LDO (I_{idle}). The latter is designed to be negligible in comparison to the bitcell current. Consequently, I_{Sense} effectively represents a downscaled counterpart of the bitcell current, I_{RRAM} . As illustrated by the waveforms in Figure 1, a signal flag, $WT_OK(j)$, is asserted following a delay of Td_WT once the write termination is detected. The implementation of WT improves programming time and efficiency but also contributes to aging mitigation, ultimately yielding a narrower statistical distribution of the resistances resulting from Set or Reset operations.

The **Write Verify (WV)** technique entails validating the bit values that have just been written by executing a subsequent read operation. Indeed, RRAM encounters an additional challenge wherein the memory window reduces as write cycling progresses, along with relaxation phenomena (refer to the silicon data in Section IV). This difficulty may invalidate the time-zero tuning to minimize the failed-bit count during wafer sorting, as it does not track the changing read window. The fundamental principle behind WV involves iterative rewriting until all bits are confirmed to be successfully programmed or until a predetermined maximum count of WV attempts is reached. The read operation is conducted employing the same voltage-based Sense Amplifier (SA) as utilized in conventional read operations, as depicted in Figure 2. In place of using RRAM elements, which might exhibit drift over time, the reference resistance is produced using an adjustable resistor network. It can be configured with a range of polysilicon resistance values, spanning from $5K\Omega$ to $200K\Omega$. This value, distinct from the regular read operation, is empirically predetermined to ensure the required balance between read stability, write speed, write endurance, and data retention. This adaptability allows for the alteration of reference values between Set and Reset phases, effectively mitigating Set/Reset tail distributions. Additionally, PMOS diodes are connected to the inputs of the SA, strategically incorporated to ensure that the voltage difference between the selected bitline, $BL(j)$, and the chosen $SL(j)$ consistently remains beneath $0.4V$ in order to prevent data from being lost during read (also known as read disturb phenomenon).

A binary linear block **Error Correction Code (ECC)** is used, characterized by its capability to correct double-bit errors. This dictates the use of a 44-bit code word comprising 32 data bits and 12 check bits. Additionally, a partial 3-bit error detection feature has been incorporated without necessitating an increase

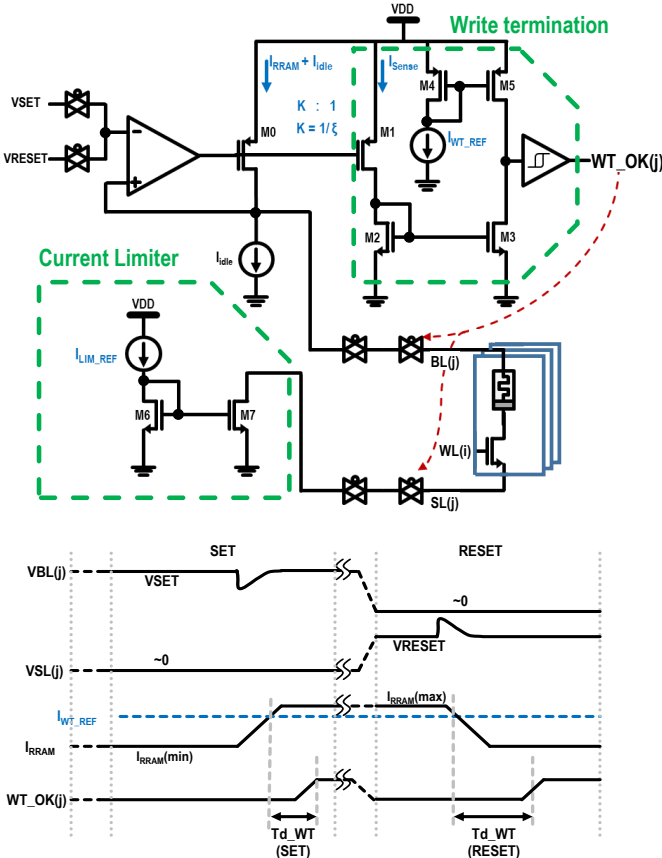


Fig. 1. Implementation and timing diagrams of the key analog circuits used for write operations.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

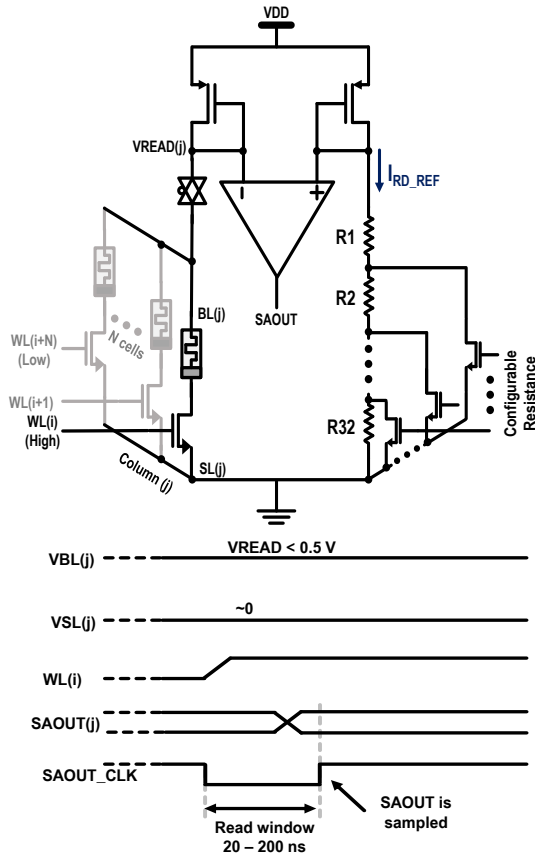


Fig. 2. Implementation and timing diagrams of the key analog circuits used for read operations.

in the number of check bits. Consequently, this configuration enables the identification of 78% of all possible 3-bit errors per code word. Another salient property of this ECC is its inversion invariance, signifying that the inversion of every bit within a code word yields another valid code word [18]. This is especially relevant for testing and forming: given that any linear ECC contains an *all-0* code word, the inversion invariance implies the existence of an *all-1* code word as well, which corresponds to the pre-forming memory state. It obviates the requirement for a specialized test access mechanism to bypass the ECC encoder for direct control over data and check-bit storage locations. Moreover, ECCs featuring inversion invariance enhance the safeguarding of RRAMs employing differential encoding for stored information. In this scenario, each bit of a memory word is encoded across two resistive devices, as a combination of two complementary binary values (such as '01' and '10' for encoding logical '1' and '0', respectively). This results in storing a memory word across two sets of resistive devices: one containing the ECC encoder-generated code word, and the other comprising the bitwise complement of this code word, which coincidentally is also a valid code word. Reading a stored memory word is viable either in a differential mode (using a 2T2R configuration) based on a single sense operation or in a non-differential mode (1T1R configuration), usually relying on two sequential sense operations. Both memory words sensed in non-differential mode can fully benefit from the ECC decoder's comprehensive error-correction capabilities.

Actually, RBW, WV and ECC are digitally handled while CL and WT use analog signals control with extra analog circuits.

In this study, an extensive design of experiments was conducted to establish optimized programming parameters (voltage, current, and duration) that facilitated the achievement of both minimal BER and high reliability (including retention and endurance). The effectiveness of the design techniques mentioned earlier will be evaluated using silicon data, as detailed in Section IV.

B. Smart Write Algorithm (SWA) Presentation

Forming and training loops play a pivotal role in securing the dependable long-term functionality of filament creation and disruption for SET and RESET operations. The algorithm governing SET and RESET operations involves a trade-off between write speed, write endurance, and data retention. For instance, executing SET and RESET operations with higher voltage/current levels and extended write pulses can enhance data retention, yet it simultaneously affects write endurance and write speed after cycling, creating a trade-off. Grounded in this principle, multiple strategies have been devised to enhance write speed, energy efficiency, and overall reliability.

Figure 3 outlines the integrated Smart Write Algorithm, which encompasses two distinct routes:

- The write path encompasses ECC encoding, RBW, a Set phase consisting of CL and WT, a Reset phase also incorporating CL and WT, along with WV.
- The read path entails SA operations and ECC_DEC (data decoding).

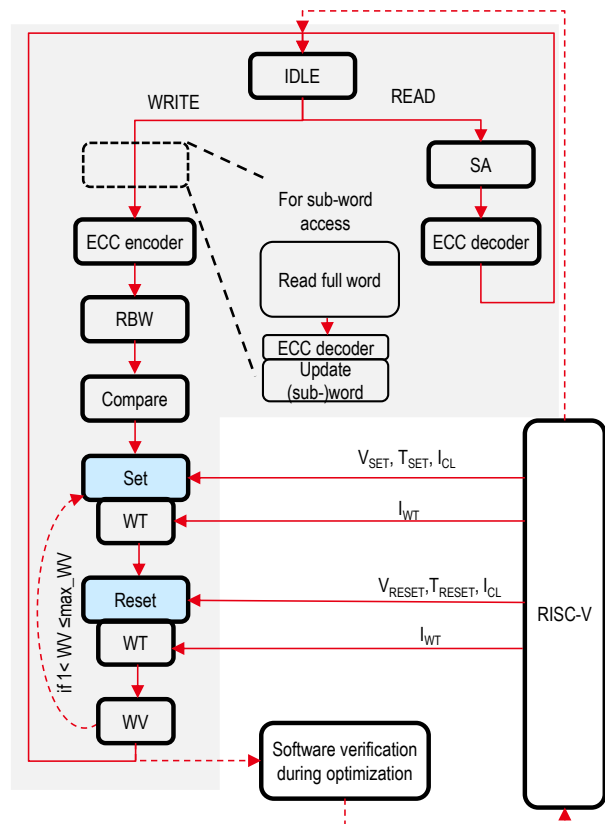


Fig. 3. SWA description.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

For sub-word writing scenarios, additional preliminary steps become necessary. Initially, the relevant full word needs to be read and corrected. Subsequently, the corresponding data bytes are modified to enable the generation of check bits for the resultant data word during the course of the complete word writing path.

