Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
SCIENZE DELL'INGEGNERIA

CICLO XXIX

COORDINATORE Prof. Trillo Stefano

EMERGING NON VOLATILE MEMORIES RELIABILITY

Settore Scientifico Disciplinare ING/INF-01

Dottorando
Grossi Alessandro

Tutore
Prof. Olivo Piero

Correlatore
Dott. Zambelli Cristian

Anni 2014/2016

To Eva, Eleonora, Lorella and Paolo

Abstract

This work presents the results of the research activity performed during the XXIX-th cycle of the Ph.D. school in Engineering Science of Università degli Studi di Ferrara. In particular the thesis focuses on the electrical characterization, physics, modeling and reliability of innovative non-volatile memories, addressing three of the most promising candidates for the floating-gate based memories replacement which are currently facing a technology dead end. The manuscript is organized as follows.

In Chapter 1 planar CT-NAND memory arrays are considered, showing that the main reliability issues affecting such technology are endurance and retention. Enhanced program and read algorithm able to reduce such limitations will be presented and characterized, highlighting the advantages obtained in terms of reliability. After that, the performances of Solid State Drives (SSD) integrating CT-based memories and using the proposed algorithms will be simulated and evaluated.

In Chapter 2 the results obtained on RRAM will be presented and discussed focusing on the variability, which is the main reliability issue of these nonvolatile memories. The impact of Forming, Set and Reset operations on variability will be evaluated, starting from single pulse operations up to program and verify algorithms. The quantum point contact model will be used to give a physical explanation of the results obtained in characterization. After that, the process parameters impact on variability and reliability will be discussed. Finally, the fundamental variability limits of such technologies will be defined through an extensive array characterization and radiation

hard application perspectives will be provided.

In Chapter 3, the results obtained on TAS-MRAM technology will be reported. The reliability and the cell-to-cell variability will be evaluated during endurance tests by extracting a set of characteristic parameters from measurements performed on 1kbits arrays. After a preliminary optimization of the writing parameters on fresh devices, the effectiveness of the optimized parameters will be verified during cycling by evaluating their advantages in terms of cell-to-cell variability and breakdown reduction. After that, a novel TAS-MRAM array with optimized read procedure (Self Referenced) will be tested and compared with the previous one to highlight its advantages in terms of reliability.

The conclusions of this work will be reported at the end of the manuscript outlining what has been accomplished, proposing possible applications for the technologies studied in this thesis and suggesting future works that could extend and improve the understanding of the reliability issues on such memory technologies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Memory devices are nowadays one of the most important electronic component in the semiconductor industry. In modern day life, more and more information is being stored on computers, laptops, tablets, smartphones and more. Today this information is usually stored on Flash memory, which conquered the market in the last decade through applications like digital cameras, SD cards and USB drives and is still the mostly used nonvolatile memory technology. A growth of the memory devices storage capacity without increasing the area occupation is constantly requested by the market: in order to satisfy such requirements, an increase of the memory density and of cell shrinking is mandatory.

Flash technology approached its scaling limits making a significant capacity increase very challenging: to overcome this limitation, the transition from planar to three-dimensional architectures appears today as one of the most viable solution for the integration of non-volatile memory cells in Tera-bit arrays. In this framework, Charge Trap (CT) NAND memory cells are considered as one of the most promising technology for 3D integration because of a better scalability than Floating Gate (FG) NAND [1]. However, the 3D-NAND solution for increased capacity is demonstrating manufacturing difficulties and does not provide any speed, energy efficiency and reliability improvement.

For this reason, several disruptive technologies have been blossoming in the last decades for non-volatile memory applications [2]: an overview of the actual memory technology scenario is provided in Fig. 1.1. The most prominent ones are Phase-Change Memories (PCM), 3D-Xpoint, Resistive RAM (RRAM) which can be divided in Oxide-based Resistive RAM (OxRAM) and Conducting-Bridge RAM (CBRAM), and Magnetic Memories (MRAM) in several fashion such as Toggle, Thermally-Assisted (TAS), Spin-Transfer Torque (STT), and perpendicular Spin-Transfer Torque (p-STT). The storage capabilities of such technologies increased rapidly in the last few years, and it is expected to be able to compete soon with NAND Flash as depicted in Fig. 1.2.

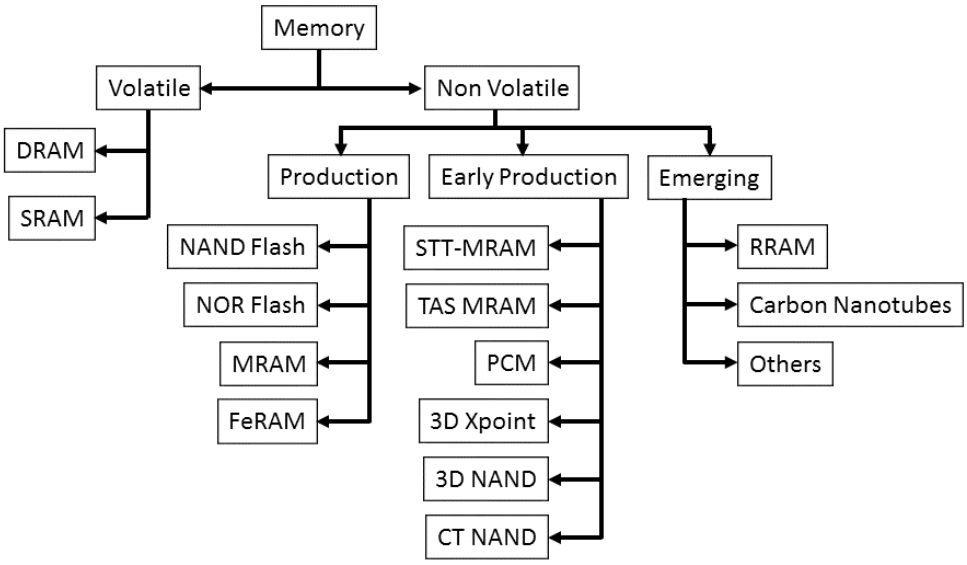


Figure 1.1: Overview of the actual memory technology scenario.

However, even if such emerging technologies offer high scalability, speed and endurance capabilities, several reliability issues still prevent them from reaching a maturity level. Reliability represents one of the major antagonist towards the unstoppable technological evolution of hyperscaled memories, since the correct operations must be assured throughout the entire lifetime. In particular, the ability of keeping unaltered the stored information even

after a consistent number of write operations and for long times must be guaranteed.

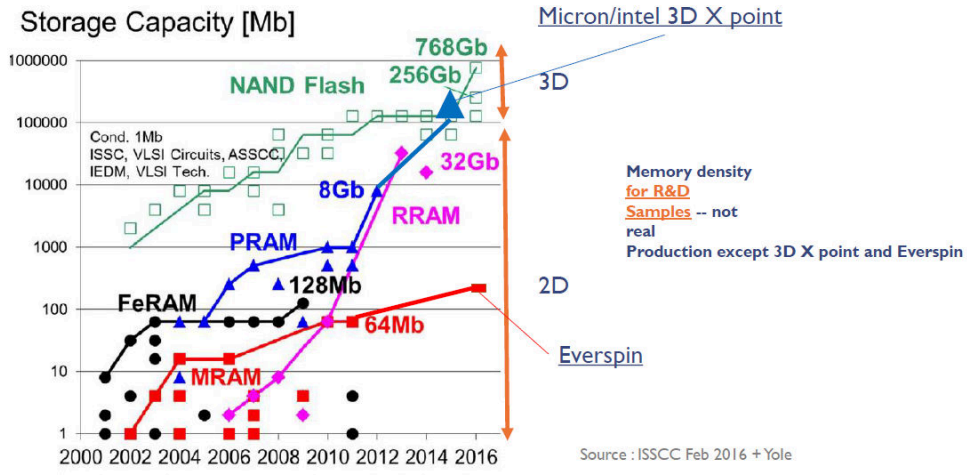


Figure 1.2: Emerging memory technologies storage capacity trend in the last years.

In this thesis, the reliability issues affecting CT NAND, RRAM and TAS-MRAM nonvolatile memory technologies will be explored and discussed in order to understand:

- The physical mechanisms affecting the reliability
- The optimal working conditions
- The maximum performance and reliability achievable when the optimal working conditions are used (write/erase/read speed, endurance, retention, read disturb immunity, etc..)
- The expected system-level performances on different applications such as consumer/enterprise SSD, automotive, space applications.

In order to do that, results obtained trough electrical characterization of memory cells and arrays as well as physical, statistical and system modeling

results will be provided and explained. Several techniques able to counteract the reliability issues will be discussed such as process optimization, read/write parameters and algorithms optimizations. Physical models will be used as a tool to correlate the experimental results with the physical phenomena causing the observed behavior of the devices, while system level models will be used to evaluate the expected system level performances in SSD applications. Finally, the thesis will speculate on promising applications and markets for each considered technology taking into account their advantages and intrinsic limitations.

Chapter 2

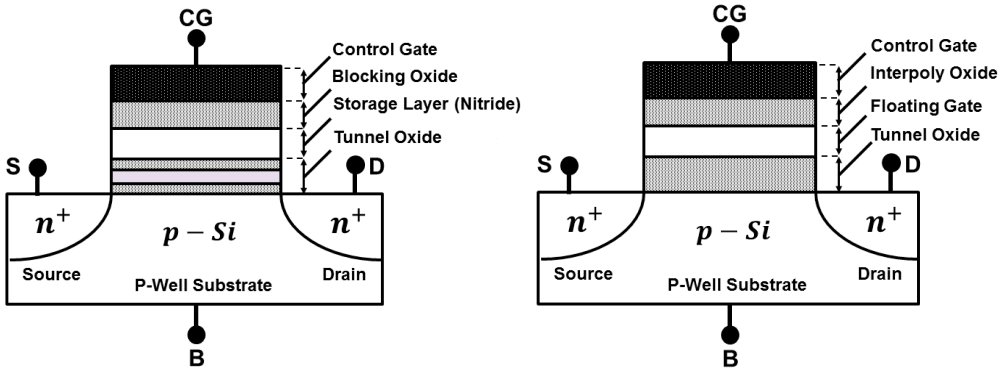
Charge Trap NAND

Charge Trap (CT) NAND memory cells are considered as one of the most promising technology for 3D integration because of a better scalability than Floating Gate (FG) NAND. Despite the high theoretical potentialities demonstrated by CT memories, several reliability issues affect such technology. Even if the transition from 2D to 3D will change the impact of the reliability issues affecting planar devices, they will still be critical. In this chapter, such reliability issues are discussed. After that, enhanced program and read techniques able to reduce their impact are presented and experimentally characterized. The results are then exploited for co-simulations at system level [3–7], assessing reliability and performance perspectives of future Solid State Drives (SSD) integrating CT-based memories.

2.1 Basics

The basic concept of a Charge Trap (CT) NAND memory cell consists of a metal oxide semiconductor device where the Floating Gate (FG) is replaced by an insulating charge trapping layer [1]. Such storage layer, typically made of silicon nitride, is isolated by means of a tunnel oxide and a blocking oxide as sketched in Fig. 2.1 where the FG cell structure is reported for comparison. The tunnel oxide plays a basic role for the control of the device threshold

voltage, whose value represents, from a physical point of view, the stored information. The blocking oxide prevents electrons from passing to/from the control gate. Electrons transferred into the storage layer give a threshold voltage variation. In quiescent conditions, thanks to the two oxides, the stored charge is supposed not to leak away, thus granting the nonvolatile paradigm fulfillment. Oxides are available in different materials depending on the Back-End-Of-Line (BEOL) process. The most common materials are: pure silicon dioxide (SiO_2) for blocking oxides, and either SiO_2 or a barrier engineered stack of Oxide-Nitride-Oxide (SiO_2-Si_3N_4-SiO_2) for tunnel oxides. A 2D planar Silicon-Oxide-Nitride-Oxide-Silicon (SONOS) cell is here used as example [1].

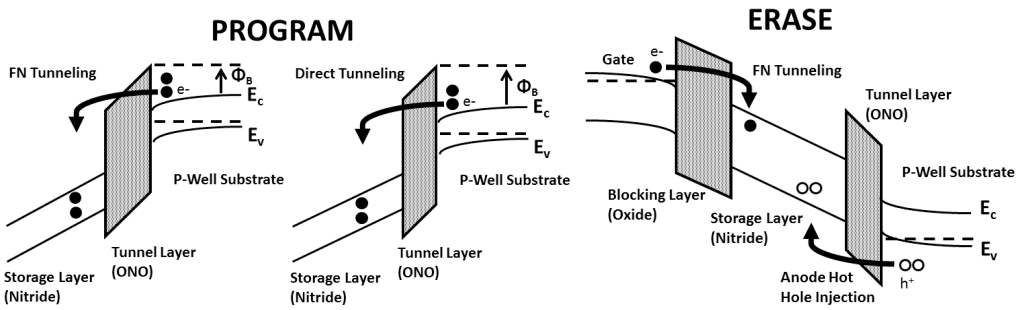


(a) (b)

Figure 2.1: Examples of Charge Trap (a) and Floating Gate (b) devices.

High electric fields applied to the tunnel oxide allow electron transfer across the thin insulator to the storage layer. The physical mechanism used for injecting electrons into the storage layer depends on the applied electric field and oxide barrier thickness. In case of high electric fields and large oxide barriers, injection mainly occurs through FN tunneling, whereas in case of low electric field and thin oxide barrier, electrons mainly transfer through Direct Tunneling (DT): in this case there is a higher read margin window but retention is worse [1]. In CT cells electron tunneling involves the MOS channel/substrate and it requires appropriate biasing of control gate and

bulk terminals (see Fig. 2.2), while drain and source are left floating. Erase operation occurs either through electron detrapping from the storage layer or hole injection from the substrate into the storage layer; at the same time, such operation causes an electron injection from the control gate to the storage layer through FN tunneling, and this is the reason for the well-known "erase saturation" problem [8]. The results of charge separation experiments [9] demonstrate that both electron detrapping and holes injection mechanisms contribute to the erase of a previously programmed CT device: electron detrapping dominates the first part of the transient, whereas hole injection prevails after the removal of the trapped electron charge due to electron emission.



(a) (b)

Figure 2.2: Band diagrams of tunneling mechanisms in planar SONOS CT cell during programming (a) and erase (b). The two different conditions triggering FN or DT are sketched for programming.

2.2 Reliability Issues

Despite the huge potential, several reliability issues affect CT memories, especially endurance and retention.

2.2.1 Endurance degradation

The band diagram depicted in Fig. 2.3 describes oxide degradation mechanisms for blocking and tunnel layers. During programming operations (left), electron injection occurs through either FN or DT, damaging the Tunnel Layer; damages to the Blocking Layer are caused by Anode Hot Hole Injection (AHHI). Moreover, electrons and holes going through blocking layer and tunnel layer from the storage layer contribute in a marginal, but not negligible, way to oxide degradation. During erasing (right), the hot hole injection from the substrate generates interface traps at the oxide/nitride interface, causing several damages to both storage and tunnel layers, as well as electrons transfer through the tunnel layer [8]. The generation of such interface traps between oxide and nitride interface is the main cause of endurance degradation: in programmed cells, electrons sitting in shallow traps can easily escape via oxide damages induced by cycling, resulting in a charge reduction that may cause read errors.

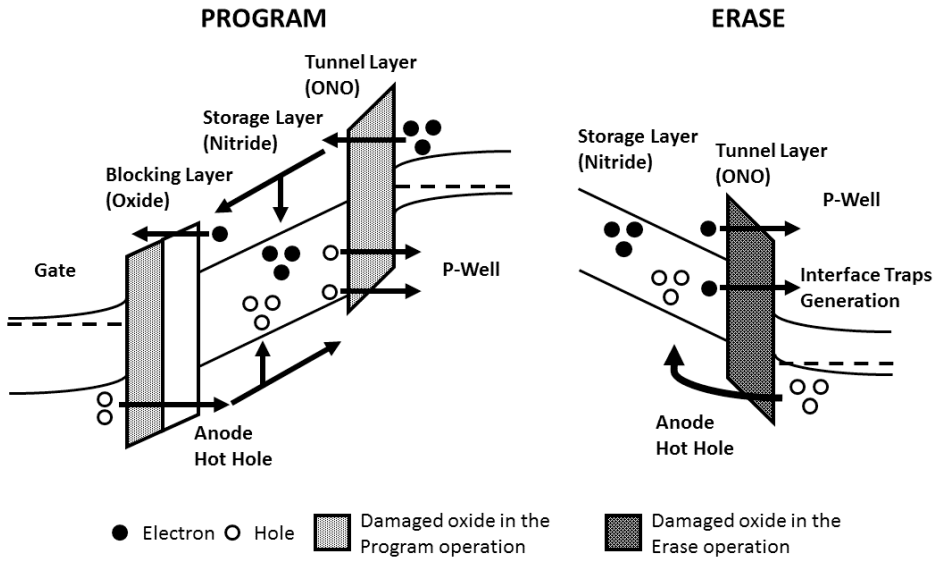


Figure 2.3: Band diagram sketch of charge transport and trapping/detrapping during Program (left) and Erase (right) in planar SONOS CT cell.

2.2.2 Data retention

Data retention is one of the major issues of CT cells, especially at high temperature. Charge loss mechanisms of CT cells has been deeply investigated [10], identifying two main discharging paths: the first is related to thermal excitation of trapped carriers, the second one is due to direct tunneling through the thin tunnel oxide.

The charge loss processes are schematically depicted in Fig. 2.4. For each electron trapped inside the silicon nitride, two discharge mechanisms have to be considered. The first one is the direct Trap-to-Band (TB) tunneling from the storage layer traps to the conduction band of the substrate or of the gate; the second one is the thermal emission from traps to the conduction band of the storage layer. When thermal emission is considered, the charge loss is the result of two subsequent steps: the emission process and the escape of the electrons towards the bulk and gate electrodes. After emission, retrapping is also possible: the tunneling rate through the oxide barrier of the electrons emitted in the storage layer conduction band could be comparable with the emission and the recapture rates. Here we consider a simplified model where electrons leave the ONO layer only if their energy is higher than the lowest between the tunnel oxide and the top oxide barriers. Consequently, the tunneling of thermally excited carriers towards the bulk and the control gate at energies lower than the oxide conduction band are neglected, assuming that carriers with such an energy are recaptured in the same traps.

In addition, a fast initial charge loss has been observed on a small percentage of cells [11] (see Fig. 2.5). This V_T transient phenomenon has been attributed to the dielectric relaxation effect in the high-k layer, to charge trapping/detrapping, or to mobile charges in the blocking layer [11]. Such mechanism, denoted as fast detrapping, is manly related to electrons trapped in shallow traps which have lower stability than electrons in deep traps; they can easily escape via oxide damages within 1 second after programming.

The same effect is observed after erase too: since the threshold voltage after program/erase does not immediately settle to the final value, there is

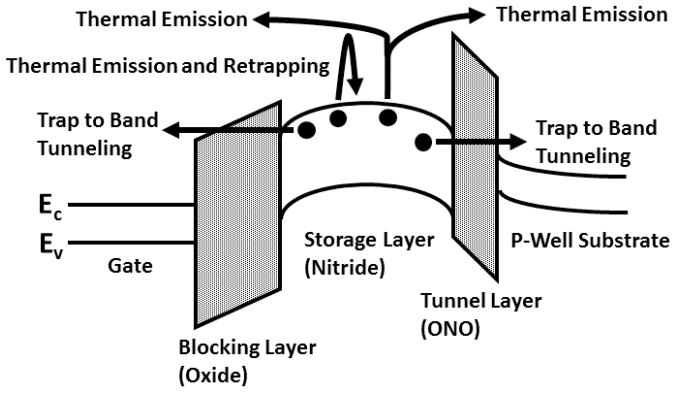


Figure 2.4: Mechanisms involved in the discharging of programmed planar SONOS CT cell: trap-to-band tunneling through the tunnel layer, Trap-to-band tunneling through the blocking layer, Thermal Emission above the oxide barriers, Thermal emission and subsequent re-trapping.

a wrong estimation of the error bits during the verification step; of course, there is a dependency from the time interval between program/erase and read operations. Waiting for the final V_T would significantly increase the total program/erase time and, of course, this is not acceptable. The transient threshold voltage shift after erase is due to hole redistribution in the charge trap layer [12].

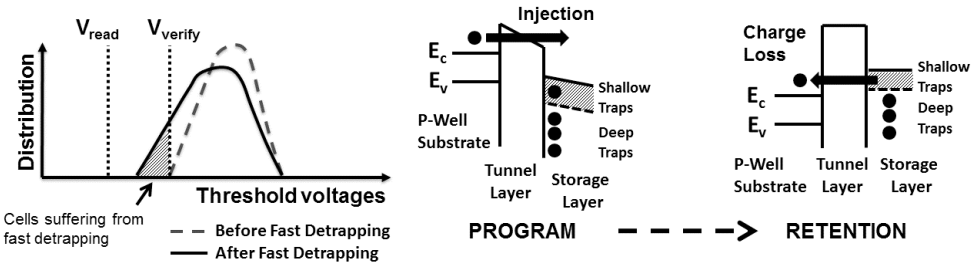


Figure 2.5: Threshold voltage shift induced by fast detrapping (left). Band diagram sketch of fast detrapping effect (right).

2.3 Enhanced algorithms for reliability improvement

CT memories suffer for a relatively low reliability since part of the trapped charge constituting the storage information is located in shallow traps that can be rapidly emptied. Therefore, the charge measured in read operations, and hence the threshold voltage determining the logical stored information, may be different to the one determined during program [11, 13, 14]. This problem is further aggravated by the oxide degradation related to the physical mechanisms exploited to introduce or remove charge into/from the storage layer, thus resulting in a low number of P/E cycles (i.e., the endurance) and a reduced retention time. From the system level point of view both these phenomena cause an increase of the Bit Error Rate (BER), which is the percentage of bits in error after a single read operation [15]. BER increase translates into the inability to correct data after a number of Program/Erase operations (i.e., P/E cycles) or after long retention times [16–18].

To overcome these reliability issues in future CT-based SSDs, a massive exploitation of enhanced read algorithms and data correction by Error Correcting Code (ECC) engines would be mandatory, thus resulting in a degraded Quality of Service (QoS) which is the overall metric considering latency, bandwidth, power consumption, endurance and retention of SSDs [19]. As a consequence, in order to make CT technology appealing in hyper-scaled SSDs, it is necessary to reduce memories' BER and, consequently, the requests for ECC and enhanced reading algorithms intervention. Since the high BER in CT-NAND is mainly due to charge loss in shallow traps, the straightforward solution is to design program algorithms able to stabilize the trapped charge. The enhanced program algorithms presented in literature [11, 20] will result in a programming time increase compared to the standard program algorithms [21], but their advantages in terms of reliability and overall QoS are terrific.

In this section, by characterizing Multi-Level-Cells (MLC) CT-NAND ar-

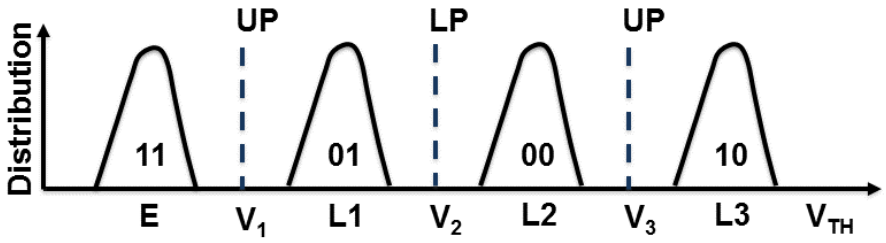


Figure 2.6: MLC target distributions and LP, UP discrimination levels (V_1, V_2, V_3)

rays it is possible to understand the impact of different program algorithms in terms of performance and reliability at a single memory level. The results are then used in simulations using a dedicated SSD co-simulation environment [3] to assess the QoS implications in future SSD architectures integrating multiple CT-based memories.

2.3.1 Read Retry procedure

In this subsection MLC architectures are considered, where each Word-Line (WL) contains an upper page (storing the MSB) and a lower page (storing the LSB). By considering the standard MLC NAND Flash coding (see table 2.1), lower page (LP) is read by applying a read voltage V_2, whereas upper page (UP) is read by applying a read voltage pair (V_1,V_3) as depicted in Fig. 2.6. Therefore, when reading a LP, an error occurs when a cell in L1 moves to L2 and vice versa. On the contrary, when reading an UP, two different errors are possible: a bit flip either between E and L1 or L2 and L3.

Table 2.1: MLC standard NAND Flash coding

	E	L1	L2	L3
UP	1	0	0	1
LP	1	1	0	0

Among several optimized reading techniques, Read Retry (RR) allows a dynamic adaptation of the read reference voltages: when BER > ECC_{th},

the algorithm shifts upwards or downwards the read reference voltage and repeats the read operation until BER $\leq$ ECC$_{th}$ as sketched in Fig. 2.7. If after a maximum number of attempts BER is still higher than ECC$_{th}$, the page is considered as failed, and therefore unrecoverable. Since any page have a different BER and since it is not known *a priori* whether an up-shift or a down-shift must be applied (the former required to deal with endurance effects [22], the latter to take into account retention problems [22, 23]), it is not predictable which solution provides the best results, eventually burdening on the read time predictability and on the reliability.

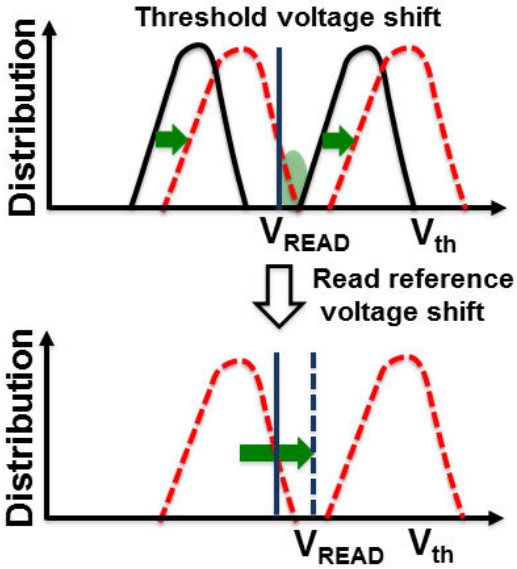
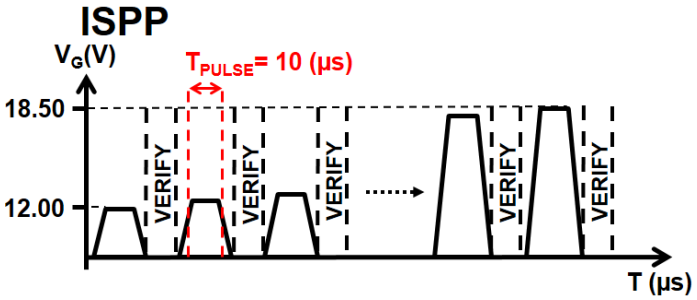


Figure 2.7: Read Retry technique schematic: a) Threshold voltage shift induced by increased number of writing operations. Several cells may result in a threshold voltage higher than the reference read voltage V$_{READ}$, thus producing a read error. b) When the number of erroneous bits is too high to be corrected by ECC the read reference voltage is shifted.

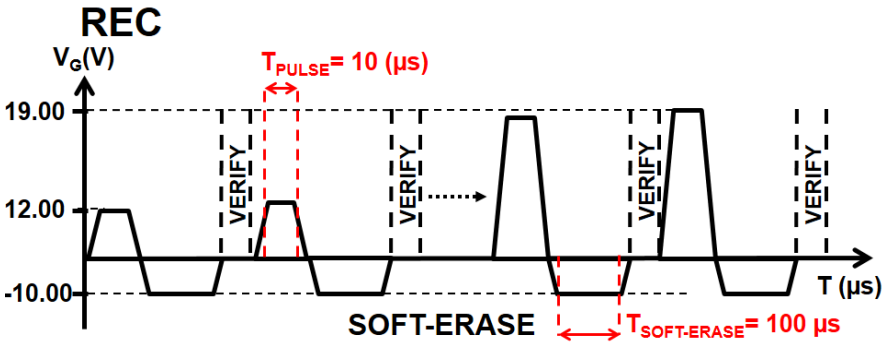
2.3.2 Program algorithms

The standard Incremental Step Pulse Program (ISPP) [21] algorithm and the enhanced one to reduce the charge loss suffered by CT-NAND, here-

after denoted as Recovery (REC) [20, 22], are depicted in Fig. 2.8. MLC paradigm has been performed by defining three target program distributions (L1, L2 and L3). Program operation has been performed by applying the Full-sequence paradigm [24]: after every program pulse a read-verify operation has been performed in order to check the cells state and to stop the algorithm execution when the target distribution is reached. In this example, ISPP algorithm has been performed by increasing the pulse voltage from 12 V up to 18.5 V (depending on the target level) with 0.25 V steps and 10 μs duration (Fig. 2.8a), with a maximum programming time of 1.23 ms for L3 distribution.



(a)



(b)

Figure 2.8: Schematic of ISPP (a) and REC (b) algorithms.

REC has been performed by increasing the pulse voltage from 12 V up to 19 V with 0.25 V steps and 10 μs duration and by applying a soft erase pulse after every program pulse with a constant voltage of -10 V and 100 μs

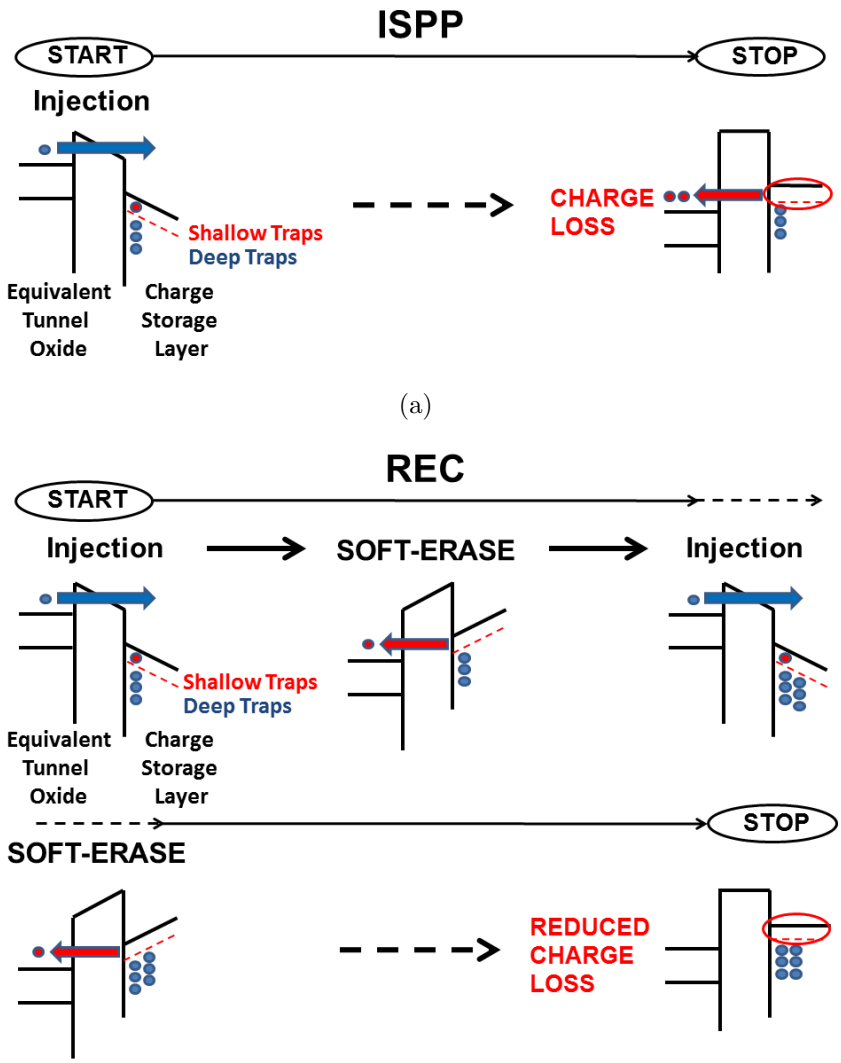


Figure 2.9: Illustration of trapped charge distribution during ISPP (a) and REC (b) program algorithms and charge loss mechanisms after writing.

duration (Fig. 2.8b), with a maximum programming time of 6.97 ms for L3 distribution.

In program operation, electrons cross the equivalent oxide barrier and are randomly captured by deep and shallow traps of the nitride storage layer [25]. In cells programmed by ISPP, electrons in shallow traps are easily de-trapped

during storage period, thus charge loss is observed (Fig. 2.9a) [11]. On the contrary, REC is able to stabilize the stored charge during programming, by reducing the presence of electrons trapped in shallow traps with a high de-trapping probability even at low electric fields (Fig. 2.9b). The soft erase pulses applied during REC are able to remove charge from shallow traps before verify operations occur. Therefore, the target voltage threshold mainly depends on charge stabilized in deep traps [20].

2.4 Electrical characterization of Charge Trap NAND arrays

The different programming techniques for CT-based memories have been experimentally tested on 4Mbits 2D CT-NAND test vehicles manufactured in a sub-4X technology node.

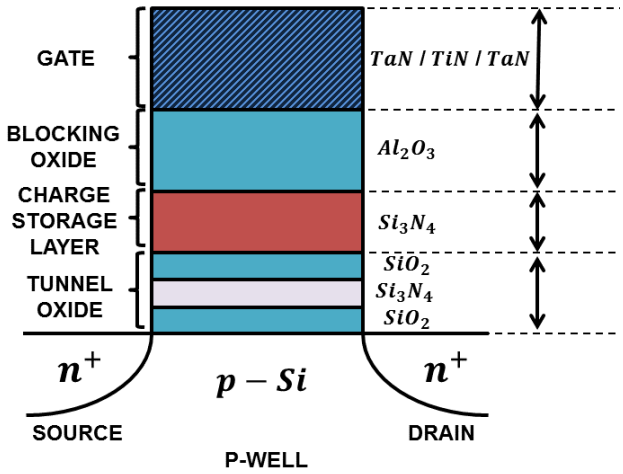


Figure 2.10: Schematic representation of the CT cell tested in this work.