The writing parameters include the Set and Reset voltages ($V_{\text{SET/RESET}}$), the duration of the write pulse ($T_{\text{SET/RESET}}$), the current limitation factor (CL_{RATIO}), and the current for write termination (I_{WT}), as illustrated in Figure 3.

Each of these parameters can be independently fine-tuned across a broad spectrum of values. In particular, the Set and Reset voltages are adjustable within the range of 1.2V to 4.6V. This adaptability is achieved by employing two 5-bit voltage Digital-to-Analog Converters (DACs), which offer a total of 1024 feasible combinations. The pulse width can span from 5ns to 5 μ s, with shorter durations often allocated for read operations. The range of the current limit and write termination current can be modulated using two 4-bit current DACs, spanning from 50 μ A to 425 μ A and from 60 μ A to 210 μ A, respectively, resulting in 256 potential combinations. In cases where the termination current exceeds the limit current, termination detection will not occur; hence termination current is always set below current limit. Moreover, WV allows up to 32 write repetitions thanks to a dedicated register.

The adjustability of all these parameters is facilitated through a RISC-V interface, enabling the tailoring of SWA patterns to suit the specific requirements of the target application. The RISC-V core is in charge of optimizing SWA parameters through software-based tuning loops during the initial stages of the chip's operational life. It proceeds as described by Algorithm 1 by estimating the distribution overlap of LRS and HRS currents on a full update of the memory array (or a subset), and then by greedily optimizing parameters in order (voltages, timings, CL_{RATIO} , WT, WV, Read reference), within configurable limits (e.g. T_{set} can be forced $\leq 1\mu$ s). This greedy heuristic was chosen for fast optimization based on estimated relative effects of the proposed techniques. Thanks to software flexibility, it would be possible to apply more complex optimization schemes such as gradient descent. Once the optimization phase is completed, these parameters are permanently stored within the RRAM macro itself. This configuration grants a high level of flexibility in individually determining the optimal settings, thereby enabling a systematic evaluation of performance and reliability trade-offs.

III. TESTCHIP DESCRIPTION

Figure 4 illustrates the comprehensive functional block diagram of our testchip that encompasses a 128kb RRAM macro seamlessly integrated with conventional SRAM interfaces, within a RISC-V subsystem employing a 32-bit data bus. SWA is integrated in a software-controlled Finite State Machine (FSM). Our RRAM system integrates the five essential digitally-configurable assist techniques (as elaborated in Section II.A.), thus facilitating the integration of an intelligent and adaptable RRAM SWA in conjunction with RISC-V. This integration serves a dual purpose of characterizing RRAM

Algorithm 1 SWA parameter optimization

Input: read voltage (V_{read}) (*design choice*),
paramValidAddress, paramValidKey (constants)
 initial SWA parameters (*estimated from simulation*)
 $Params = (V_{\text{set}}, V_{\text{reset}}, T_{\text{set}}, T_{\text{reset}}, I_{\text{CLset}}, I_{\text{CLreset}}, I_{\text{WTset}}, I_{\text{WTreset}}, R_{\text{WVset}}, R_{\text{WVreset}}, Max_{\text{WV}}, R_{\text{read-ref}})$

Output: *Params* containing final SWA parameters
 HRS_LRS_DISTRIBS(*Params*)

for all possible R_{read}
 $countLRS[R_{\text{read}}] \leftarrow 0$, $countHRS[R_{\text{read}}] \leftarrow 0$
 CONFIGURE deactivation of RRAM ECC
for $a \in$ RRAM addresses (*all or a subset*)
 WRITE *all-1* @ a
 $prev \leftarrow all-0$
 for R_{read} in all possible values in increasing order
 $next \leftarrow \text{READ } @a$
 for $i \in [0:31]$
 if $next[i] \neq prev[i]$ **then** increment $countLRS[R_{\text{read}}]$
 $prev \leftarrow next$
 Similarly *for* $countHRS$, write *all-0*, check in decreasing R_{read}
return $countLRS, countHRS$

PARAMSCORE($countLRS, countHRS$)
 $score \leftarrow 0$
for all possible R_{read}
 $score \leftarrow score + \min(countLRS[R_{\text{read}}], countHRS[R_{\text{read}}])$
 (*count misread values as area overlap of distributions*)
return $score$

OPTIMIZEPARAMETERS(*Params*)
for p iterating over *Params* except $R_{\text{read-ref}}$ (*in order as above*)
 $Params[p] \leftarrow \text{argmin}(\text{PARAMSCORE}(\text{HRS_LRS_DISTRIBS}(Params)))$
 (*minimize over valid values of Params[p]*)
 $R_{\text{read-ref}} = \text{mean}(R_{\text{WVset}}, R_{\text{WVreset}})$
return *Params*

STOREPARAMETERS
 WRITE *Params* @ *paramValidAddress*+1 sqq. in RRAM
 WRITE *paramValidKey* @ *paramValidAddress*

TESTSTOREDPARAMETERS
 $Params \leftarrow$ LOAD initial P from data SRAM
 $k \leftarrow$ READ @ *paramValidAddress* from RRAM
if read succeeded and $k = paramValidKey$ **then**
 $Params \leftarrow$ READ @ *paramValidAddress*+1 sqq. from RRAM
return 1, *Params*
return 0, *Params*

behavior and optimizing trade-offs among speed, energy efficiency, and endurance. Each of the 44 Input/Output columns of the RRAM subsystem is coupled to write circuits that encapsulate design assist techniques like CL and WT, along with read circuits, which incorporate SA.

The physical layout of our system is shown in the microphotograph and cross-sectional view illustrated in Figure 5. The used RRAM bitcells correspond to the Oxide-based type, namely OxRAM. They are manufactured in the back end of line of the 130nm CMOS process technology, strategically positioned between metal layers 4 and 5. The voltage supply for analog and digital circuits is tailored to 4.6V and 1.2V, respectively. The main features of the RRAM macro are listed in Table I. The RRAM macro surrounded by programmable circuitries (CL, WT, ECC) orchestrated by the SWA. Two 88kb RRAM arrays are integrated, each comprising 64k data bits and 24k check bits (for ECC). Each array, organized with 256 rows and 352 columns, is further divided into 22 column groups. Columns in each group are multiplexed to the data IO circuits (read and write) using MUX-16 IO slices. Memory capacity was chosen to provide sufficient representative statistical data concerning cell-to-cell variations, a fundamental requirement for the ongoing study.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

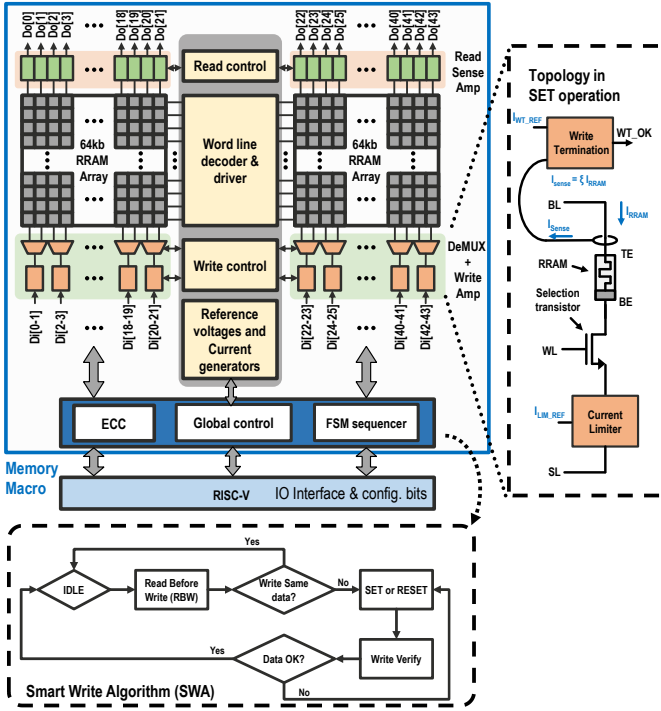


Fig. 4. Functional block diagram of the RRAM macro including Smart Write Algorithm.

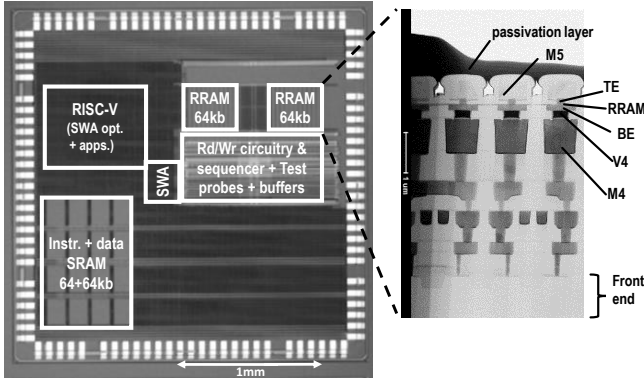


Fig. 5. Die photography and technology cross sectional view.