The memory cells feature a p-Si/SiO_2/Si_3N_4/SiO_2/Si_3N_4/Al_2O_3 stack overwhelmed by a high work function TaN/Ti/TaN metal gate (Fig. 2.10). Such a stack will likely be present also in 3D-NAND architectures [26, 27], and therefore the issues retrieved on a traditional 2D technology are inherited

by 3D architectures. The array architecture consists of a standard NAND array, whose pages organization is indicated in Fig. 2.11. The program and the read operations are performed page-wide. The erase operation is performed block-wide with a single voltage pulse featuring 19 V amplitude and 100 μs duration. The program conditions and parameters are the ones reported in the previous section. Those values have been chosen to minimize the disturbs and other unwanted phenomenon due to the soft-erase operation applied on all the cells within a common block [28].

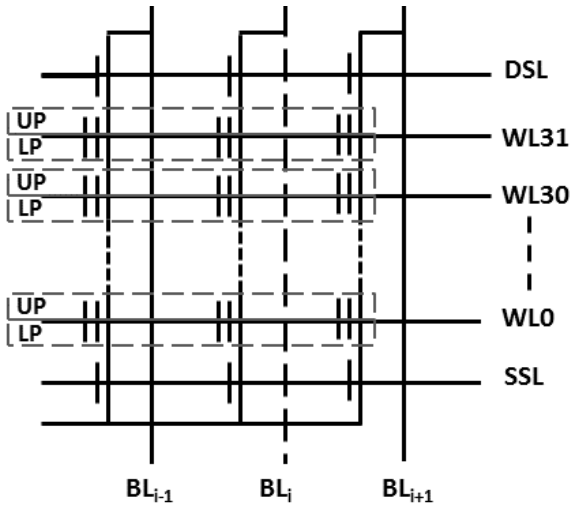


Figure 2.11: Single block schematic of the 4 Mbits CT-NAND array considered in this work.

Fig. 2.12 shows V_T distributions shifts and broadening at different endurance cycles when REC algorithm is performed. Results are even worse for the ISPP algorithm. As it can be seen, the most dangerous effect is due to L1 distribution crossing the V_2 read reference voltage. For this reason, in the rest of the section we will consider LP errors as source of reliability decrease, which represent the worst case without lack of generality.

Fig. 2.13 shows the BER calculated on LP during endurance cycles obtained with ISPP before (a) and after (b) RR application. Thanks to RR, a BER reduction can be observed. Nevertheless a rapid BER increase occurs after 6k P/E cycles. Moreover, due to the edge wordline effects [29], all the

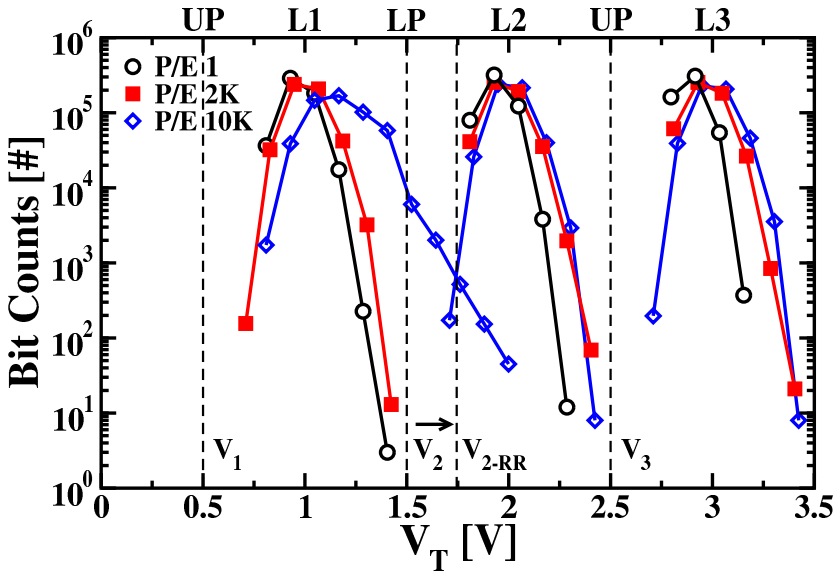


Figure 2.12: REC V_T distributions after a Recovery algorithm at P/E=1, 2k, 10k.

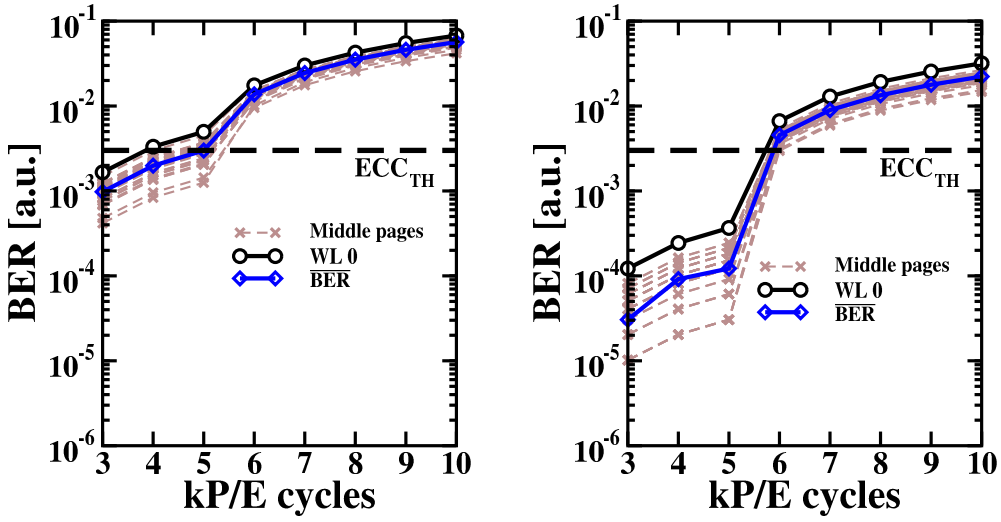


Figure 2.13: ISPP programmed cells LP-BER vs. P/E cycle calculated without (a) and with (b) RR procedure, respectively. ECC_{TH} limit corresponding to 100 errors per read page is shown.

cells on WL0 of a memory string show significantly higher BER compared to the average value, further increasing the average BER. Fig. 2.14 shows the BER calculated on LP during endurance cycles obtained with REC before (a) and after (b) RR application: in both cases the BER is reduced compared to ISPP and it is shown to be lower than 10^{-3} up to 20k cycles thanks to RR procedure.

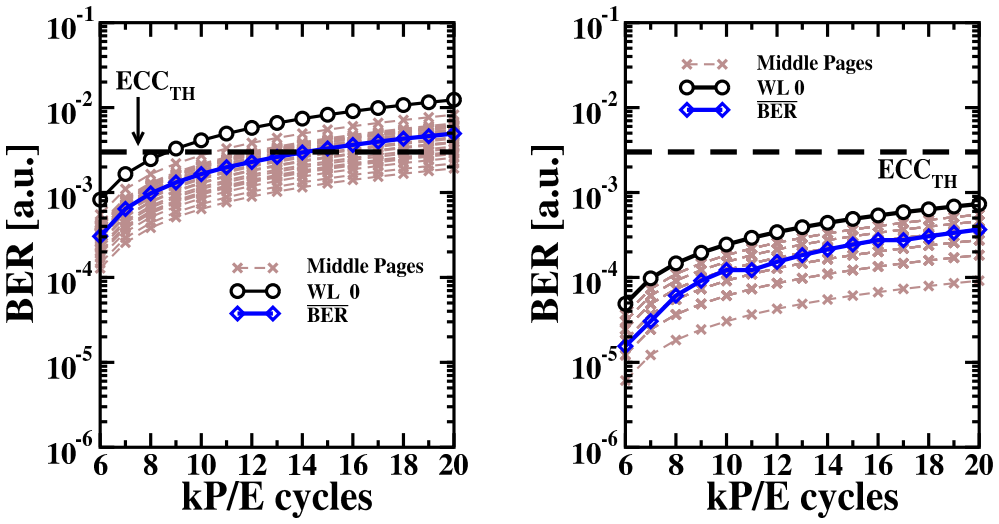


Figure 2.14: REC programmed cells LP-BER vs. P/E cycle calculated without (a) and with (b) RR procedure, respectively. ECC_{TH} limit corresponding to 100 errors per read page is shown.

Fig. 2.15 shows the percentages of uncorrectable pages calculated during cycling for both ISPP and REC algorithms: a page has been considered uncorrectable by the ECC when more than 100 errors are detected on the page (BER > ECC_{TH}) and all RR attempts have been applied. Although RR procedure allows improving ISPP performances, all pages become uncorrectable after 5k cycles. The usage of REC program algorithm combined with RR procedure, on the contrary, allows keeping null the percentage of uncorrectable pages up to 20k cycles. As a comparison, the same uncorrectable pages percentages have been calculated not considering the contributions coming from

cells on WL0 pages. However, no significant advantages are obtained in this case. In fact, when ISPP with RR is considered, all the pages are shown to suddenly cross the ECC_{TH} limit after 6k P/E cycles, hence the failures are almost simultaneous. On the contrary, when REC with RR is considered, all pages' BER is below ECC_{TH} limit up to 20k P/E cycles, hence there is no perceived impact.

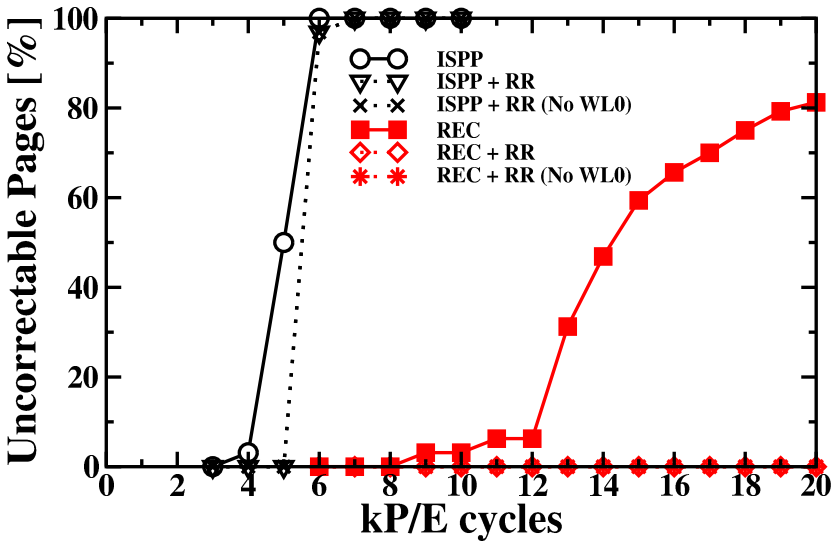


Figure 2.15: Uncorrectable pages percentage calculated at different endurance cycles.

Fig. 2.16 show the BER evolution for MLC during retention measures at 85°C after P/E = 2k. Dotted lines represent the results obtained with RR procedure: 8 Retry Steps of 25 mV each have been used on MLC resulting in a final threshold shift of 200 mV. The results obtained for long term retention are similar for the two algorithms, whereas Recovery is significantly better if the early retention domain is considered. The experimental results evidence that retention is the main reliability issue on this technology.

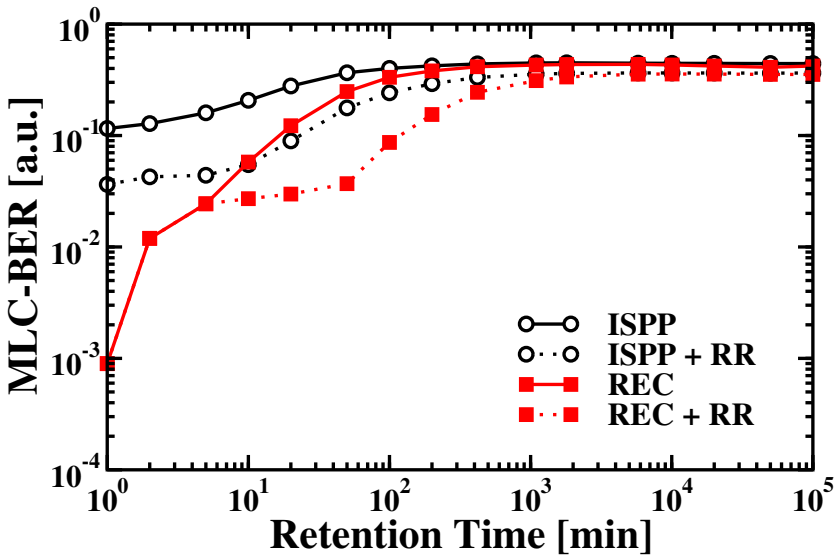


Figure 2.16: MLC-BER vs. Retention Time at 85°C after P/E=2k. Dotted and full lines are calculated with and without RR procedure, respectively.

2.5 SSD applications perspectives

In this section, the different programming algorithms performance and reliability figures obtained through the characterization of the CT-NAND arrays are evaluated from a SSD-QoS perspective [7]. This task has been performed exploiting a co-simulation framework able to extract performance and latency of a target disk architecture as a function of memory wearout, while allowing detailed reliability analysis [3–5]. Fig. 2.17 shows the baseline architecture modeled by the simulator: it is composed by a processor, a host interface, a DRAM buffer, a channel controller, a multi-threaded BCH ECC engine [30] and a regular matrix of non-volatile memory targets. Different CT-NAND chips sharing the same bus are defined as channels. Each single channel is connected in parallel on the same channel controller.

In the following system level analysis both the ISPP and the REC algorithms embodying RR algorithms will be considered since this represents a realistic study case used in all SSD platforms [31,32]. A QoS threshold value has been set for an enterprise scenario, in terms of host interface bandwidth.

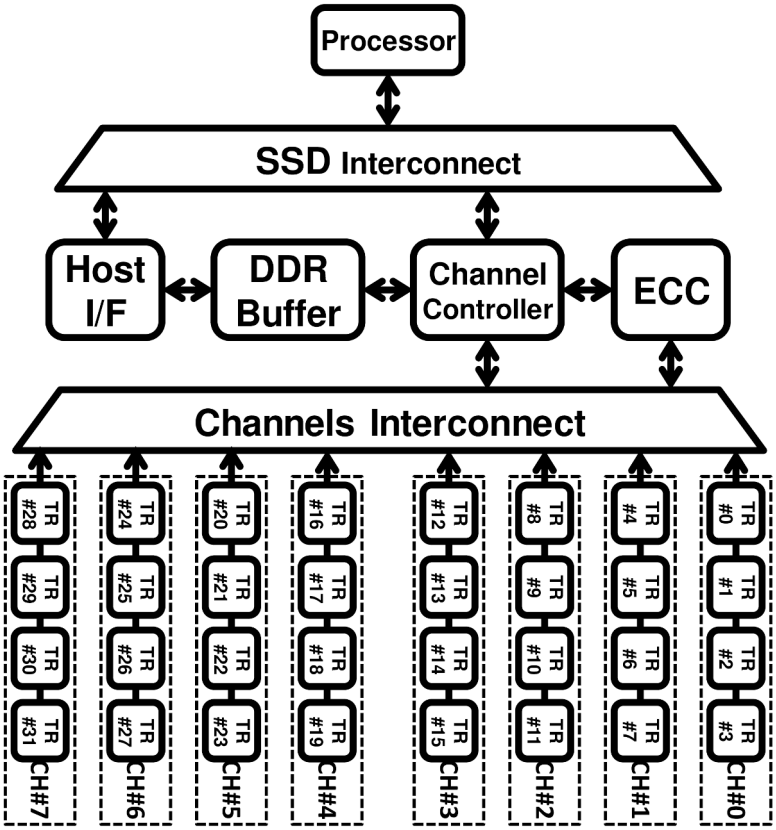


Figure 2.17: SSD baseline architecture modeled by the simulator.

For a standard interface like the PCIexpress Gen2 x8 [33] the limits corresponds to 4 GB/s which is further reduced to 1150 MB/s due to the host system I/O drivers overhead [34, 35].

In the simulated architectures a SSD is considered to miss the QoS target when the achieved bandwidth is below such threshold. It is worth to point out that if the bandwidth achieved by the SSD is greater than the target QoS, any performance fluctuation introduced by the programming algorithm, the ECC, and the RR are not exposed to the end user, since the perceived overall SSD bandwidth is the one imposed by the QoS limit. Program algorithm's QoS implications have been investigated on multiple SSD configurations exploiting different number of channels and targets (see Tab. 2.2).

Table 2.2: Architecture configurations

	A	B	C	D	E
Channels	8	8	4	2	1
Targets	4	2	4	8	8

The results of the simulations on different SSD architectures, performed by using a 100 % sequential read workload after a program operation at different endurance cycles using either the ISPP or the REC, are shown in Fig. 2.18: only the configuration A is able to satisfy enterprise QoS requirements for a defined number of P/E cycles that varies by using one of the two programming algorithms. When considering SSD architectures using the ISPP algorithm, in all cases a sudden performance drop is experienced after 3k P/E cycles because of the too high bit error density. Configurations A, B, and C show a faster QoS degradation around 3k P/E cycles with respect to D and E. This effect is due to the high number of flash targets saturating the channel bandwidth, hence masking the initial performance drop. A similar behavior is observed for SSDs using the REC algorithm, although a larger number of P/E cycles can be experienced before performance degradation.