TABLE I
MAIN DEMONSTRATOR FEATURES

Feature	This work
CMOS and RRAM Technology	130nm and OxRAM
Supply voltages	1.2/4.6 V
Cell Size	61F ²
Capacity	128kb (+ 48kb for ECC)
Number of arrays	2
Array organization	352 columns x 256 rows
Write access granularity (bits)	MUX-16
Bits masking feature	from 1 to 22
Array efficiency for 1Mb (array/total area)	yes
SWA + ECC area w.r.t. RRAM macro area	53% ^a
	9% + 2%

^a estimated

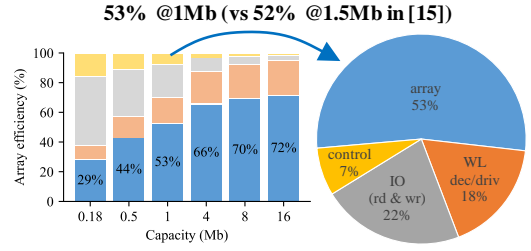


Fig. 6. Estimated array efficiency (ie. ratio of array area / full macro area) versus memory capacity.

An imperative consideration when assessing the benefits of design assist techniques pertains to the associated area overhead. Given the relatively constrained memory capacity, the impact of surrounding circuits' area penalty becomes an essential factor. To gauge the scale of this effect, we performed an extrapolated estimation of area overhead encompassing control circuits, IO, Word Line (WL) drivers, and decoders, spanning memory capacities from 0.18Mb to 16 Mb, as illustrated in Figure 6. The outcome reveals an area penalty of 47% for 1Mb, which is consistent with prior research [15], and a reduction to as low as 28% for 16Mb. Additionally, it's noteworthy that in many application scenarios, the need for flexible programmability diminishes, thereby leading to further enhancements in array efficiency.

IV. EXPERIMENTAL MEASUREMENTS

In this section, the silicon data showcased have been obtained from a specific circuit chosen as the optimal candidate during the measurement campaign. We will detail our approach for fine-tuning programming duration and current parameters, aiming to attain optimal optimization through the utilization of SWA. This method takes into account factors such as write time and energy consumption.

A. Improvements of each Design Technique

Figure 7 illustrates the read current distribution measured on the 128kb memory macro, specifically for High-Resistance State (HRS) and Low-Resistance State (LRS), excluding check bits. The distributions have been acquired by extensively adjusting the reference resistance values over a broad range, spanning from 5K Ω to 200K Ω (as depicted in Fig. 2), utilizing varying programming durations and CL_{RATIO} set at 0.5, all conducted at ambient temperature. The applied pulse widths encompass a spectrum from 500ns to 3 μ s. It is notable that the read margin experiences a notable augmentation as the pulse width is extended. To strike a balance between functionality, performance, and energy considerations, a pulse width of 1 μ s (depicted by the green curve in Fig. 7) has been selected for both Set and Reset operations. This choice aligns with the compromise that will be further enhanced in subsequent sections.

Similarly, Figure 8 provides an overview of the measured read current characteristics on the 128kb memory module under ambient temperature conditions, encompassing a range of CL_{RATIO} settings spanning from 0.35 to 0.50. RM experiences

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

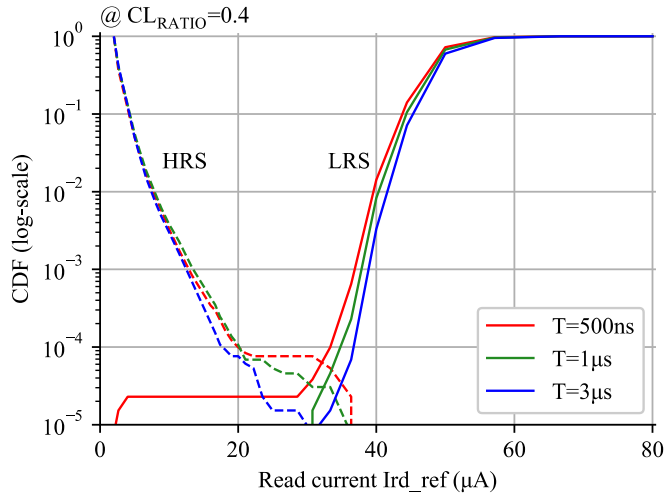


Fig. 7. Writing effect on Read Margin: read current statistics evolution for pulse widths at room temperature.

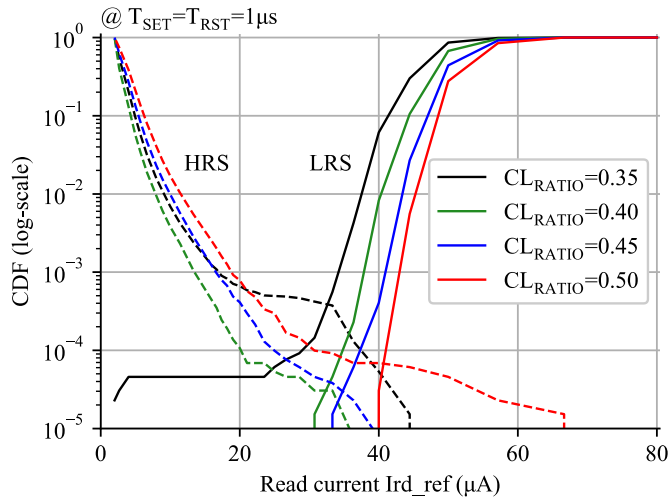


Fig. 8. Current limitation effect on Read Margin: read current evolution for different CL_{RATIO} at room temperature.

degradation, plunging into negative values at a BER of 10^{-4} for a CL_{RATIO} of 0.35, while displaying notably improved performance for CL_{RATIO} of 0.40 and 0.45. However, it is essential to consider that higher CL_{RATIO} introduces adverse effects on energy consumption and aging, leading to the selection of the optimal trade-off point at 0.40 for CL_{RATIO} .

Moreover, the integration of the write verify feature facilitates a reduction in programming voltage by 10% and 18% for Set (V_{SET}) and Reset (V_{RESET}) operations, respectively, as shown in Figure 9. Besides, keeping V_{SET} and V_{RESET} at their nominal value, the average durations for Set and Reset operations experience a significant reduction from 120ns down to 40ns. It's important to emphasize that the priority of the SWA optimization is geared towards enhancing reliability, even if it entails a marginal increase in energy consumption and programming durations.

Read operation takes 40ns, corresponding to a RISC-V request cycle of 20ns, aligned with an RRAM macro access time of 20ns. These timings mirror the designs outlined in [13-14],

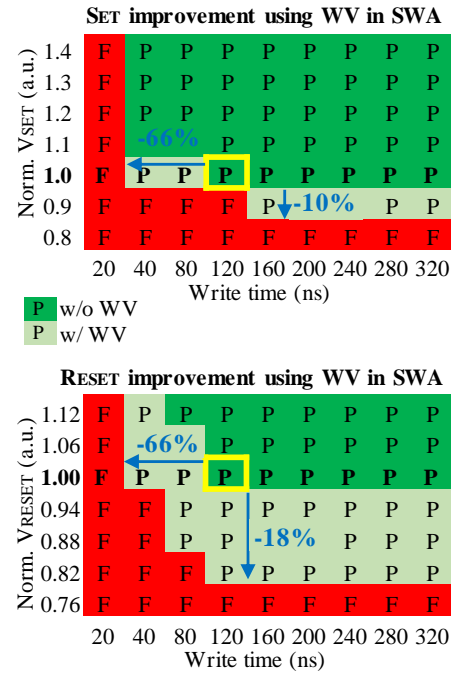


Fig. 9. Shmoo plots for write operations w/ & w/o WV (write verify) at room temperature.

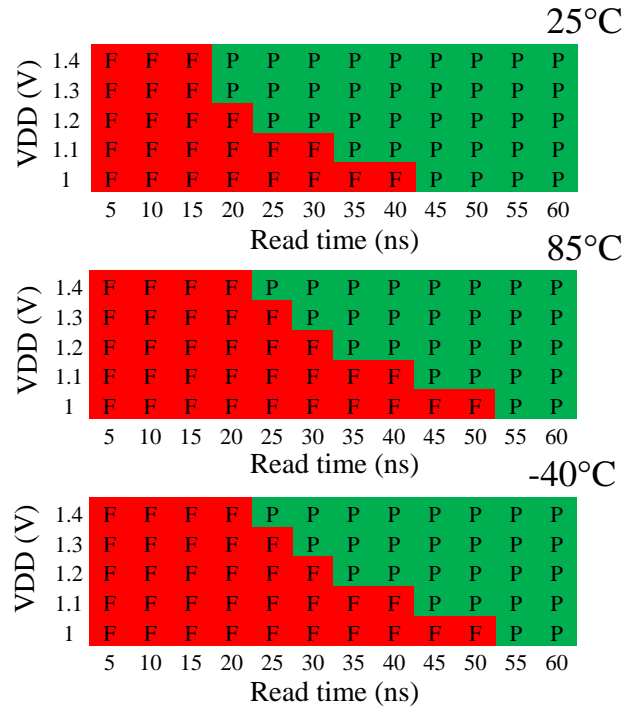


Fig. 10. Shmoo plot for read operations at ambient, 85°C and -40°C temperatures.

which leverage a more advanced technology node. It's noteworthy that the read access time displays a range between 20ns and 45ns as the supply voltages decrease from 1.3V to 1V, as illustrated in Figure 10. In light of this, it is important to mention that ECC activation does not contribute to an improvement in this shmoo representation, since the ECC

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

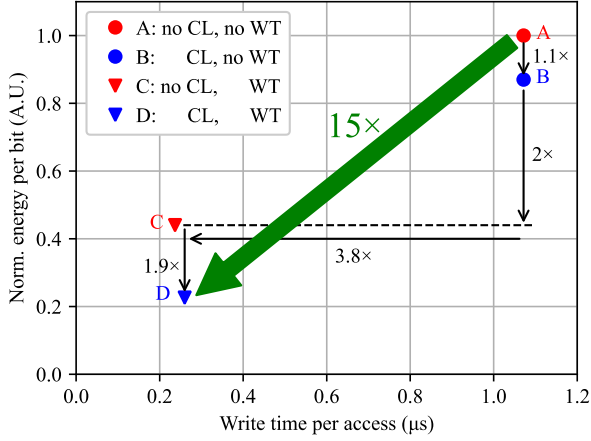


Fig. 11. Write time and energy improvements thanks to WT and CL at room temperature.

duration is around 3.5ns, which is smaller than the granularity of the x-axis, set at 5ns.

The variation of energy over time during a Set operation is presented in Figure 11 across different configurations of CL and WT. When the optimal CL_{RATIO} is selected, there is an enhancement of 10% in energy reduction. An additional 2x

TABLE II
Main Benefits of Design Assist Techniques

Design techniques	Energy	Access time	Additional benefits
RBW	-47%	+4% WR	NA
WT	-56%	-74% WR	NA
CL	-13%	<1% WR	NA
WV	+5%	+6% WR	RM x2
ECC	+38%	+18% RD	BER/100

reduction is observed when only WT is activated, and a further 1.9x reduction is achieved when both WT and optimized CL are combined. While CL significantly contributes to energy reduction, WT offers the advantage of reducing both energy consumption and write time. The optimal combination of CL and WT results in an impressive 15x saving in Energy Delay Product.

Moreover, CL prevents excessive energy allocation to individual bitcells, thus averting the risk for a RRAM device to become stuck. On the other hand, WT ensures equitable programming energy distribution to all bitcells, leading to a much narrower distribution of LRS and HRS. Consequently, this configuration enhances reliability throughout cycling operations. Table II summarizes the experimental data of benefits of each design assist technique independently from the other ones. For the sake of clarity, the improvements focus on energy, access time, RM and BER. Endurance is not analyzed here, and will be discussed in section IV. B.

RBW effectively achieves a reduction of 47% in energy consumption, at the cost of a marginal increase of 4% in write access time. This enhancement is particularly significant as RBW prevents approximately half of the bits from undergoing overwriting, on average. Although the read energy is only a fraction of the write energy, accounting for a mere 3%, RBW's

efficacy lies in its ability to stave off unnecessary overwrites for a substantial portion of the bits.

WT presents a notable energy efficiency improvement, contributing to a reduction of 56% in energy consumption, and concurrently, a 74% reduction in write access time. These improvements are visible on the vertical and horizontal transitions in Fig. 11, as the system moves from points A to C.

In contrast, CL achieves an energy saving of 13% without influencing the write access time, as evidenced by the transition from point A to point B. Employing a single CL_{RATIO} for each data IO slice (comprising 16 bitcell columns) proves more effective in mitigating cell-to-cell variabilities compared to configurations where each bitcell selector restricts current.

WV introduces a trade-off, resulting in a 5% increase in energy consumption and a 6% increase in write access time in average, in favor of doubling the Read Margin (RM). The additional 5% energy expenditure accounts for the verification of all bits once (considering the read operation as representing 3% of write energy) and the need to rewrite 2% of bits, on average, following the verification phase. Similarly, the 6% expansion in access time stems from a 4% increase in read time, along with an additional 2% dedicated to rewriting the bits.

Incorporating ECC corresponds to 38% extra energy mainly due to check bits writing. It introduces minimal write energy overhead for encoder and decoder while contributing to an 18% increase in read time. This enhancement, however, significantly improves the Bit Error Rate (BER) by more than 100 times, transitioning from 10^{-5} to less than $3 \cdot 10^{-8}$, as demonstrated later in endurance testing. The supplementary 18% read time corresponds to 20ns, excluding ECC decoding time, and extends to 23.5ns when considering ECC decoding time.

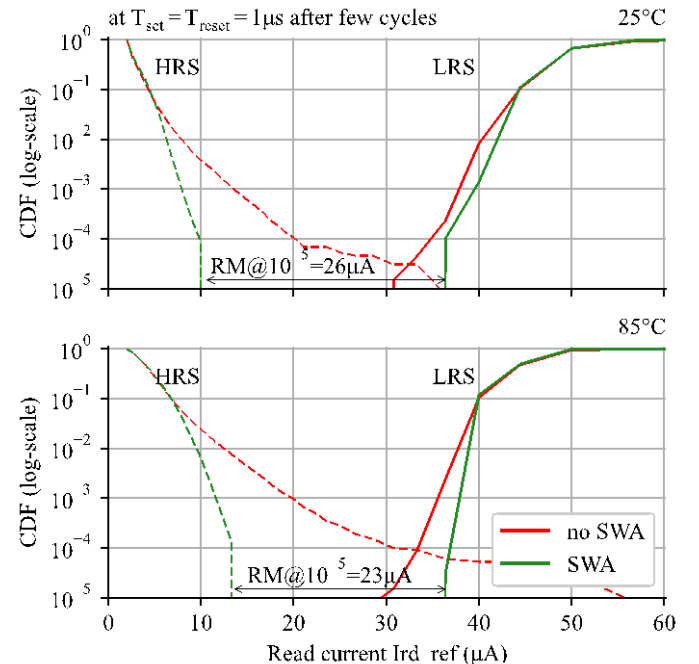


Fig. 12. Memory windows improvements achieving 26 μ A and 23 μ A margin at 25 $^{\circ}$ C and 85 $^{\circ}$ C, respectively.

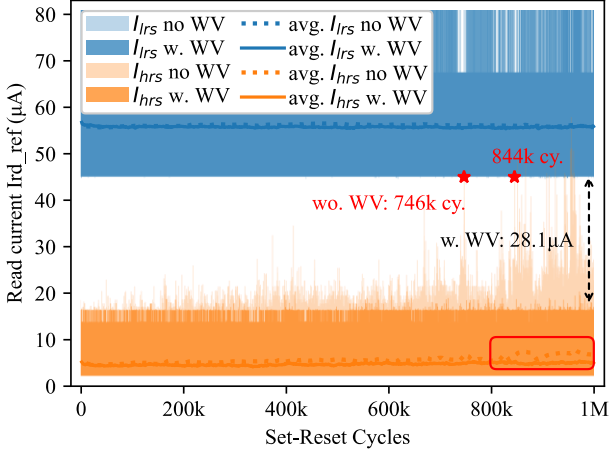


Fig. 13. Read current of the same 32-bit word over 1M Set-Reset cycles at room temperature.

B. Improvements of all Combined Design Techniques

Following the SWA optimizations, the integration of RBW, WT, and ECC is enabled, while CL_{RATIO} is fixed at 0.40. Additionally, three iterations of WV were incorporated. The impact of these enhancements is depicted in Figure 12, where the intrinsic Read Margin (RM) and its augmentation resulting from SWA optimization are showcased. Specifically, at a temperature of 25°C, the RM experiences a significant boost from 13 μ A at a Bit Error Rate (BER) of 10^{-4} to 26 μ A at a BER of 10^{-5} , without the influence of cycling. These improvements are represented by the red and green dashed lines, respectively, surpassing the outcomes achieved in [12, 15-16].

Moreover, at 85°C, which corresponds to the conditions prevalent in industrial applications, the impact of SWA optimization is even more pronounced. Here, the RM is elevated from 3 μ A at a BER of 10^{-4} to 23 μ A at a BER of 10^{-5} . It's essential to note that without the activation of SWA, the RM closes at 0 μ A for BER values of 3×10^{-5} and 9×10^{-5} , at temperatures of 25°C and 85°C, respectively.

Figure 13 shows the endurance of the RRAM using SWA and particularly WV. Read currents of the same 32-bit word have been measured over 1 million Set-Reset cycles at room temperature. Read current values were obtained with a dichotomy search on reference currents. For both LRS and HRS, filled areas represent the span between minimum and maximum read currents among the 32 bits, while the average for 32 bits is plotted with a continuous line with WV and a dashed line without WV. During the WV process, resistance references were set to two different thresholds for Set and Reset operations to maximize the Read Margin (RM). We show that WV exerts minimal impact on the LRS current (overlapped pale and solid blue). However, it significantly improves the upper tail of the HRS current distribution: without WV a large area of pale orange with increasing worst-case read current closes the RM, while WV effectively pushes all measured HRS currents below 20 μ A. Notably, without WV the RM starts to close sporadically at approximately 750k and 850k cycles (as indicated by the red stars in Figure 13), yet remains largely unchanged with WV over the course of 1 million write cycles.