Since the scope of this work is to show CT-NAND-based SSDs' exploitation in hyper-scaled systems, the following analysis will be focused on enterprise class SSD's, hence only configuration A will be considered further on. This configuration meets the high density requirements provided by 3D memories exploiting CT technology that is leading to use these solutions for cold storage applications, in which data are written once and read many times [36]. As a consequence, read-intensive workload is the most expected use case for such memories, whereas write-intensive workload represent a corner case.

The 100 % sequential read throughput performances obtained with ISPP and REC algorithms are reported in Fig. 2.19: while with ISPP no more than 3k P/E cycles are possible before QoS degradation, the endurance obtained with REC is improved by a factor of 4, reaching almost 12k P/E cycles.

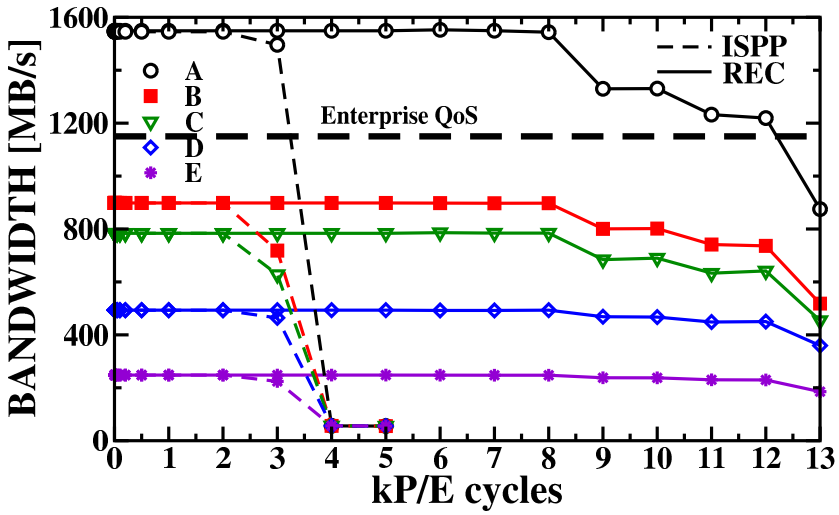


Figure 2.18: 100 % sequential read throughput calculated at different endurance cycles with different architecture configurations for ISPP and REC, respectively.

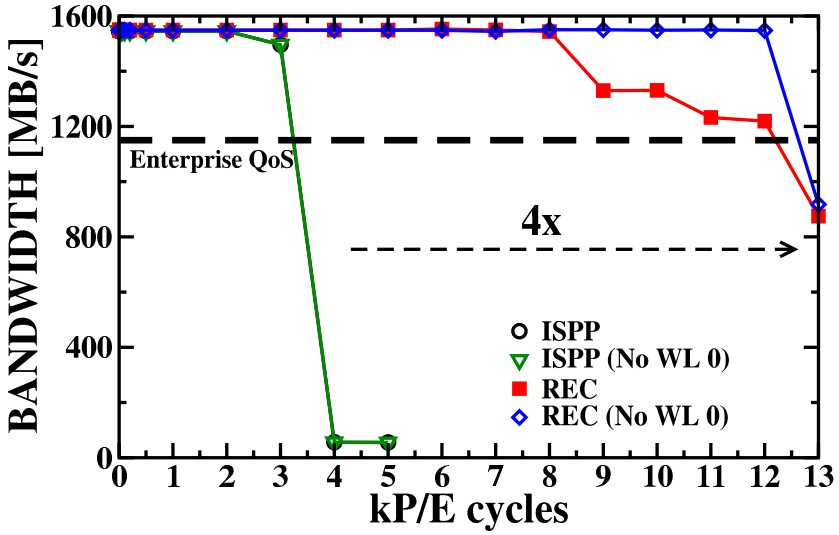


Figure 2.19: Sequential read throughput calculated at different endurance cycles.

The experienced bandwidth degradation is due to the joint effect of the RR and ECC engines whose execution time increases with the number of errors to be corrected. Moreover, by considering that WL0 pages usually show

higher BER than others, if it is used a Flash Translation Layer in the SSD that excludes those pages from programming, REC throughput can get a further endurance boost of ≈ 500 P/E cycles. On the contrary, when ISPP is considered, no relevant advantages are obtained in terms of bandwidth by using this approach, because of the too high bit error density. Finally, in the case of a 100 % random read workload, results obtained perfectly matched those achieved with the 100 % sequential read workload. This phenomenon is due to the high parallelism offered by the simulated SSD architectures, which is able to sustain the output bandwidth even when random operations are issued. To this extent, random results are not shown. The usage of enhanced program algorithms like the REC in SSD architectures carries the drawback of an increased program time. To this extent, in order to understand if a longer program time could impact the perceived SSD's bandwidth, two mixed traffic scenarios have been considered: 75% write and 25% read, 75% read and 25% write (both using 4 kB interleaved random read and write operations) [37]. The latter is the one closer to the scenario that are targeted in this work. SSD throughput results are shown in Fig. 2.20. Even if in write-intensive conditions ISPP shows a slightly higher throughput compared to REC, both algorithms do not satisfy the QoS enterprise requirements. REC allows satisfying enterprise QoS requirements up to 12k P/E cycles if read-intensive or 100 % read conditions are considered. Bandwidths obtained with mixed scenarios involving more than 25% write operations cross the QoS limit for both ISPP and REC algorithms. However, it is worth pointing out that such workloads represent a worst-case corner for hyper-scaled SSDs which are foreseen in cold storage scenarios.

In conclusion, the use of the proposed REC program algorithm combined with RR read algorithm shows promising advantages in terms of SSD's QoS for cold storage scenarios, where retention issues could be counteracted at cost of endurance cycles by periodically re-programming the memory cells in background without any impact on the SSD throughput. The use of 3D CT cells with improved retention capabilities it is expected to reduce this issue.

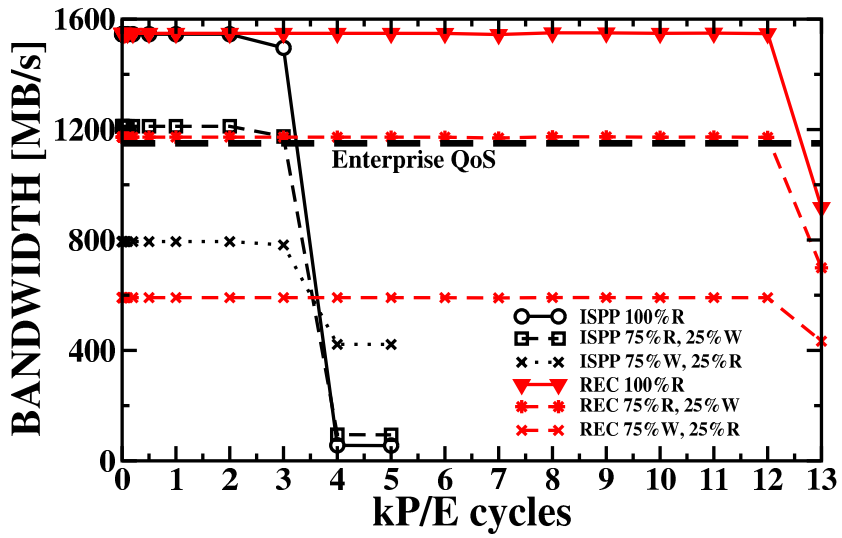


Figure 2.20: Random mixed traffic throughput calculated at different endurance cycles.

Chapter 3

RRAM

Resistive RAM (RRAM) is considered as one of the most promising emerging nonvolatile memory technology since it features high scalability, low power consumption, high endurance, fast switching and possibility of 3D stackability. Recent advances in the RRAM performance have led to a significant interest in system-on-chip applications in Si-based CMOS technologies, in particular for microcontrollers in wireless sensor nodes, low power and wearable IoT, automotive electronics and neuromorphic computing. HfO_2 is one of the most promising transition metal oxides for RRAM with an ideal CMOS back-end-of-line compatibility. Thus, considerable progress has been made in 1T-1R device integration as well as in understanding the physical/chemical properties of the resistance change behavior. Although memory arrays in the 1T-1R architecture demonstrated excellent performance parameters, the intercell variability (variations between cells) and the intracell variability (cycle-to-cycle variations of any given cell) still prevent RRAM manufacturing from fast commercialization. While the intracell variability can be optimized for the particular memory cells, the intercell variability in memory arrays must be minimized. Hence, the investigation of the extrinsic process-induced and the intrinsic microscopic origin of the intercell and intracell variability is a critical step in order to bring such technology to a maturity level. In this chapter, the results obtained on RRAM will be pre-

sented and discussed focusing on the variability, which is the main reliability issue of these nonvolatile memories. The impact of the switching operations on variability will be evaluated, starting from single pulse operations up to program and verify algorithms. The quantum point contact model will be used to provide a physical understanding of the experimental results. After that, the process parameters impact on variability and reliability will be discussed. Finally, the fundamental variability limits of such technology will be defined through an extensive array characterization and radiation hard application perspectives will be provided.

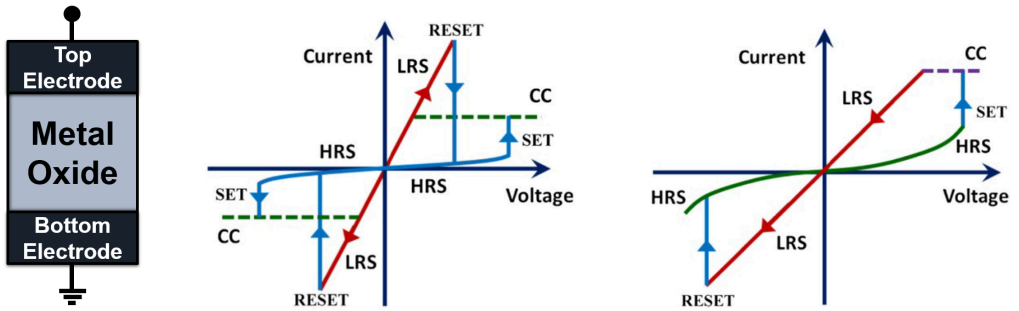
3.1 Basics

RRAM device typically consists of an insulating layer, usually a metal oxide, interposed between a Top Electrode (TE) and a Bottom Electrode (BE), both generally consisting of metallic layers or stacks [38–41]. Typical material used for the electrodes are Ti, TiN, W and Pt [38] whereas the most common metal oxides used in RRAM are HfO_2 [42], Ta_2O_5 [43] and TiO_2 [44]. A sketch of the typical RRAM cell is reported in Fig. 3.1 (a).

These devices exhibit resistive switching between a low-resistance state (LRS) and a high-resistance state (HRS) due to the creation/disruption of a conductive filament (CF) through the metal oxide: the switching event from HRS to LRS is called Set while the switching event from LRS to HRS is called Reset. Usually for the fresh samples in its initial resistance state, a voltage larger than the Set voltage is needed to trigger on the resistive switching behaviors for the subsequent cycles: such process is called Forming.

The switching modes of metal oxide RRAM can be broadly classified into two switching modes: unipolar and bipolar. Fig. 3.1 shows a sketch of the I-V characteristics for Unipolar (b) and Bipolar (c) switching modes. In unipolar devices the switching direction depends on the amplitude of the applied voltage but not on its polarity, thus Set/Reset can occur at the same polarity. In bipolar devices, on the contrary, the switching direction depends on the

polarity of the applied voltage, hence Set can only occur at one polarity and Reset can only occur at the reverse polarity. For either switching modes, to avoid a permanent dielectric breakdown in the Set process, it is recommended to enforce a compliance which is usually provided by the semiconductor parameter analyzer or by a memory cell selection transistor or diode. To read the data from the cell a small read voltage is applied that does not affect the state of the memory cell to detect whether the cell is in HRS or LRS.



(a) (b) (c)

Figure 3.1: (a) Schematic of MIM structure for metal oxide RRAM, and I-V curves schematic showing two modes of operation: (b) Unipolar and (c) Bipolar.

The CF creation during Forming and its evolution during Set and Reset are depicted in Fig. 3.2. During the initial Forming operation a CF of oxygen vacancies is formed by dielectric breakdown: oxygen ions drift to the anode interface by the high electric field where they are discharged as neutral non-lattice oxygen if the anode materials are noble metals or react with the oxidizable anode materials to form an interfacial oxide layer. Thus, the electrode/oxide interface behaves like an oxygen reservoir [45].

Forming operation is performed by applying a positive voltage on TE, causing an increase of the current flowing through the cell: such increase is usually limited by the compliance system or a series selector/transistor that allows to control the size of the CF and avoids the destructive (hard) breakdown of the switching layer. After Forming, the device manifests improved

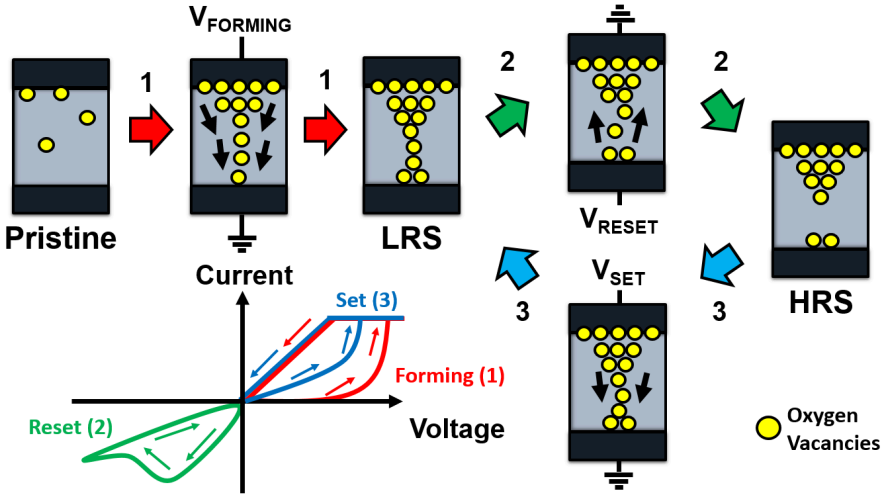


Figure 3.2: Schematic illustration of the switching process in bipolar metal oxide RRAM.

conductance as the CF connects the TE and BE by shunting the insulating layer, thus resulting in the low-resistance state (LRS) of the RRAM. In bipolar RRAM a Reset operation is then carried out to disconnect the CF by applying a positive voltage on BE: Joule heating from the current thermally activates the diffusion of oxygen ions that migrate back to the bulk either to recombine with the oxygen vacancies or to oxidize the metal, returning the memory cell to the HRS [38]. Alternating the Set and Reset operation, the CF can be repeatedly connected/disconnected, thus allowing multiple transition cycles between HRS and LRS. Note that the conductance of HRS is higher compared to the initial state before forming: this can be understood by the microscopic structure of the HRS, where the CF is not entirely dissolved after Reset, rather only disconnected via a relatively small depletion gap.

In general, Fig. 3.3 shows all the possibilities for an electron pass from cathode to anode [23]:

1. Schottky emission: thermally activated electrons injected over the barrier into the conduction band.

2. Fowler-Nordheim (F-N) tunneling: electrons tunnel from the cathode

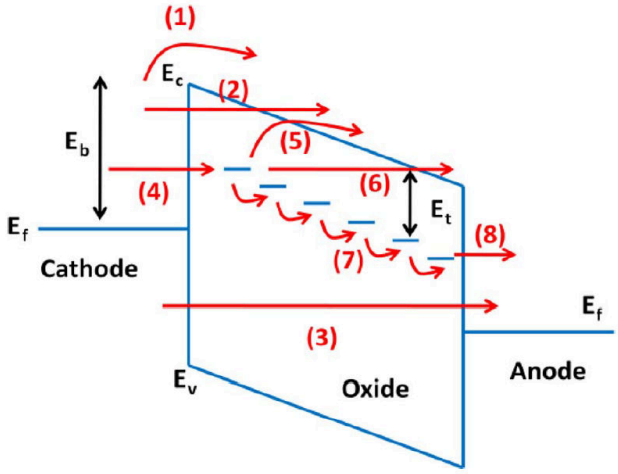


Figure 3.3: Schematic of the possible electron conduction paths through a MIM stack.

into the conduction band; usually occurs at high field.

3. Direct tunneling: electron tunnel from cathode to anode directly; usually occurs when the oxide is thin enough.

If the oxide has substantial number of traps (e.g., oxygen vacancies), trap assisted tunneling contributes to additional conduction including the following steps:

4. Tunneling from cathode to traps.

5. Emission from trap to conduction band (e.g. Poole-Frenkel emission).

6. F-N-like tunneling from trap to conduction band.

7. Trap to trap hopping or tunneling: may be in the form of Mott hopping when the electrons are in the localized states or may be in the form of metallic conduction when the electrons are in the extended states depending on the overlap of the electron wave function.