Without WV, the average value starts to be really degraded from 800k cycles as indicated by the red rectangle of Figure 13.

Figure 14 characterizes the endurance degradation by showing the trend of Read Margin (RM) moving average over 50,000 cycles within the span of 1 million Set-Reset cycles. In the absence of Write Verify (WV), the resulting worst-case RM average initiates at 28 μ A and degrades to 14 μ A towards the end of the cycling sequence. This degradation can be well-fitted with a slope of $-5.7 \mu\text{A/Mcy}$ up to 800k write cycles, followed by a steeper slope of $-49.9 \mu\text{A/Mcy}$ during the last 200k write cycles.

Upon enabling WV, the worst-case RM average experiences a much narrower range, fluctuating between 29.9 μ A and 28.1 μ A over the cycling sequence, translating to a significantly reduced degradation of only 6%. The implementation of WV effectively curtails the RM degradation to less than $-2 \mu\text{A/Mcy}$ across 1 million cycles. Throughout the 32 bits subjected to the million cycles, not a single cycle exhibited three or more weak bits beyond the ECC correction capability. This is notable, as RRAM endurance failures are often attributed to either a resistance window collapse, where resistance settles between HRS and LRS [19], or a dielectric breakdown leading to resistance stuck in LRS [20].

The data presented in Figure 14, supported by the implementation of WV, underscores the fact that optimized programming schemes and controlled programming energy contribute to mitigating the acceleration of degradation. Summarizing these improvements in Table III, the activation of all design assist techniques integrated into SWA leads to a substantial energy enhancement of -83%, a reduction in write access time by -55%, and a doubling of memory windows. These benefits, however, come with a tradeoff of an 18%

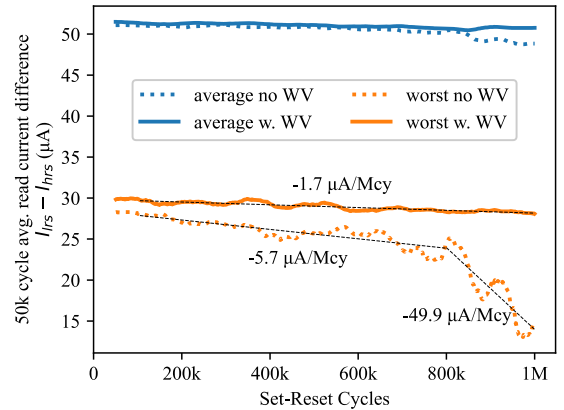


Fig. 14. Read Margin: current difference averaging on 50k cycles over 1M Set-Reset cycles at room temperature.

TABLE III

Main measured results with all design assist techniques

Design techniques	Energy	Access time	RM	BER	Endurance
All enabled	-83%	-55% WR +18% RD	x2	$<10^{-7}$	$>10^6$

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

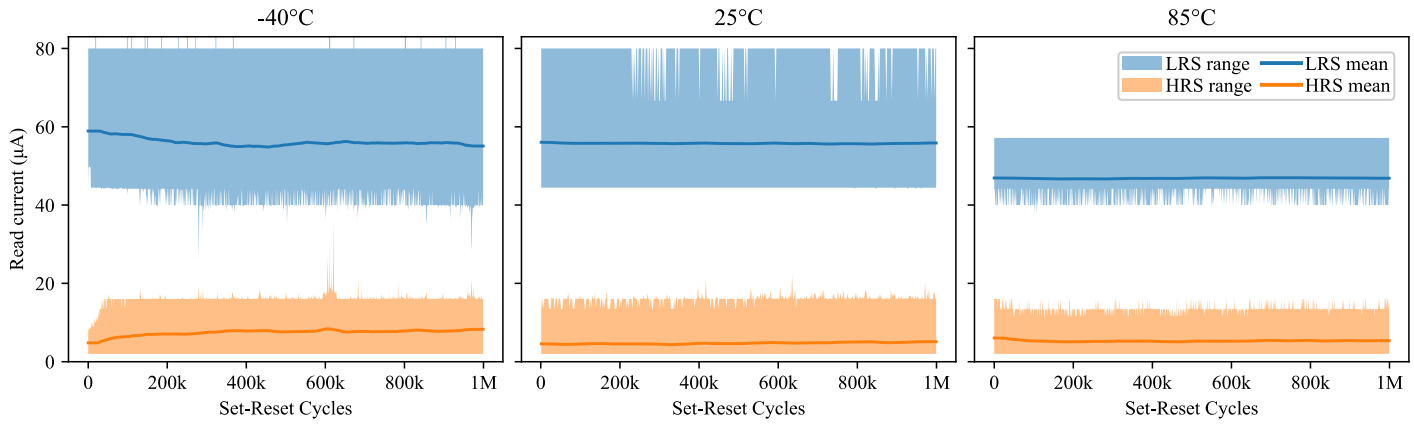


Fig. 15. Read current of the same 32-bit word over 1M Set-Reset cycles at -40°C , room temperature and 85°C with WV enabled.

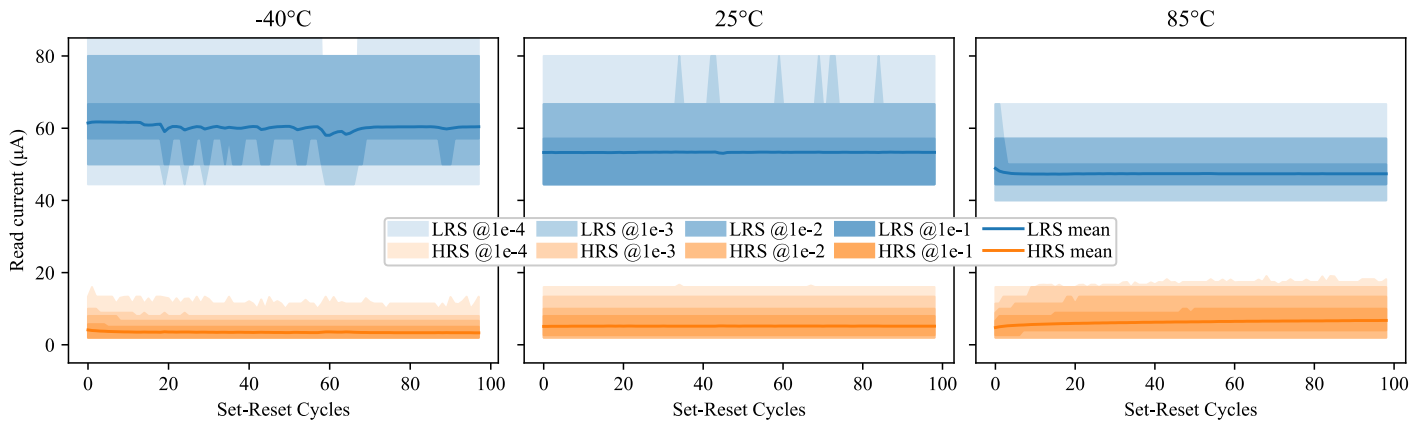


Fig. 16. Read current of the entire memory arrays over 100 Set-Reset cycles at -40°C , room temperature and 85°C with WV enabled.

increase in read access time due to ECC decoding. The combined effect of these optimizations results in a BER below 10^{-7} and endurance exceeding 1 million cycles.

Figure 15 shows the measured endurance of one 32-bit word over 1 Million cycles at 3 different temperatures (-40°C , 25°C and 85°C) while activating WV. The ambient and high temperature cases maintain RM all over the 1M Set-Reset cycles rather constant at $28\mu\text{A}$ and $24\mu\text{A}$, respectively. Besides, applying -40°C increases the average difference between LRS and HRS. At around 600k cycles, several erratic issues in Reset operation prevent HRS current to remain sufficiently low [1]. This phenomenon does not correspond to a major issue, as this is reversible and can be solved thanks to specific programming pattern or design techniques. This can be due to intermittent Random Telegraph Noise [21-22] associated to electrons temporary trapped close to the filament leading to cell current variations.

Figure 16 displays the read current evolution of all the 32-bit words of entire circuit over 100 Set-Reset cycles. In order to manage the huge quantity of data to be plotted, envelops for different quantiles are illustrated. Curves at -40°C display here higher median RM than ambient and high temperature cases, while presenting no failure at all. One must note that while low temperature corresponds to the initial best RM (Fig. 16), it also degrades the RM by $10\mu\text{A}$ with increasing cycle number (Fig.

15). Improvements for variations over temperature were implemented in newer generations of this circuit.

A noteworthy observation is the presence of a relaxation phenomenon associated with the sudden activation of Write Termination (WT) as soon as the filament is established. This issue, illustrated in Figure 17, corresponds to a time-induced degradation of the RRAM filament. The effect creates an increase of the resistance value in LRS over time, leading to a

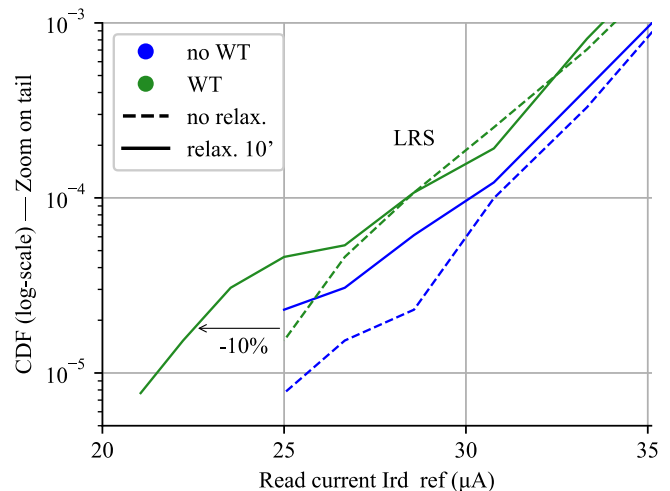


Fig. 17. LRS relaxation phenomenon on read current statistics with and without WT at room temperature.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

reduction of read current by over 10% within just 10 minutes (indicated by the green curves). Such relaxation is not observed in HRS distribution after Reset. This phenomenon aligns with findings from prior research [23-24] and highlights the impact of WT on RRAM performance over extended timeframes.

Table IV presents a comprehensive comparison between our current work and previous studies. An essential distinction lies in the fact that our system pioneers the integration of an on-chip programmable SWA, consolidating all endorsed design techniques into a unified framework. Notably, the optimization of this SWA benefits from the inclusion of RISC-V. In contrast to prior publications, our ECC exhibits the unique property of inversion invariance, thereby streamlining memory testing and forming operation.

In terms of data retention, we demonstrate that the circuit is qualified for 10-year retention at 85°C. We also previously reported 10-year retention using the same RRAM process at temperatures exceeding 125°C, demonstrated on a 16kb characterization structure [25].

Leveraging the SWA optimization, our RRAM macro attains error-free operation across 1 million cycles, signifying a tenfold increase in endurance compared to earlier benchmarks. While demonstrating robustness against read disturbances, our RRAM macro displays measured Set and Reset times of 0.32μs and 1.13μs, respectively, translating to approximately a ninefold reduction in write time compared to previous works. Noteworthy as well is the confirmed resilience against relaxation effects and temperature fluctuations up to 85°C.

The SWA's implementation translates to a fourfold enhancement in energy efficiency and a noteworthy reduction in average write time by a factor of 3.8, ultimately resulting in a remarkable 15-fold increase in Energy-Delay Product (EDP) efficiency.

For a full array write followed by a read, our proposed SWA is able to expand memory windows to 26μA, all while maintaining a commendable BER below 10^{-5} without ECC. This stands as a notable advancement, nearly twofold above the best-published works. Moreover, as the temperature rises beyond 80°C, the SWA's efficiency yields a substantial enhancement to 23μA at a 10^{-5} BER level – representing an impressive improvement of over 60%.

When subjected to a rigorous stress test of one million cycles on 32-bit words, the SWA showcases a remarkable BER of below 10^{-7} after ECC, marking a substantial improvement from the existing state-of-the-art. Moreover, when employing a conservative SWA tuning strategy, the memory windows are significantly augmented, reaching up to 28μA at a 10^{-7} BER level. This conservative SWA involves the application of more stringent conditions on Write Verify (WV), achieved by adjusting two different reference resistances for Set and Reset operations, unlike LP SWA, which uses only one reference resistance. This optimization results in the effective elimination of critical distribution tails, leading to the expansion of memory windows and consequently, improved overall performance.

V. CONCLUSION

This paper presents an efficient 128kb RRAM macro that enables a quantitative evaluation of the distinct and collective advantages conferred by various programming techniques. Our testchip incorporates assist techniques including Current Limiter, Write Termination, Read-Before-Write, Write Verify, and an inversion-invariant Error Correction Code, all orchestrated through the versatility of a flexible smart programming algorithm, SWA, seamlessly integrated with an on-chip RISC-V core. This innovative approach effectively mitigates the intrinsic variability inherent to RRAM technology, demonstrating robust retention even at

TABLE IV
Comparison with RRAM prior work.

		ISSCC'18 [12]	ISSCC'19 [13]	VLSI'20 [14]	VLSI'20 [15]	ISSCC'21 [16]	This work	
Embedded design techniques	WT	No	Yes	Yes	Yes	Yes	Yes	
	CL	Yes	Yes	Yes	Yes	No	Yes	
	ECC	No	No	No	No	No	Inversion invariant	
	Programing algorithm	No	Yes	No	No	No	Yes	
	System integration	No	No	No	No	No	RISC-V	
Device perf.	Technology (nm)	40	22	22	28	14	130	
	10 yr retention (°C)	85	-	85	85	85	85	
	Write Endurance	>1k	-	10k	10k	~ 100k	1M [†]	
Circuit performances at macro level	Min. acc. time (ns)		9	<10	10	3.3	-	20
	Read	Disturb immunity (<0.45V)	Yes	Yes	-	Yes	Yes	Yes
		max. V _{READ} (V)	0.26	0.45	0.6	0.34	0.2	0.4
	Write time	Set (μs)	-	< 10	-	-	-	0.32
		Reset (μs)	-	< 10	-	-	-	1.13
	RM Memory windows (μA)	short term (1cy)@25°C	14@10 ⁻⁵	-	10@10 ^{-5*}	30@10 ^{-2*}	-	26 @10 ⁻⁵ w/ LP SWA
		short term (1cy)@>80°C	14@10 ⁻⁵	-	10@10 ^{-5*}	-	-	23 @10 ⁻⁵ w/ LP SWA
		long term (1Mcy)@25°C with ECC	-	-	-	30@10 ^{-2*}	-	11 @10 ⁻⁷ w/ LP SWA 28 @10 ⁻⁷ w/ conservative SWA

* Estimated from publication

†@25°C

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

temperatures up to 85°C. Each design assist technique exerts a unique impact on critical metrics such as energy consumption, access time, Bit Error Rate (BER), and read margin. By combining these techniques, we achieve an impressive energy improvement of -83%, reduce write access time by -55%, and double memory windows. Furthermore, our system demonstrates exceptional endurance, enduring 1 million cycles without any read errors in 32-bit word operations, resulting in a BER well below 10^{-7} . Notably, the read margin reaches 28 μ A over 1 million cycles, or approximately 2 times higher for single-cycle access at both ambient (26 μ A) and elevated (85°C, 23 μ A) temperatures. This read margin is comparable to, and even outperforms, previously reported works that have a BER of only 10^{-2} .

These collective achievements position our work as a promising contender for recent industrial RRAM adoption. Our approach paves the way for enhanced memory technology, setting new benchmarks for energy efficiency, reliability, and endurance in the context of advanced memory systems.