8. Tunneling from traps to anode.

Whether any one particular process dominates is determined by its transition rate: electrons would seek the least resistive paths among all the possibilities. Therefore, various metal oxide RRAMs may have different dominant conduction mechanism depending on the dielectric properties (band gap or

trap energy level, etc.), the fabrication process conditions (annealing temperature, annealing ambient, etc.), and the properties of the interface between the oxides and the electrodes (interfacial barrier height). For HfO_2 RRAM, it has been verified that at low bias regime the HRS conduction is dominated by trap assisted tunneling, whereas the conduction is mainly Ohmic if the CF connects the TE and BE by shunting the insulating layer [46, 47]. The I-V relationship at the low bias regime is mainly determined by the electron conduction process for a given configuration of the CF, while at the high bias regime, the motion of atoms (such as oxygen ions/vacancies) would change the configuration of the CFs and trigger a change of the current [38].

3.2 Experimental Setup

The test environment consists of both wafer-level testers and a dedicated Automated Test Equipment (ATE) for packaged devices analysis. For wafer testing the Advantest V93000 SOC, Keithley 4200-SCS and Agilent B1500 (synchronized with an Arduino ATmega2560 for the array digital addressing part) were used, whereas for packaged devices testing it has been used the Active Technologies RIFLE system including arbitrary waveform generators and programmable measurement units resources for testing purposes [48–51]. The tests performed indicated a negligible deviation between results from wafer and packaged devices. Different memory cells and arrays were considered, hence each one of them will be described in its related session.

3.3 DC and Pulse-induced Forming analysis

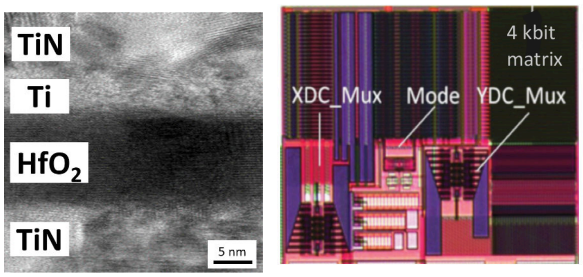
In this section are investigated the inter-cell variability of the initial state in memory arrays and the impact of direct current (DC) and pulse Forming on inter-cell variability as well as on intra-cell variability. To assess and confirm the nature of the variability during forming operation and during cycling the quantum point-contact (QPC) model was considered [52]. This

analysis was performed on 4kbit RRAM arrays tested with RIFLE ATE.

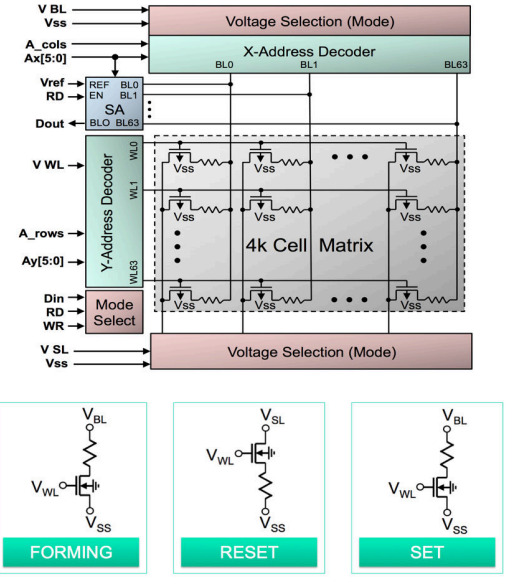
The 1T-1R memory cells considered in this section are constituted by a select NMOS transistor manufactured in 0.25 μm BiCMOS technology featuring W=1.14 μm and L=0.24 μm in series to a variable resistor. The variable resistor is a Metal-Insulator-Metal (MIM) stack fabricated on a 150 nm TiN bottom electrode deposited by physical vapor deposition (PVD) sputtering with sheet resistances in the order of 10-50 Ω/sq. A 9 nm HfO_2 was then deposited at 320°C by the reaction of O_2 and [$Hf(NMeEt)_4$)] onto TiN in an atomic vapor deposition (AVD) chamber. Finally, HfO_2 was capped by 7 nm ionized metal plasma (IMP) Ti and 150 nm PVD TiN [53]. The cross-Sectional STEM Image of the integrated MIM stack is shown in Fig. 3.4 (a). Such cells were integrated in 4 kbit RRAM arrays [49–51,54]: the structure of the array shown in Fig. 3.4 (b) is described by four architectural blocks: the array of 4096 1T-1R RRAM cells; a wordline (WL) address decoder (XDC MUX); a bitline (BL) address decoder (YDC MUX); and an operation control circuitry (Mode). The array schematics is depicted in Fig. 3.4 (c).

3.3.1 Inter-cell variability before and after DC Forming

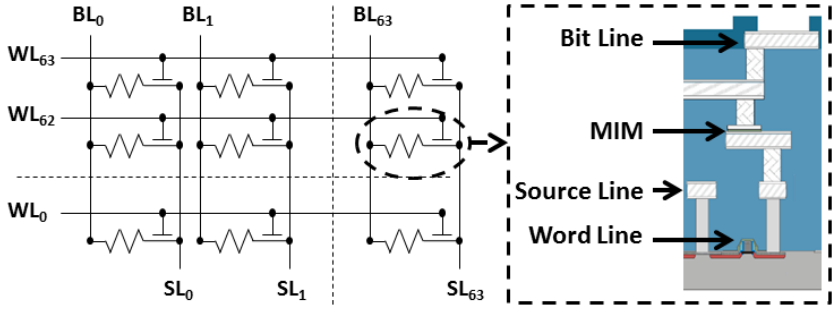
In this subsection, the influence of DC Forming process on the inter-cell variability is reported. Forming is a prerequisite to induce stable resistance changes in our devices. The common Forming process consists in the application of a DC sweep on the bitline (BL) grounding the sourceline (SL). To prevent hard breakdown, the saturation current of the select transistor is controlled by the wordline (WL) voltage. Table 3.1 summarizes the Forming, Set/Reset and reading parameters used in this section. In general, the Forming process produces a non-destructive soft breakdown and progressive breakdown regime of the dielectric and requires a sufficiently high electric field across the HfO_2 film thickness [41, 52, 55]. Due to the stochastic nature of the Forming process or due to process-induced variations, the inter-cell variability of memory elements in a 4 kbit memory array cannot be neglected.



(a) (b)



(c)



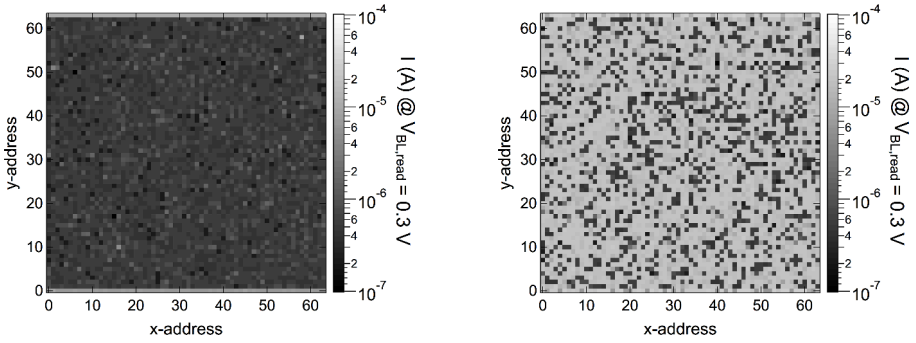
(d)

Figure 3.4: Cross-Sectional STEM Image of the integrated MIM stack in the RRAM Cell (a). Photograph of the 4 kbit memory array with control circuits (b). Simplified block diagram of the 4 kbit memory array (c). Schematic of the 1T-1R cells array and structure of a RRAM cell (d).

To address this point, we first show in Fig. 3.5 the current distribution in the initial state which is defined by the combination of the Ti and HfO_2 film thicknesses [51] and the current distribution after DC Forming.

Table 3.1: Summary of DC forming, pulse-induced forming, pulse-retry forming, DC-retry forming, Set/Reset and read voltages.

	V_{BL}	V_{SL}	V_{WL}	$t_{pulse/sweep}$
DC Forming	0 V - 2.3 V	0 V	1.4 V	23 s
DC Set	0 V - 2.3 V	0 V	1.4 V	23 s
DC Reset	0 V	0 V - 2.3 V	2.2 V	23 s
Pulse-induced Forming	3.5 V	0 V	1.4 V	10 μs
Pulse-retry Forming	3.5 V	0 V	1.4 V	10 μs
DC retry Forming	0 V - 3.5 V	0 V	1.4 V	-
Pulse Set cycling	3 V	0 V	1.4 V	10 μs
Pulse Reset cycling	0 V	3 V	2.2 V	10 μs
Reading	0.3 V	0 V	1.4 V	-



(a) (b)

Figure 3.5: Current distributions for a 64 × 64 bit (4 kbit) RRAM array with 1T-1R devices with 600 × 600 nm^2 MIM area. (a) Initial state. (b) After Forming in DC step sweep. WL voltage in Forming was set to $V_{WL} = 1.4$ V with a bitline voltage sweep to $V_{BL} = 2.3$ V with ramp rate $dV/dt = 0.1$ Vs^{-1}. Current reading was performed at $V_{WL} = 1.4$ V and $V_{BL,\,read} = 0.3$ V.

Statistical analysis in a cumulative distribution plot in Fig. 3.6 reveals that only 78 % of the devices (red curve) were successfully formed by the DC sweep, whereas 22% of the devices were not formed due to the high inter-cell variability of the initial state current. The thickness inhomogeneity of the PVD Ti layer with $\sigma < 5\%$ standard deviation over the whole wafer diameter is believed not to play a central role in the inter-cell variability of the initial currents in Fig. 3.6 since the memory array with circuitry area is 365×752 μm^2. When compared with Ti, the thickness inhomogeneity of the HfO_2 film with $\sigma < 2\%$ standard deviation should also be insignificant for the initial current inter-cell variability [56].

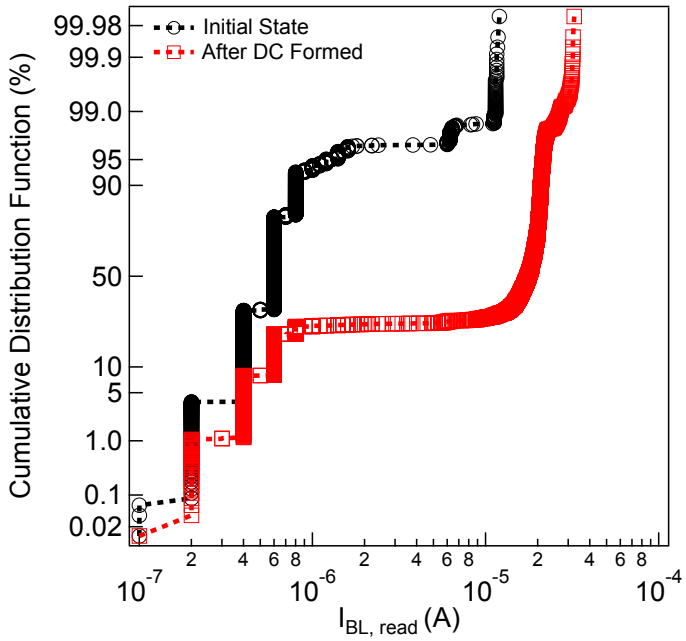
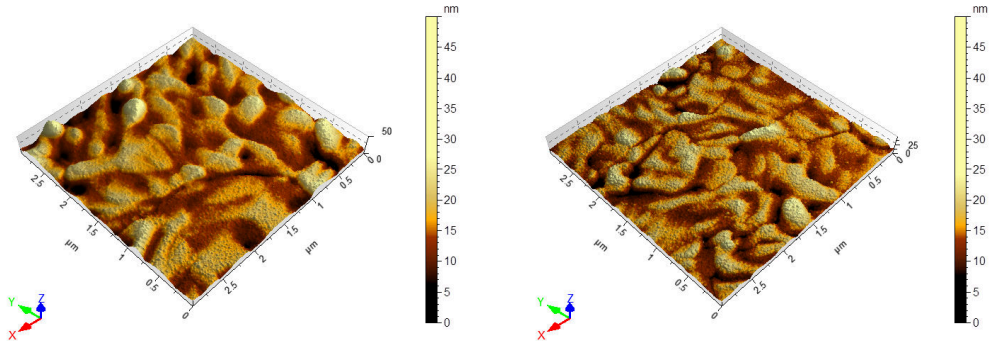


Figure 3.6: Statistical current distributions of the initial state and after Forming in DC sweep. Current reading was performed at $V_{WL} = 1.4\,V$ and $V_{BL,\,read} = 0.3\,V$.

One reason could be the root mean square surface roughness of HfO_2 films due to the columnar structure of the TiN bottom metal electrode [57], although a post-metallization annealing (PMA) step was applied which was reported to reduce the surface roughness [58]. To support this hypothesis it

is shown in Fig. 3.7 the different root mean square values of the roughness profiles extracted from TiN and HfO_2 films used in RRAM arrays processing. The randomly distributed inter-cell variability after forming in DC sweep is often related to intrinsic processes which involve microscopic transport of oxygen vacancies and permutations of filamentary clusters [59, 60].



(a) (b)

Figure 3.7: Roughness profile of the TiN (a) and HfO_2 (b) films used in the array processing.

3.3.2 Inter-cell variability after pulse-induced Forming

The pulse-induced Forming has attracted wide attention as an alternative to DC Forming since it may reduce the operation current and the energy per unit volume imparted to the dielectric [61, 62]. Moreover, the Forming time for the whole memory array can be reduced. From a physical point of view, the pulse-induced Forming has an impact on the shape of the created conductive filament: while pulses with low V_{WL} and V_{BL} voltages or short pulse width $t_{BL,\,pulse}$ create narrow Filaments showing low currents, higher V_{WL} and V_{BL} voltages or long pulse width $t_{BL,\,pulse}$ create stable and wider filaments [42]. The filament diameter and areal density may be correlated with the concentration of oxygen vacancies that are supposed to participate in the filament formation process. The oxygen vacancies are predominantly produced at the Ti metal electrode as the result of redox processes of the oxide

by the Ti metal electrode [63]. To understand the pulse-induced Forming properties we applied $V_{WL} = 1.4\,V$, $V_{BL} = 3.5\,V$, and $t_{BL,\,pulse} = 10\,\mu s$ to a fresh memory array. With these parameters, the pulse-induced Forming energy density (voltage × current × time/area) imparted to the dielectric for each cell reaches an average value of $1.9 \times 10^{-9}\,J/\mu m^2$ in comparison to $3.7 \times 10^{-3}\,J/\mu m^2$ after DC Forming. Figures 3.8(a) and 3.8(b) show the current distributions of the initial state and after pulse-induced Forming.

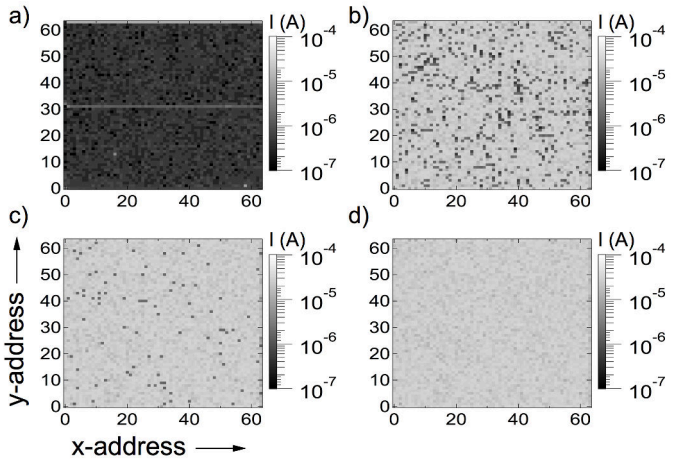


Figure 3.8: Current distributions for a 64 × 64 bit (4 kbit) RRAM array with 1T-1R devices with 600 × 600 nm^2 MIM area. (a) Initial state. (b) After pulse-induced Forming. V_{WL} was set to 1.4 V with $V_{BL} = 3.5\,V$ and $t_{BL,\,pulse} = 10\,\mu s$: total 86 % formed. (c) After applying a retry-algorithm to unformed devices: total 97.6 % formed. (d) After DC Forming of the devices not formed by the retry step: total 100 % formed. Current reading was performed at $V_{WL} = 1.4\,V$ and $V_{BL,\,read} = 0.3\,V$.

Statistical analysis in a cumulative distribution function plot in Fig. 3.9 reveals that 86 % of the devices were formed by the pulse process. A retry-algorithm was then implemented to form the remaining devices. The retry operations have been mostly required on cells with low initial currents. One reason could be that devices with a lower initial current need a higher energy density imparted to the dielectric to be formed. Forming with the same

parameters as before was applied 5 times on BLs with unformed devices. This increased the fraction of formed devices in the array to 97.6 % (Figs. 3.8(c) and 3.9). The remaining 2.4% devices could be formed by a DC Forming step (Figs. 3.8(d) and 3.9). After Forming, positive V_{SL} step sweeps Reset the devices to the high resistance OFF state, whereas positive V_{BL} step sweeps Set the devices to the low resistance ON state (see Table 3.1). The OFF/ON resistance contrast on repeated switching cycles remains typically between 1.4 and 10.

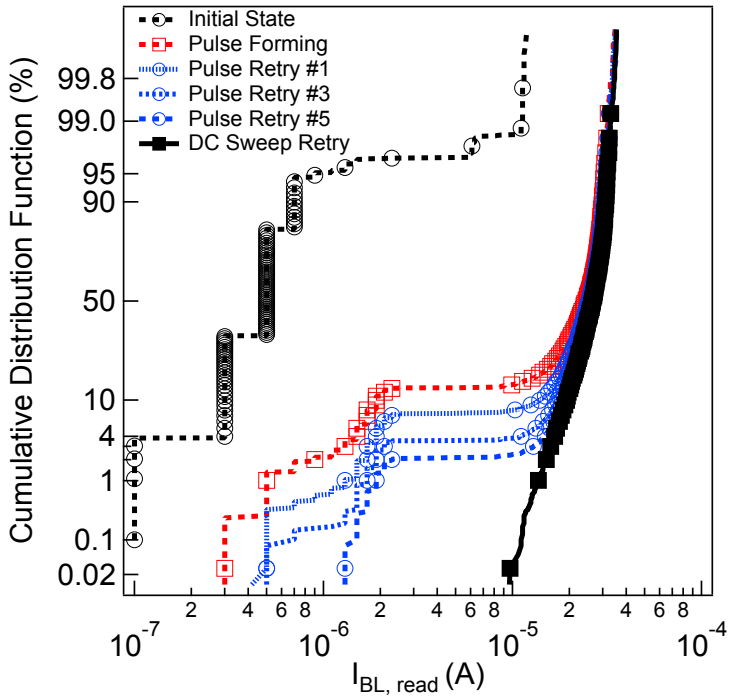


Figure 3.9: Statistical current distributions of the initial state, after pulse-induced Forming and after application of retry-algorithms. Current reading was performed at $V_{WL} = 1.4\,V$ and $V_{BL,\,read} = 0.3\,V$.

3.3.3 Impact of Forming on endurance

Set/Reset single pulse operations were performed after pulse Forming and after pulse-retry Forming to evaluate the cycling endurance (Fig. 3.10).

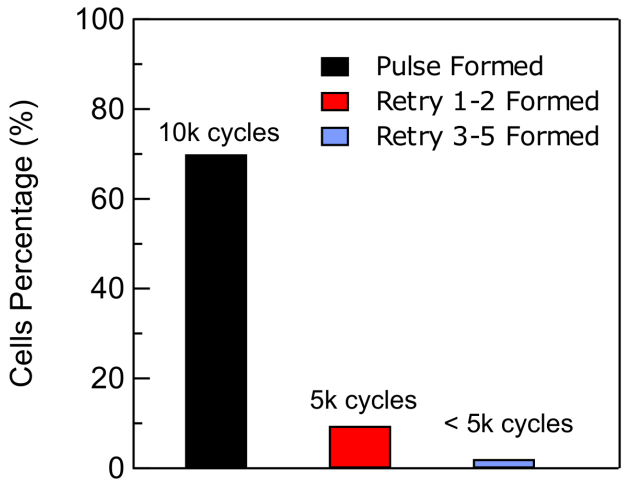


Figure 3.10: Set/Reset cycling endurance after pulse-induced forming and after pulse-retry forming.

70 % of the devices formed with the first Forming pulse show the highest endurance, whereas almost 16% of them evidenced resistance switching issues due to a possible Over-Forming process [64].

As the number of Forming-retry pulses increases, the resistance switching capability of the conductive filament is strongly impacted: retry-formed cells show lower cycling endurance compared to single-pulse formed cells. The retrieved endurance of 10^4 cycles is significantly lower than what retrieved for similar arrays [42,65] and this could be attributed to the impact of impurities in the metal-organic AVD precursor, in particular to carbon [63]. To support such a consideration, Fig. 3.11 shows a XPS depth profile measurement of a 50 nm thick HfO_2. Since the cells embedded into the array are too small for the analysis, XPS profile has been measured on cells with a higher HfO_2 stack, but processed with the same deposition parameters used for the MIM cells in the array. A high presence of carbon and nitrogen atoms can be observed: these impurities are caused by the liquid metal-organic precursor, used for the AVD deposition process. Their presence severely limits the oxide performance and wears the cells in the array faster than any other degradation mechanism.

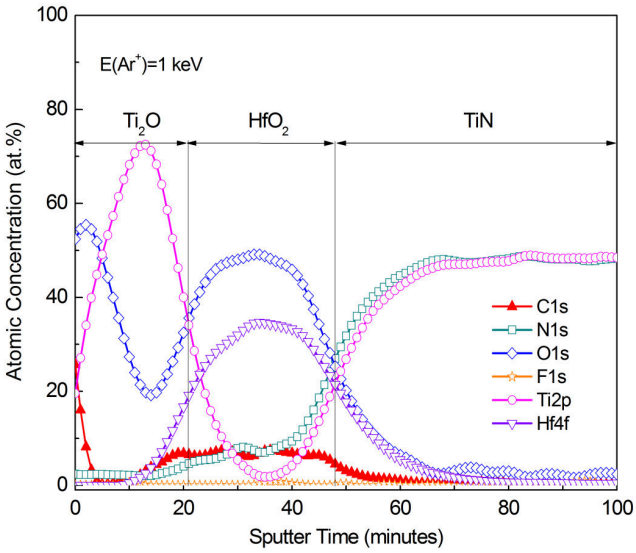
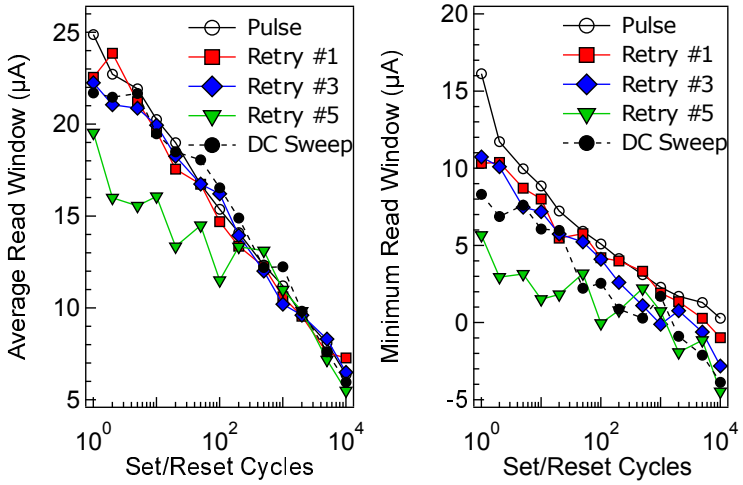


Figure 3.11: XPS depth profile measurement of a 50 nm thick HfO_2 RRAM cell.

To evaluate the relationship between Forming conditions and endurance, formed cells with a single pulse, after 1, 3, 5 Forming-retry pulses and after DC step sweep were considered (Table 3.1). Figures 3.12 (a) and 3.12 (b) show the average and minimum (i.e. considering the worst-case condition) read window calculated during cycling for each cell subset [50]. For all cell subsets both average and minimum read window decrease with increasing number of cycles.

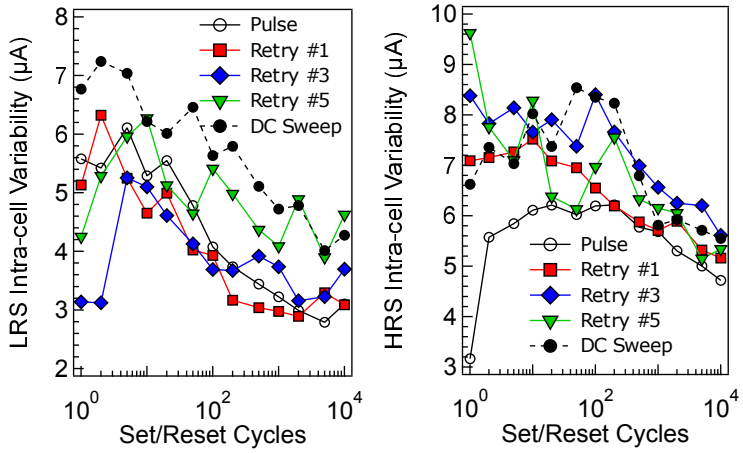
Figure 3.13 (a) shows the ON state (LRS) intra-cell variability during cycling, which indicates decrease for each cell subset. The highest variability is observed on cells formed after 5 pulse-retry algorithms and after DC sweeps. Similar behavior is observed for the OFF state (HRS) intra-cell variability during cycling in Fig. 3.13 (b). The reduction of the intra-cell variability during cycling is given by the fact that a consecutive Set/Reset operation brings all cells toward a uniform wear-out condition of the conducting filament, therefore reducing the fluctuations due to structural modifications of the conduction path in the memory cells [66].

In order to investigate the reason underlying different degradation speed



(a) (b)

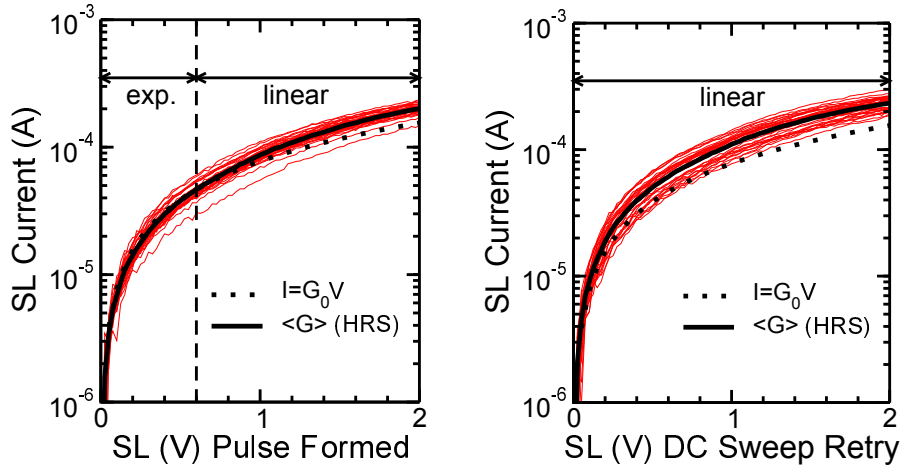
Figure 3.12: (a) Average and (b) minimum read window after a single pulse, after 1, 3, 5 forming-retry pulses and after DC sweep.



(a) (b)

Figure 3.13: Intra-cell variability during 10^4 pulse set/reset cycling for (a) LRS and (b) HRS.

for different Forming retry times, I-V measurements of the Reset operation have been performed at endurance cycle 10^4. Fig. 3.14 shows HRS I-V curves of the cells formed with a single pulse (a) or with DC sweep retry operations (b). The black dashed line shows the limit $I = G_0V$ with $G_0 = 2e^2/h$ equal to the quantum conductance unit corresponding to the creation of a single mode nanowire according to Quantum Contact Point (QPC) model [52], where e is the electron charge and h the Planck's constant. The average $< G >$ curves measured after Reset is shown for sake of comparison.



(a) (b)

Figure 3.14: HRS I-V curves measured on cycled devices after Reset on cells formed with a single pulse (a) and with DC retry operation (b).

Within this framework, in case of $I > G_0V$ the presence of a residual conductive filament has to be taken into account. Cells formed with a single pulse show lower average conductance values after Reset, even after cycling. Moreover, on single pulse formed cells an exponential ($V_{SL} < 0.6V$) plus linear ($V_{SL} > 0.6V$) current behavior can be observed since $< G >$ is very close to $I = G_0V$ limit, whereas only linear behavior is observed on $< G >$ of DC sweep retry formed cells. This indicates that for single pulse formed cells, the filament geometry in HRS after cycling still evidences a potential barrier giving rise to direct tunneling transport through a material barrier

or through a residual confinement barrier (QPC) [67] (i.e., the exponential part of the curve), whereas for DC sweep retry formed cells only the ohmic component is present due to worn-out filament (see Fig. 3.15).

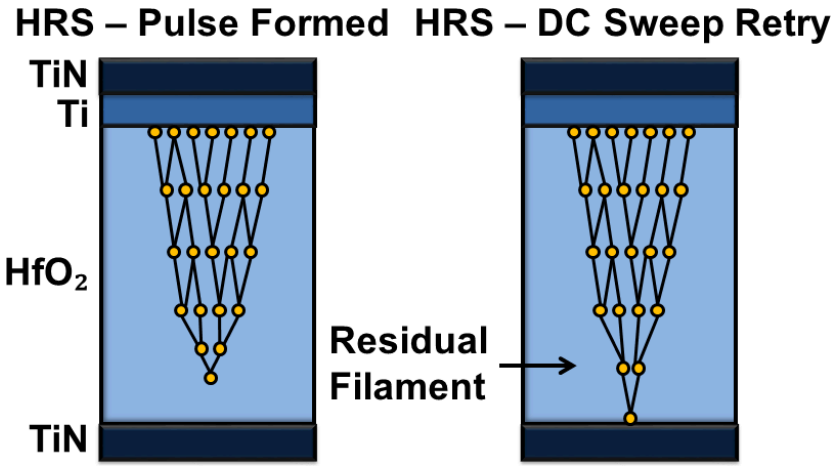


Figure 3.15: Schematic showing the conductive filament shape after Reset (HRS) for Pulse and DC Retry formed cells.

Fig. 3.16 shows the cumulative distribution of conductance values measured after Reset operation of cells formed at different retry steps: cells formed with a higher number of retry operations show higher conductance values and intra-cell variability. More retry steps allow the cells to reach a more stable state resulting in higher conductance values yet displaying lower switching capabilities. More energy is indeed required in these cells in order to create the conductive filament compared to single pulse formed cells, hence it can be expected that even during cycling a higher Set/Reset energy was required to create/disrupt the filament on such cells. Due to such lack of energy, a higher intra-cell variability is observed since Set/Reset operations were not correctly performed on all retry-formed cells.

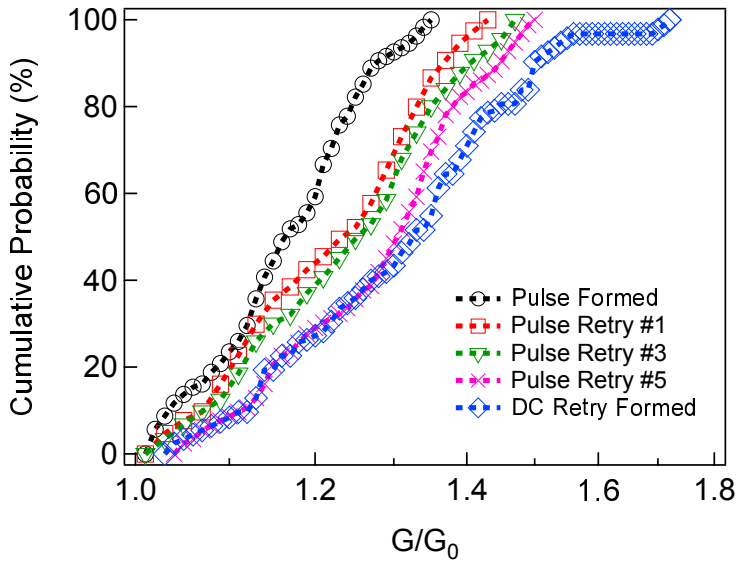


Figure 3.16: Cumulative distribution of conductance values measured after Reset on cycled cells.

3.4 Switching Parameters analysis

In this section a Set/Reset pulse investigation on RRAM arrays is performed in order to find the most reliable Set/Reset operation conditions. The analysis will compare DC and pulsed Set/Reset operations featuring different durations and voltages on previously formed 1×1 μm^2 TiN/HfO_2/Ti/TiN 1T-1R 4 kbit memory arrays. A thorough analysis of the RRAM reliability joining the cell-to-cell variability analysis to that of cycling evaluations in complete arrays is addressed. A comparison between DC and Pulse Set/Reset in terms of switching yield, read window and device-to-device uniformity is reported. Finally, the impact of a temperature bake at 125 °C on a cycled array is shown to study the temperature impact on the array variability. The memory cell and array full descriptions are provided in 3.3.

In order to get a first glance of the average switching voltages, Set and Reset I/V characteristics have been measured with a DC staircase on 30 memory cells: the large device variability evaluated on the Set/Reset switching kinetics between LRS and HRS is shown in Fig. 3.17. By plotting the aver-

age Set and Reset I/V characteristics along with the behavior of 30 memory cells it can be observed that each cell features its own characteristic switching voltage. Moreover, as usually evidenced in RRAM technology [68], HRS shows a larger range of variability compared to LRS, as evidenced in Fig. 3.18 showing the cumulative distributions of the Set/Reset switching voltages calculated on the entire array. These results indicate that an optimization of the Set/Reset operations is mandatory to reduce the impact of the device variability whereas minimizing the array yield loss due to non-switching cells. To this purpose, the following analysis will compare DC Set/Reset operations and pulsed ones featuring different durations and voltages. Furthermore, it is necessary to evaluate the reliability of a chosen Set/Reset operation in terms of read current ratio I_{HRS}/I_{LRS} stability during cycling.