REFERENCES

- [1] C. Peters, F. Adler, K. Hofmann and J. Otterstedt, "Reliability of 28nm embedded RRAM for consumer and industrial products," IMW 2022.
- [2] G. Molas and E. Nowak, "Advances in Emerging Memory Technologies: From Data Storage to Artificial Intelligence," Appl. Sci. 2021, 11, 11254.
- [3] A. Grossi, C. Zambelli, P. Olivo, E. Nowak, G. Molas, J.-F. Nodin and Luca Perniola, "Cell-to-Cell Fundamental Variability Limits Investigation in OxRRAM Arrays," in Electron Device Letters, 2018.
- [4] L. Grenouillet, N. Castellani, A. Persico, V. Meli, S. Martin, O. Billoint, R. Segaud, S. Bernasconi, C. Pellissier, C. Jahan, C. Charpin-Nicolle, P. Dezest, C. Carabasse, P. Besombes, S. Ricavy, N.-P. Tran, A. Magalhaes-Lucas, A. Roman, C. Boixaderas, T. Magis, M. Bedjaoui, M. Tessaire, A. Seignard, F. Mazen, S. Landis, E. Vianello, G. Molas, F. Gaillard, J. Arcamone, E. Nowak, "16kbit 1T1R OxRAM arrays embedded in 28nm FDSOI technology demonstrating low BER, high endurance, and compatibility with core logic transistors," IMW 2021.
- [5] J. Sandrini, L. Grenouillet, V. Meli, N. Castellani, I. Hammad, S. Bernasconi, F. Aussenac, S. Van Duijn, G. Audoit, M. Barlas, J.F. Nodin, O. Billoint, G. Molas, R. Fournel, E. Nowak, F. Gaillard and C. Cagli, "OxRAM for embedded solutions on advanced node: scaling perspectives considering statistical reliability and design constraints," 2019 IEEE International Electron Devices Meeting (IEDM), 2019, pp. 30.5.1-30.5.4.
- [6] C. Nail, G. Molas, P. Blaise, G. Piccolboni, B. Sklenard, C. Cagli, M. Bernard, A. Roule, M. Azzaz, E. Vianello, C. Carabasse, R. Berthier, D. Cooper, C. Pellissier, T. Magis, G. Ghibaudo, C. Vallée, D. Bedeau, O. Mosendz, B. D. Salvo and L. Perniola, "Understanding RRAM endurance, retention and window margin trade-off using experimental results and simulations," 2016 IEEE International Electron Devices Meeting (IEDM), 2016, pp. 4.5.1-4.5.4.
- [7] EENews Europe, "Infineon's microcontrollers to go RRAM with TSMC", eeneurope.com. <https://www.eeneurope.com/en/infineon-microcontrollers-to-go-rram-with-tsmc/> (accessed Feb. 5, 2023).
- [8] Design & Reuse, "eMemory and UMC Bring New ReRAM Intellectual Property to Market", design-reuse.com. <https://www.design-reuse.com/news/50886/ememory-umc-reram-intellectual-property.html> (accessed Feb. 5, 2023).
- [9] X. Xue, W. Jian, J. Yang, F. Xiao, G. Chen, S. Xu, Y. Xie, Y. Lin, R. Huang, Q. Zou and J. Wu, "A 0.13 μ m 8 Mb Logic-Based Cu xSi yO ReRAM With Self-Adaptive Operation for Yield Enhancement and Power Reduction", IEEE Journal of Solid-State Circuits (Volume: 48, Issue: 5, May 2013).
- [10] X. Xue, J. Yang, Y. Zhang, M. Wang, H. Lv, X. Zeng, M. Liu, "A 28nm 512Kb adjacent 2T2R RRAM PUF with interleaved cell mirroring and self-adaptive splitting for extremely low bit error rate of cryptographic key", IEEE Asian Solid-State Circuits Conference (A-SSCC), 2019, pp. 29-32.
- [11] M.-F. Chang, J.-J. Wu, T.-F. Chien, Y.-C. Liu, T.-C. Yang, W.-C. Shen, Y.-C. King, C.-J. Lin, K.-F. Lin, Y.-D. Chih, S. Natarajan, J. Chang, "Embedded 1Mb ReRAM in 28nm CMOS with 0.27-to-1V Read Using Swing-Sample-and-Couple Sense Amplifier and Self-Boost-Write-Termination Scheme", IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014.
- [12] C.-C. Chou, Z.-J. Lin, P.-L. Tseng, C.-F. Li, C.-Y. Chang, W.-C. Chen, Y.-D. Chih, T.-Y. J. Chang, "An N40 256K \times 44 embedded RRAM macro with SL-precharge SA and low-voltage current limiter to improve read and write performance," ISSCC, pp. 478-480, 2018.
- [13] P. Jain, U. Arslan, M. Sekhar, B. C. Lin, L. Wei, T. Sahu, J. Alzate-vinasco, A. Vangapaty, M. Meterelilyoz, N. Strutt, A. B. Chen, P. Hentges, P. A. Quintero, C. Connor, O. Golonzka, K. Fischer and F. Hamzaoglu, "A 3.6 Mb 10.1Mb/mm² Embedded Non-Volatile ReRAM Macro in 22nm FinFET Technology with Adaptive Forming/Set/Reset Schemes Yielding Down to 0.5V with Sensing Time of 5ns at 0.7V," ISSCC, pp. 212-213, 2019.
- [14] C.-C. Chou, Z.-J. Lin, C.-A. Lai, C.-I. Su, P.-L. Tseng, W.-C. Chen, W.-C. Tsai, W.-T. Chu, T.-C. Ong, H. Chuang, Y.-D. Chih and T.-Y. J. Chang, "A 22nm 96KX144 RRAM Macro with a Self-Tracking Reference and a Low Ripple Charge Pump to Achieve a Configurable Read Window and a Wide Operating Voltage Range," Symp. VLSI Circuits, 2020.
- [15] J. Yang, X. Xue, X. Xu, H. Lv, F. Zhang, X. Zeng, M.-F. Chang and M. Liu, "A 28nm 1.5Mb Embedded 1T2R RRAM with 14.8 Mb/mm² using Sneaking Current Suppression and Compensation Techniques," Symp. VLSI Circuits, 2020.
- [16] J. Yang, X. Xue, X. Xu, Q. Wang, H. Jiang, J. Yu, D. Dong, F. Zhang, H. Lv and M. Liu, "A 14nm-FinFET 1Mb Embedded 1T1R RRAM with a 0.022 μ m² Cell Size Using Self-Adaptive Delayed Termination and Multi-Cell Reference," ISSCC, pp. 336-338, 2021.
- [17] B. Giraud, S. Ricavy, Y. Moursy, C. Laffond, I. Sever, V. Gherman, M. Pezzin, F. Lepin, M. Diallo, K. Zenati, S. Dumas, M. Vershkov, A. Bricalli, G. Piccolboni, J.-P. Noel, A. Samir, G. Pillonnet, Y. Thonnart and G. Molas, "Benefits of Design Assist Techniques on Performances and Reliability of a RRAM Macro," IMW 2023.
- [18] V. Gherman, S. Evain and B. Giraud, "Binary Linear ECCs Optimized for Bit Inversion in Memories with Asymmetric Error Probabilities," Design, Automation and Test in Europe (DATE), 2020.
- [19] S. Balatti, S. Ambrogio, Z. Wang, S. Sills, A. Calderoni, N. Ramaswamy and Daniele Ielmini, "Voltage-Controlled Cycling Endurance of HfOx-Based Resistive-Switching Memory," IEEE Transactions on Electron Devices, vol. 62, no. 10, pp. 3365-3372, Oct. 2015.
- [20] D. Alfaro Robayo, G. Sassine, Q. Raffay, G. Ghibaudo, G. Molas and E. Nowak, "Endurance Statistical Behavior of Resistive Memories Based on Experimental and Theoretical Investigation," IEEE Transactions on Electron Devices, vol. 66, no. 8, pp. 3318-3325, Aug. 2019.
- [21] S. Ambrogio, S. Balatti, V. McCaffrey, D. C. Wang and D. Ielmini, "Noise-Induced Resistance Broadening in Resistive Switching Memory—Part II: Array Statistics," in IEEE Transactions on Electron Devices, vol. 62, no. 11, pp. 3812-3819, Nov. 2015.
- [22] L. Reganaz et al., "Investigation of resistance fluctuations in ReRAM: physical origin, temporal dependence and impact on memory reliability," International Reliability Physics Symposium (IRPS), 2023.
- [23] A. Fantini, G. Gorine, R. Degraeve, L. Goux, C.Y. Chen, A. Redolfi, S. Clima, A. Cabrini, G. Torelli and M. Jurczak, "Intrinsic Program Instability in HfO₂ RRAM and consequences on program algorithms," IEDM 2015.
- [24] E. Esmanhotto, L. Brunet, N. Castellani, D. Bonnet, T. Dalgaty, L. Grenouillet, D. R. B. Ly, C. Cagli, C. Vizioz, N. Allouti, F. Laulagnet, O. Gully, N. Bernard-Henriques, M. Bocquet, G. Molas, P. Vivet, D. Querlioz, JM. Portal, S. Mitra, F. Andrieu, C. Fenouillet-Beranger, E. Nowak and E. Vianello, "High-Density 3D Monolithically Integrated Multiple 1T1R Multi-Level-Cell for Neural Networks," IEDM 2020.
- [25] G. Molas, G. Piccolboni, A. Bricalli, A. Verdy, I. Naot, Y. Cohen, A. Regev, I. Naveh, D. Deleruyelle, Q. Raffay, N. Castellani, L. Reganaz, A. Persico, R. Segaud, J. F. Nodin, V. Meli, S. Martin, F. Andrieu and L. Grenouillet, "High temperature stability embedded ReRAM for 2x nm node and beyond," IMW 2022.



Bastien Giraud was born in La Seyne sur Mer, France in 1982. He received the M.S. and Ph.D. degrees from Polytech Marseille and Telecom ParisTech France, in 2005 and 2008, respectively. The PhD thesis focused on SRAM design in FDSOI.

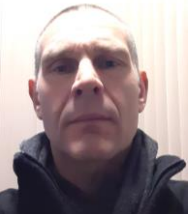
> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

In 2009, he was postdoctoral researcher at UC Berkeley working on low power circuits and SRAM variability. From 2010, he works at CEA as a circuit designer specialized in memory and low power circuit in advanced technologies. His research interests include SRAM and CAM design, in memory computing, using advanced CMOS technologies and non-volatile RRAM technologies such as OxRAM, FeRAM and PCM. He has published more than 50 papers in international conferences and journals. He is the main author of a book chapter and the main inventor or co-inventor of 25 patents.



Sebastien Ricavy was born in Marseille in 1977. He received the M.Sc. degree in electrical engineering and computer sciences fields in 2001 from Grenoble Alpes University, Grenoble, France.

He started his career focusing on design of standard cells and memories on SOI technologies (PD/FDSOI technologies) at SOISIC SA. From 2006 to 2018, he worked at ARM LTD where he was involved in SRAM, register files or VROM design activities as Senior Memory Technical Lead role or Memory Team Manager. Since 2019, he has been working as Memory Research Engineer at CEA, Grenoble, France. His research activities include various type of non-volatile technologies such as OxRAM, FeRAM or PCM designs.



Cyrille Laffond was born in Voiron, France in 1975. He graduated from ENSERG France in 1999.

From 2000, he started his career in a subcon company working for STMicroelectronics that he joined in 2005. He was technical leader on several mixed-signal SOC designs until 2020. Since this year, he is with CEA as a principal engineer responsible for digital design and architecture as well as top verification.