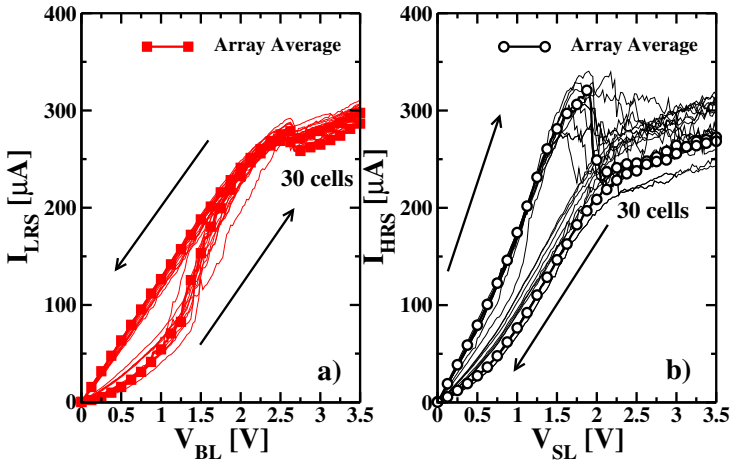


Figure 3.17: Set (a) and Reset (b) switching kinetics performed on 30 cells in the RRAM array test structure

Set operation in DC mode has been performed increasing the bitline voltage V_{BL} from 0 to 3.5 V with V_{step} = 0.1 V ($T_{step,DC}$ = 50 μs) and the wordline voltage fixed to V_{WL} = 1.4 V. Reset operation in DC mode has been performed increasing the source line voltage V_{SL} from 0 to 3.5 V, with V_{step} = 0.1 V ($T_{step,DC}$ = 50 μs) and V_{WL} = 2.5 V.

In pulsed mode operation the wordline voltage has been fixed to V_{WL}

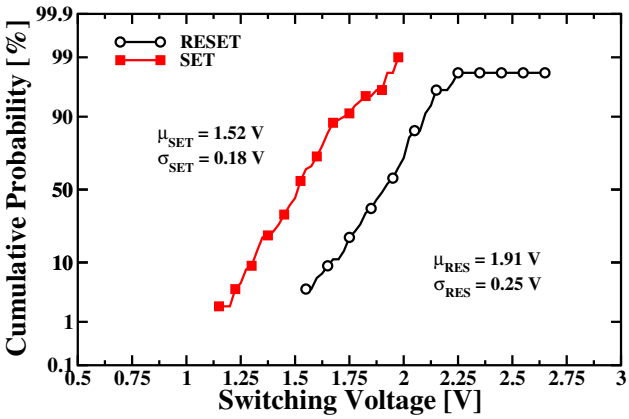


Figure 3.18: Cumulative distribution of the Set and Reset switching voltages in the entire RRAM array.

= 1.4 V during Set and V_{WL} = 2.8 V during Reset, while different BL/SL voltages and durations have been investigated. In Fig. 3.19 a comparison between DC and pulse mode with different durations at fixed V_{pulse} = 3 V is depicted. Fig. 3.19(a) shows I_{HRS}/I_{LRS}, normalized with respect to that calculated at cycle 1, as a function of the Set/Reset cycle number for different pulse durations. In all cases a non-monotonic behavior is observed, eventually ending up with a significant I_{HRS}/I_{LRS} reduction with the exception of the shortest pulse duration (T_{pulse} = 1 μs). Fig. 3.19(b) shows the switching yield (i.e. the percentage of cells in the array that actually toggles between Set/Reset states) of each Set/Reset mode proving an interesting trade-off: pulses with a too short or too long duration result in a lower yield compared to an average timing condition. Similar considerations can be derived by the analysis of Fig. 3.20, where the dependencies of the normalized I_{HRS}/I_{LRS} and that of the switching yield are evaluated in cycling for different pulse voltages considering the optimal pulse duration (T_{pulse} = 10 μs).

From a physical point of view, this phenomenon is related with the size of the conductive filament created during Set: while pulses with too low voltages or durations create too small filaments showing low current in Set condition, too high voltages or durations create too big filaments hard to

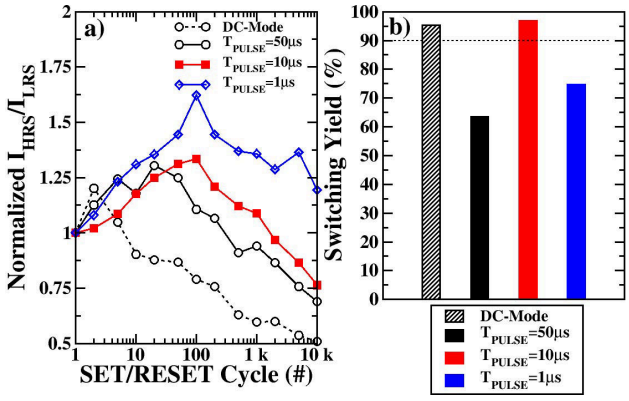


Figure 3.19: Normalized read current Ratio (a) and Switching Yield (b) evaluation for different Set/Reset modes (DC and pulses with different durations) during cycling. V_{pulse} = 3 V for pulsed modes.

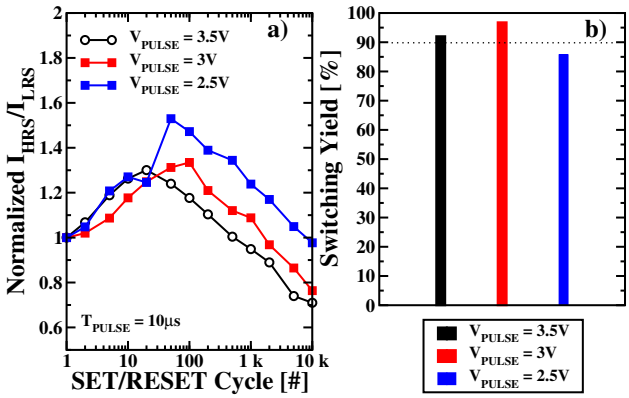


Figure 3.20: Normalized read current Ratio (a) and Switching Yield (b) for different Set/Reset pulse amplitudes during cycling. T_{pulse} = 10 μs.

disrupt in the following Reset operations. Both cases result in a lower yield compared to an average condition.

Starting from the best pulse conditions (T_{pulse} = 10 μs, V_{pulse} = 3 V) the read window closure has been analyzed as a function of Set/Reset cycling (see Fig. 3.21). Current reading was performed at V_{WL} = 1.4 V, $V_{BL,read}$ = 0.2 V, T_{read} = 10 μs. The average read current trend and the $\pm\sigma$ deviations are plotted for Set and Reset. It can be observed that the device variability of the cells in the array remains almost constant during cycling.

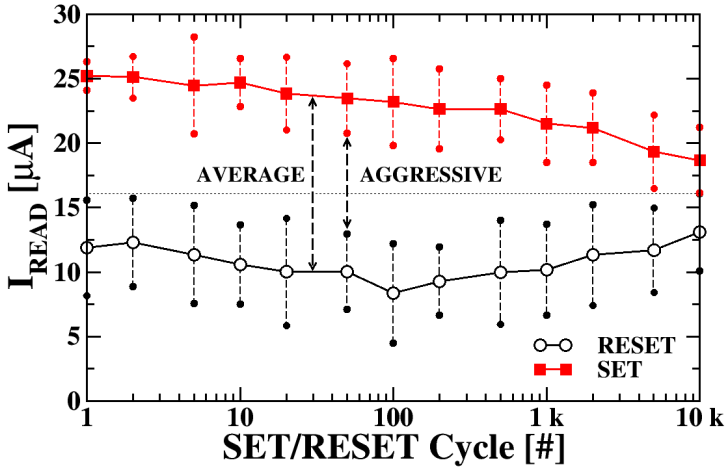


Figure 3.21: Set and Reset read current behavior during cycling with V_{pulse} = 3 V, T_{pulse} = 10 μs. Average and aggressive read window calculation points are indicated.

Fig. 3.22 shows the read window ($I_{LRS} - I_{HRS}$) closure calculated on the array average and aggressive (considering the worst-case) conditions. In this study, the endurance failure criterion is defined as the point where the aggressive read window case falls below 3 μA, that is the limit for the sense amplifier to discriminate between states.

The read windows show the same behavior in each pulse condition : an increase can be observed during the first cycles due to a variability reduction, followed by a closure after the degradation of the HfO_2 material stack. Short pulse durations and voltages result in a smaller read window due to a higher device variability caused by incomplete Set/Reset switching.

Fig. 3.23 shows the cumulative distributions at Set/Reset cycle 1 and 10K for a pulsed Set/Reset operations with V_{pulse} = 3 V, T_{pulse} = 10 μs before and after a temperature bake at 125 °C. I_{LRS} does not show any relevant variation after bake, whereas impacting on the cell-to-cell variability: the average current is still ≈18.5 μA, while σ decreased from 2.68 to 1.83 μA. I_{HRS} shows a slight shift from ≈13 to ≈12 μA after bake, while σ increased from ≈3 to ≈4 μA. These results demonstrate that a temperature

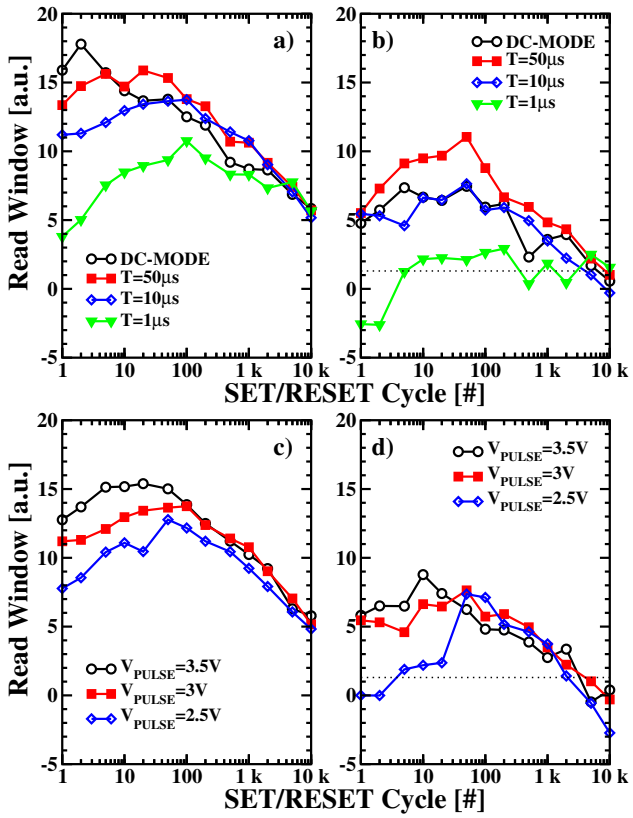


Figure 3.22: Read Window trends in cycling. Average (left column) and aggressive (right column) measurements: a) and b) same conditions as Fig. 3.19; c) and d) same conditions as in Fig. 3.20. The limit for the Set/Reset discrimination is depicted at read window = 3 μA.

bake shows a rather confined impact on a cycled array: while cycling seems to be the main reliability detractor for the RRAM technology, the temperature bake produces a relative shift of the read window due to the modifications concerning the cell-to-cell variability.

3.5 Incremental Programming with verify

In order to control the intrinsic RRAM variability, an incremental programming with verify algorithm is defined and characterized in this section.

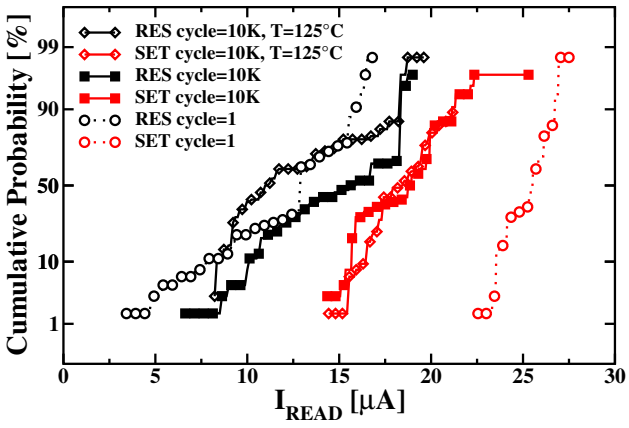


Figure 3.23: Read current cumulative distributions at Set/Reset cycle 1 and 10K for a pulsed Set/Reset V_{pulse} = 3 V, T_{pulse} = 10 μs and temperature bake effect.

The results obtained during Forming, Set and Reset are reported, showing the advantages obtained in terms of reliability compared to the previous section. QPC modeling is used to evaluate the effectiveness of the proposed algorithm in controlling the conductive filament shape. Such algorithm is also used as a tool to study the cell behavior during Forming and to relate the observed current oscillations with conductive filament properties, stability and reliability. The full description of memory cell and array used in this section is provided in 3.3.

3.5.1 Incremental Forming with Verify

Even if forming process is performed just once, as showed in the previous section it plays a fundamental role in determining the subsequent array and system performance [69–72]. The effectiveness of the forming process depends on its ability in creating homogeneous conductive conditions among the cells thus easing successive Set/Reset operations. Standard forming is performed by applying either a voltage ramp or pulse to each cell individually: such methods do not allow a tight control of the filament conductance, resulting in larger cell-to-cell variability and larger disturb sensitivity [73, 74]. As an

alternative, forming process can be performed through a sequence of pulses featuring the same voltage. While the forming time can be minimized using a single pulse with high compliance and voltage parameters, the forming yield may be limited since the applied energy may not be sufficient to complete the forming process in all cells [49]. Several pulse-based forming alternatives have been proposed to increase the applied energy and therefore the yield by using long pulses or sequences of short pulses at constant voltage [73].

In this section an incremental pulse and verify Forming technique is proposed and compared with different pulse-based forming techniques in terms of forming time, yield and cell-to-cell variability on 4 kbits RRAM arrays. Considering the peculiarity of each cell in terms of the switching behavior activation, it is shown that a tight control of the forming process allows taking profit on a long term basis during the successive Set/Reset operations [59, 68]. Among all the investigated techniques, the proposed procedure (already used for Set/Reset operation [68]) showed the most promising results. Such method, hereafter referred to as Incremental Form and Verify (*IFV*) leverages on the application, to any single cell, of a sequence of trapezoidal waveforms with increasing maximum voltages, each step being followed by a current read operation that monitor the cell resistance. When a cell reaches a predefined read current value after the pulse application the procedure is interrupted, so that all cells are brought into a comparable electrical condition. In fact, even if the cell requires longer forming time, *IFV* offers a superior advantage by significantly reducing the cumulative number of Set/Reset pulse and verify operations during cycling and, consequently, the overall power consumption of the memory peripheral circuitry. After-forming modeling of Reset I-V operations has been performed through Quantum-Point Contact (QPC) model [52], showing that if a too high conductance is reached during forming the filament became hard to disrupt in the successive Reset operation, resulting in faulty behavior [74, 75]. Moreover, it is shown that thanks to the verify procedure implemented during forming such faulty behavior can be avoided. Tab. 3.2 summarizes the Forming, Set, Reset and

Read parameters used in this section.

Three different forming schemes have been characterized:

a) Single pulse, denoted as *Pulse* in Fig. 3.24(a): V_{BL} = 3.5 V, V_{WL} = 1.4 V, with pulse duration T_{pulse} = 10 μs and a finite $t_{rise} = t_{fall} = 1\mu$s to avoid overshoot effects [76].

b) Incremental Form, denoted as *IF* in Fig. 3.24(b): the bitline voltage V_{BL} was increased with a sequence of increasing voltage pulses from 2 to 3.5 V with ΔV_{BL} equal to 0.1 V, a wordline voltage V_{WL} of 1.4 V and T_{pulse} = 10 μs, $t_{rise} = t_{fall} = 1\mu$s.

c) Incremental Form and Verify (*IFV*), Fig. 3.24(c): pulses were applied with increasing V_{BL} from 2 V up to 3.5 V with two different ΔV_{BL} equal to 0.1 V and 0.01 V, respectively, V_{WL} = 1.4 V and T_{pulse} = 10 μs, $t_{rise} = t_{fall} = 1\mu$s. After each forming pulse the cell read current I_{read} was measured: if $I_{read} >$ 19 μA the forming process was interrupted and the cell marked as formed.

Table 3.2: Summary of Forming, Set, Reset and Read parameters.

Operation	V_{BL} [V]	V_{SL} [V]	V_{WL} [V]	T_{pulse} [μs]
Pulse Form	3.5	0	1.4	10
IF Form (ΔV_{BL} = 0.1 V)	2-3.5	0	1.4	10
IFV Form (ΔV_{BL} = 0.1 V)	2-3.5	0	1.4	10
IFV Form (ΔV_{BL} = 0.01 V)	2-3.5	0	1.4	10
Set (ΔV_{BL} = 0.1 V)	1.5-3.5	0	1.4	10
Reset (ΔV_{SL} = 0.1 V)	0	1.5-3.5	2.8	10
Read	0.2	0	1.4	10

Fig. 3.25 shows the distributions of the IFV forming voltages, confirming that the specific voltage conditions triggering the forming behavior are quite different for each cell within the array. All the forming schemes considered in this section base on pulse widths of 10 μs. This is mainly due by two factors: performance/variability concerns, and technological limitations of the integrated array.