Ilan Sever, received the B.Sc.E.E. degree from the Technion—Israel Institute of Technology. He has over 25 years of design and project-management expertise in the field of semiconductor IP and SOC design with deep expertise in volatile and non-volatile memory design. He currently serves as the Vice President of research and

development at Weebit-Nano, an Israeli startup in the field of resistive memory (ReRAM). Prior to joining Weebit-Nano, he spent 11 years as the Memory-Group CTO and Subsidiary GM at French Semiconductors Company Dolphin Design, leading the development of numerous memory architectures, as well as innovation projects in the fields of security, rad-hard memories, and RISC-V memory subsystems. He previously managed the VLSI activities of Sandlinks Systems, a company in the field of IoT, developing a breakthrough Low-Power, Mixed-Signal/RF SOC. In his previous roles, Mr. Sever was Director of IP & Libraries at Tower Semiconductors and Design Manager of Flash-Memory at ST Microelectronics. He has several granted patents and awards.

Valentin Gherman received the master degree in physics from the University of Siegen, Germany, in 2000, and the PhD degree in the field of built-in self-test from the University of Stuttgart, Germany. Since 2006, he is with CEA. As expert in error masking and error correction, he was involved in the specification and development of reliability solutions for storage and memory systems. He is author of seventeen patents and several scientific papers. His main research interests include: design for test; in-situ monitoring; error-control coding; logic optimization.



Florent Lepin was born in Tassin-la-Demi-Lune, France in 1981. He received the engineering degree in electronic and industrial IT engineering from Polytech' Grenoble, France (U. Joseph Fourier's Engineering School) in 2007.

In 2012, he joined CEA where he worked on digital mixed integrated circuit designer. Currently he is Head of Digital Low power Architecture & IC Design Lab at CEA List. His research incorporates asic architecture and software for power-efficient computing focused on memory systems.



Mariam Diallo was born in Casablanca, Morocco in 1984. In 2008, she received an engineering degree from the French University in Cairo, Egypt in communication and information technology and a master's degree from Paul Sabatier University in Toulouse, France in the field of microwave, electromagnetism and optoelectronics.

In 2011, she started her career as quality assurance expert of place and route software in Mentor Graphics Rabat Morocco (now Siemens EDA). In 2018, she joined the department of numerical systems and integrated circuits in CEA Grenoble, France as asic physical design engineer.



Khadija Zenati received the engineering degree in electronics from the National Engineering School of Sfax in Tunisia and a Master's degree in the integrated electronic systems from the University of Montpellier in France in 2019. She is now a digital design and verification engineer at the department of numerical systems and

integrated circuits in CEA Grenoble, France.



Sylvain Dumas was born in Aubenas, France in 1977. He received the engineering degree in electronic and IT engineering from Polytech'Grenoble France (U. Joseph Fourier's Engineering School) in 2000.

In 2000, he started his career in power IC test laboratory in STMicroelectronics.

In 2005, he joined CEA, Grenoble, France. Currently, he is IC test engineer at the Department of Systems. He has co-authored several publications.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <



Olivier Guille was born in Grenoble, France, in 1978. After a M.S. in Embedded Systems Electronics from INPG-ESISAR school (2001), he worked for many companies like Thales, Texas Instruments, EASii-IC, in the domains of firmware and test software for microelectronics or embedded systems. Joining CEA-LETI test laboratory in 2019, he has been developing test platforms for ASIC validation.



Maxim Vershkov was born in Bila Tzerkva, Ukraine in 1986. He received B.S.c degree in electric and computer engineering, with microelectronic and vlsi specialization, from Ben-Gurion university of Negev, Beer Seva, Israel in 2018. He joined Dolphin Design, as circuit design engineer, working on SRAM memories (2018-2020). In 2021 he joined Weebit-Nano LTD, as circuit design engineer, working on design and mixed-signal and analog verification of RRAM memory circuit.



Alessandro Bricalli received the B.S. and M.S. degrees in engineering physics and the Ph.D. degree in information technology from Politecnico di Milano, Milan, Italy in 2013, 2015 and 2019, respectively. In 2019, he joined Weebit Nano (Hod Hasharon, Israel) as R&D engineer. He works on the development of Weebit's ReRAM in the laboratories of CEA-Leti, Grenoble, France.



Giuseppe Piccolboni was born in Verona, Italy on the 19-08-1987. He attended Politecnico di Milano (Italy) and graduated with summa cum laude in physics engineering both bachelor (2009) and master degree (2012). After one year as a filled engineer for Schlumberger, he did a PhD on vertical resistive ram in CEA-Leti during which we participated to several conferences (IMW 2016, SSDM 2016, IEDM 2015). In 2016 he joined weebit-nano as a R&D engineer in Grenoble, France. His main job consists in running electrical tests and doing data analysis.



Jean-Philippe Noel was born in Alès, France, in 1985. He received the M.S. and Ph.D. degrees in microelectronics from Polytech Marseille, France, and the University Grenoble Alpes, France, in 2008 and 2011, respectively. During his Ph.D. studies, he worked at the CEA, Grenoble, France, on the design optimization of low power digital and memory circuits in UTBB FD-SOI technology. In November 2011, he joined STMicroelectronics, Crolles, France, as a memory design engineer. His work was focused on the design of innovative

SRAM and TCAM circuits in UTBB FD-SOI technology. Since December 2016, he has joined CEA, Grenoble, France, to work on innovative in- and near-memory computing architectures and emerging non-volatile memories (ReRAM, PCM ...) circuit design. He has authored and co-authored over 60 publications and 30 patents.



Anass Samir, Anass Samir received the M.S and Ph.D. degrees in microelectronics from Polytech Marseille and Aix-Marseille University, France, in 2008 and 2011, respectively. His doctoral research was on the on-chip design optimization of energy management for low power microcontroller applications, work realized at STMicroelectronics, Rousset, France.

In 2011, he joined STMicroelectronics, Rousset, France, as an Analog/IO Design Engineer Architect, where he works on the research and the development of low-power analog solutions for standby/shutdown modes and USB architectures (USB OTG battery charging, USB power delivery Type C) for the STM32 microcontroller.

Since December 2021, he joined CEA-Leti, Grenoble, France, as an IC Design Research Engineer in advanced technologies specialized in innovative emerging non-volatile memories and low-power circuit in advanced technologies.

He takes part in some publications in international journals and conference proceeding and in several patents on analog domains.



Gaël Pillonnet (M'05, SM'16) was born in Lyon, France, in 1981. He received his Master's degree in Electrical Engineering from CPE Lyon, France, in 2004, a PhD and habilitation degrees from INSA Lyon, France in 2007 and 2016, respectively. Following an early experience as analog designer in STMicroelectronics in 2008, he joined the University of Lyon as associate Professor. During the 2011-12 academic year, he held a visiting researcher position at the University of California at Berkeley. Since 2013, he is in CEA-Leti, Grenoble, France, involved in developing various projects in design technology co-optimization. During the 2022-2023 academic year, he joins energy-efficient microsystems group at UCSD as a visiting researcher. He is now the scientific advisor for the Silicon component division in CEA-Leti. His research focuses on power-conversion oriented circuits such as DC-DC converters, audio amplifiers, adiabatic logics, electromechanical transducers, harvesting electrical interfaces. He is also involved in various device/circuit design enablement for RRAM, Si-qubit, solid-state energy storage and MEMS sensors.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <



Yvain Thonnart, (M'16) received the MS degree from Ecole Polytechnique and an engineering diploma from Telecom Paris, France in 2005. He then joined the Technological Research Division of CEA, the French French Alternative Energies and Atomic Energy Commission, within the CEA-Leti institute until 2019, then within the CEA-List institute. He is now senior expert on communication and synchronization in systems on chip, and scientific advisor for the mixed-signal design lab. His main research interests include asynchronous logic, networks on chip, physical implementation, emerging technologies integration such as photonics, cryoelectronics and interposers, leading to more than 70 publications and 10 patents. He is currently serving in the technical program committee of the ISSCC.



Gabriel Molas was born in 1979 in Paris, France. He received the B.S. and M.S. in Physics Engineering, with microelectronics specialization (2001), and the Ph.D. degree in micro- and nano-electronics from the Polytechnics Institute of Grenoble, France (2004), with a thesis on few electron memories. In 2021, he passed Habilitation à Diriger des Recherches, école doctorale EEATS (Univ. Grenoble Alpes), allowing to be PhD director. He joined LETI (Laboratory of Electronics, Technology and Instrumentation of CEA, Grenoble) as a research engineer in 2004, working on emerging charge trap memories (2004-2011) and resistive memories (2011-2021), responsible from the LETI side of various industrial and institutional projects. In 2016, he became LETI senior expert on non-volatile memories and in 2021 CEA research director. In 2020 he was also nominated scientific delegate of microelectronic service of LETI. In 2022 he joined Weebit Nano Ltd as Chief Scientific Officer, developing the emerging resistive memory technology. He is author or co-author of more than 160 publications in international conferences (including 23 IEDM and more than 20 invited talks), 2 short courses (IEDM, InMRAM), 1 tutorial, 50 papers in refereed journals, 5 book chapters, 25 patents. He is/was a reviewer of several international journals. He is/was a member of several technical committees (IEDM, IRPS, ESSDERC, IMW) and organizing committees (Fall MRS Memory Symposium, IEEE IMW organizing committee, IEEE ESSDERC Financial chair). In 2019 he became IEEE senior member.