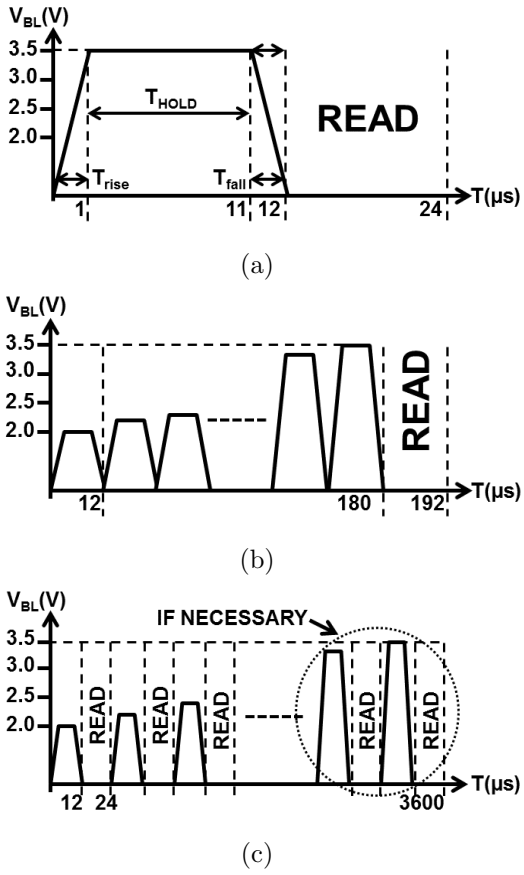


Figure 3.24: Single Pulse (a), *IF* (b) and *IFV* (c) Forming schemes.

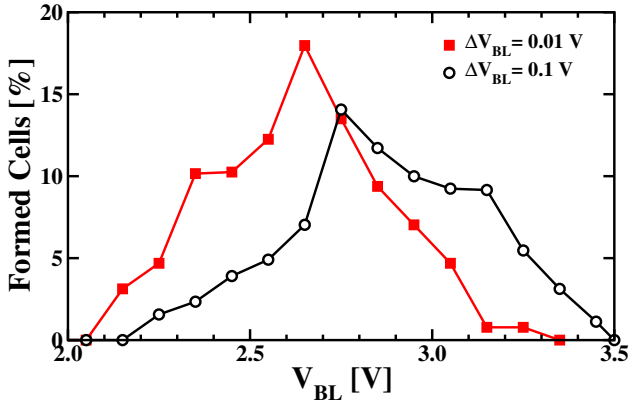


Figure 3.25: Distributions of the forming voltages during *IFV* for the ΔV_{BL} = 0.01 V and ΔV_{BL} = 0.1 V cases.

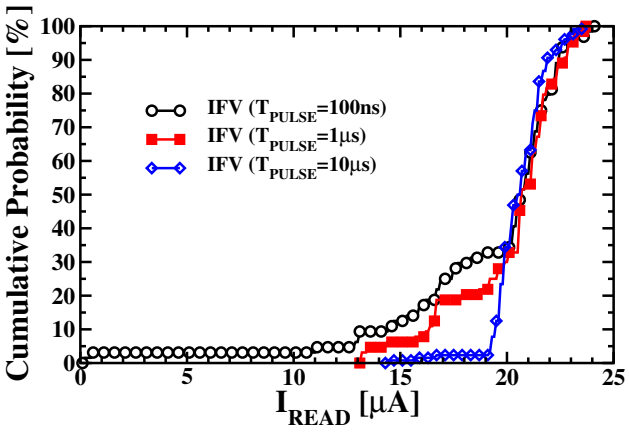


Figure 3.26: Cumulative I_{read} distributions after IFV with ΔV_{BL} = 0.01 V and different pulse widths.

Concerning the former factor, different pulse widths were tried for IFV forming scheme using ΔV_{BL} = 0.01V: 100 ns, 1 μs, and 10 μs. As shown in Fig. 3.26, the lower is the pulse width, the lower is the energy supplied to the cell for the forming operation, leading to poor forming yield and large inter-cell variability [70]. The trend of the results reported here applies also for IFV with larger ΔV_{BL} and for IF, and Pulse scheme. From the technology viewpoint, it must be reminded that the developed 4 kbits arrays feature a peripheral circuitry that drives and routes all the signals on the memory cells through large multiplexers and selectors that, along with the process-induced variability, limits the duration of pulse width on a narrow range of values. However, the results shown in this section and the experimental methodology, especially concerning the IFV scheme, apply without lack of generality to any optimized RRAM technologies, where shorter pulse widths can be adopted.

The cell current I_{read} was measured by applying V_{BL} = 0.2 V, V_{WL} = 1.4 V and a read time T_{read} = 10 μs, $t_{rise} = t_{fall}$ = 1μs. All operations are performed by applying the V_{SL}, V_{BL} and V_{WL} cell-by-cell, sequentially. Cell forming time and yield obtained with each method are reported in Tab. 3.3. Forming yield is calculated as the cell percentage that shows a read verify

current after forming I_{read} >19 μA, ensuring the creation of a conductive filament.

Table 3.3: Forming Methods Timings and Yield.

	Time [μs] average	Time[μs] worst case	Yield [%]
Pulse	12	12	54
IF (ΔV_{BL} = 0.1 V)	180	180	77
IFV (ΔV_{BL} = 0.1 V)	216	360	87
IFV (ΔV_{BL} = 0.01 V)	1584	3600	99

The choice of 19 μA both as a read current threshold criterion to assess and claim the actual creation of a conductive filament during forming, and as a forming yield criterion is ascribed to stability and variability concerns that have been taken into account from previous results shown in literature [70, 77]. Concerning the stability, it is important to ensure that the creation process of the conductive filament in the HfO_2 will be utmost homogeneous, and that the so formed filament would be easily disrupted in consecutive Reset operations. As demonstrated in [77], if it is not adopted a sufficiently high read current threshold, large oscillations during forming and unstable Set/Reset behavior may appear during the memory cell lifetime. In order to account for the intrinsic variability of the RRAM technology exploited in this work, it must be ensured also that the average I_{read} calculated on the entire array is in the range between 18 μA and 24 μA. Forming with read current threshold lower than 18 μA generally produces unstable memory cells that will fail after few Set/Reset cycles [77], whereas forming with a threshold current higher than 20 μA usually display larger array inter-cell variability, as indicated in the values provided in Table 3.4. *IFV* average and worst case time (i.e. requiring the highest number of pulses) are reported since *IFV* forming time is different from cell to cell. Even if *Pulse* is the fastest, the very low yield result confirms that the energy provided by this forming technique is insufficient for most of the memory cells in the array. For this

reason this forming scheme will be no further considered in this study. The highest forming yield is obtained using $\Delta V_{BL} = 0.01$ V (about 99%), therefore only this *IFV* variant will be considered further on.

Table 3.4: Array average read currents and Inter-cell variability after forming with different thresholds.

Forming Threshold [μA]	Array average I_{read} [μA]	Array inter-cell variability [μA]
19	20.58	1.26
20	20.88	1.77

Fig. 3.27 shows I_{read} distributions of correctly formed cells (i.e. reaching the 19 μA target) for all the schemes considered in this work. As it can be observed, the distributions related to the *IFV* scheme exhibit a lower dispersion of the current values, thus resulting in a better control of the cell-to-cell variability.

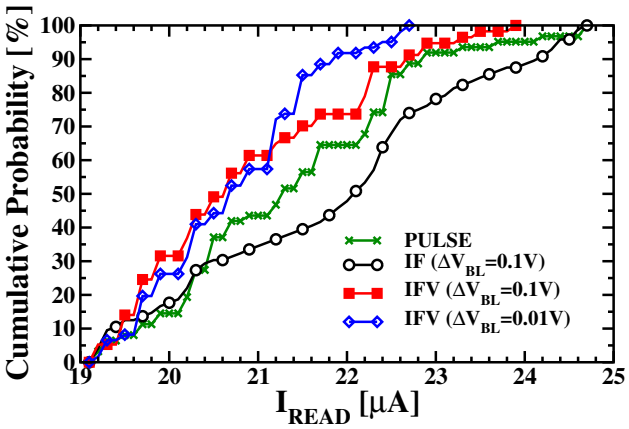


Figure 3.27: Cumulative I_{read} distributions of correctly formed cells after Pulse (54%), IF (77%), IFV with $\Delta V_{BL} = 0.1$ V (87%) and IFV with $\Delta V_{BL} = 0.01$ V (99%) forming operations.

3.5.2 Incremental Set/Reset with verify

The Set/Reset operations were performed by using an Incremental Step Pulse algorithm [66, 78, 79], by increasing V_{BL} from 1.5 V up to 3.5 V with ΔV_{BL} = 0.1 V, V_{WL} = 1.4 V and T_{pulse} = 10 μs, $t_{rise} = t_{fall}$ = 1μs during Set and by increasing the sourceline voltage V_{SL} from 1.5 V up to 3.5 V with ΔV_{SL} = 0.1 V, V_{WL} = 2.8 V and T_{pulse} = 10 μs, $t_{rise} = t_{fall}$ = 1μs during Reset. LRS and HRS read current cumulative distributions measured after the first Set and Reset operations are reported in Fig. 3.28. It can be observed that Reset failed on some *IF* formed cells (denoted as *Hard to disrupt*) showing read current values above the HRS threshold fixed to I_{read} =10 μA, whereas Reset operation has been correctly performed on all *IFV* formed cells.

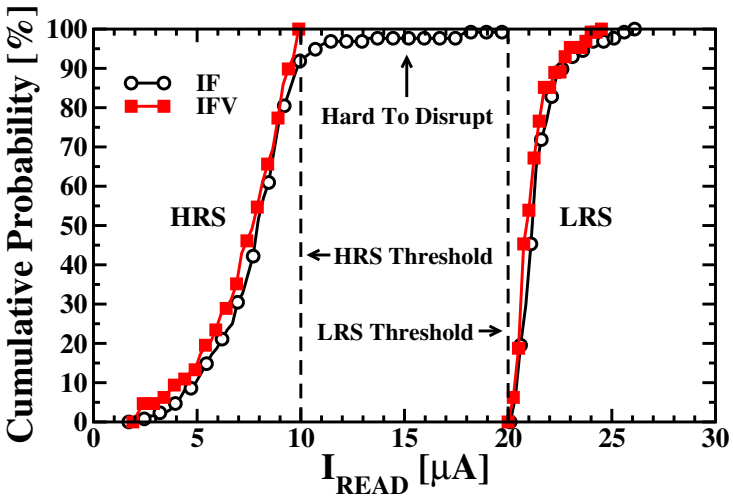


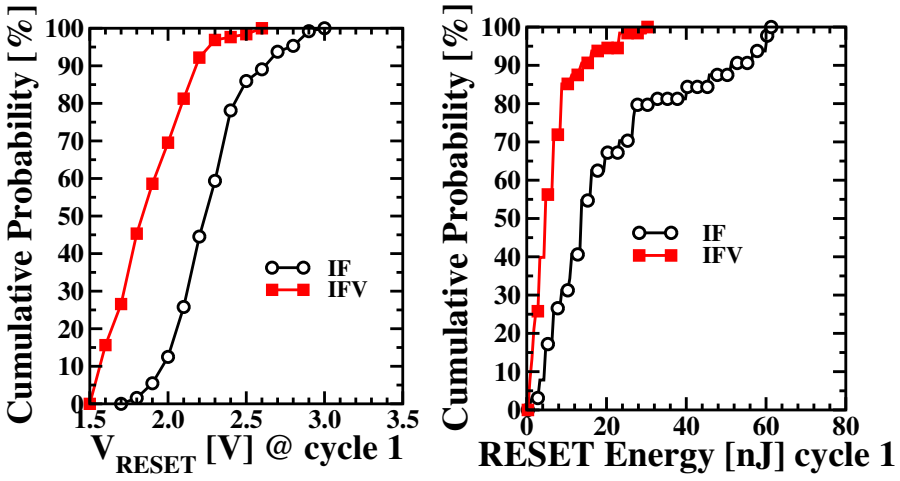
Figure 3.28: IF and IFV cumulative I_{HRS} and I_{LRS} distributions measured at endurance cycle 1.

Experimental results show that the *IFV* technique exhibits a lower Reset switching voltage and a reduced operation current: Fig. 3.29 (a) show the cumulative distribution of the reset switching voltages measured on fresh devices for *IF* and *IFV* (ΔV_{BL} = 0.01 V) forming schemes (no relevant variation was observed on Set switching voltage). The advantages are ascribed to a

higher filament geometry control devised by the *IFV* scheme. Fig. 3.29 (b) shows the cumulative distribution of the energy required to perform Reset operations. The overall energy required to disrupt the conductive filament has been calculated as:

$$E = \sum_{i=1}^{n} V_{pulse,i} * I_{pulse,i} * T_{pulse} + V_{read} * I_{read,i} * T_{read} \tag{3.1}$$

Where n is the number of Reset pulses applied during incremental pulse operation, $V_{pulse,i}$ is the pulse voltage applied at step i, $I_{pulse,i}$ is the current flowing through RRAM cell during pulse i application, $T_{pulse} = 10\mu s$ is the pulse length, V_{read} =0.2 V is the read voltage applied during verify operation, $I_{read,i}$ is the current read during read verify step i, and $T_{read} = 10\mu s$ is the verify pulse length.



(a) (b)

Figure 3.29: Cumulative distribution of the reset switching voltages (a) and overall Energy required to perform Reset operation (b) at cycle 1 for *IF* and *IFV* ($\Delta V_{BL} = 0.01$ V) forming schemes.

According to these results, the *IFV* advantages after forming are clear. As a matter of fact, as demonstrated in Fig. 3.25, each cell is formed with its proper filament geometry in order to achieve the 19 μA current, thus allowing

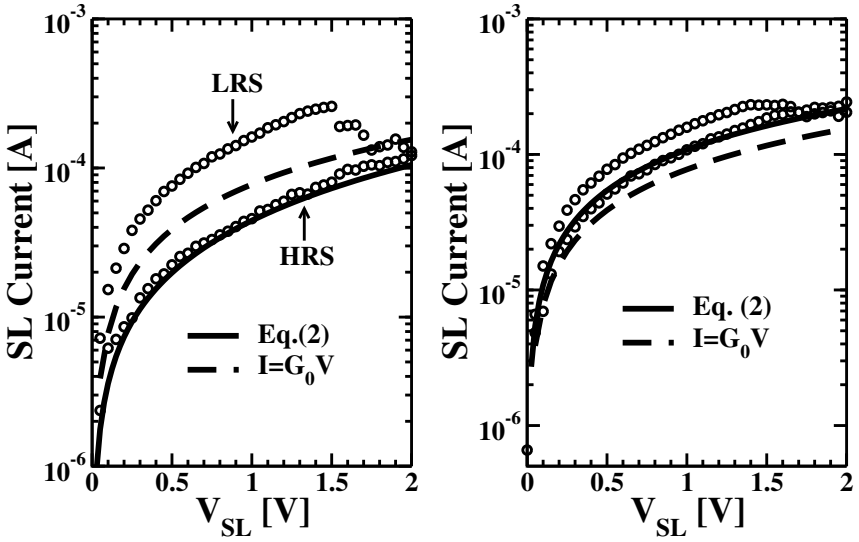
to counteract the intrinsic technological variability that would be impossible without a verify procedure.

3.5.3 Post-Forming modeling

Reset I-V characteristics have been measured after-forming to analyze the conductive filament properties through QPC model [52,67,80]. Two different behaviors can be observed for correctly working cells (a) and hard to disrupt cells (b), reported in Fig. 3.30. The differences in the HRS current values can be ascribed to the differences in the filament size [81]. The black dashed line shows the limit $I = G_0V$ with $G_0 = 2e^2/h$ the quantum conductance unit corresponding to the creation of a single mode nanowire, where e is the electron charge and h the Planck's constant. Within this framework, $I = G_0V$ sets a limit: in case of $I > G_0V$ more than a single conductive filament or a single filament with more than one mode must be taken into account: this means that on hard to disrupt cell a residual part of the conductive filament is still present after Reset operation since High Resistive State (HRS) curve measured is over the limit. HRS I-V curves fitting has been performed through QPC (lines). HRS current is calculated according to the expression:

$$I = \frac{2e}{h} G/G_0 \Big(eV + \frac{1}{\alpha} Ln \Big[\frac{1 + e^{\alpha(\Phi - \beta eV)}}{1 + e^{\alpha[\Phi + (1-\beta)eV]}} \Big] \Big) \tag{3.2}$$

where Φ is the barrier height (bottom of the first quantized level), $\alpha = t_B \pi^2 h^{-1} \sqrt{2m^*/\Phi}$ is a parameter related to the inverse of the potential barrier curvature (assuming a parabolic longitudinal potential), $m^* = 0.44m_0$ is the effective electron mass and t_B is the barrier thickness at the equilibrium Fermi energy. β takes into account how the potential drops at the two ends of the filament: since the constriction is highly asymmetric β=1 has been used (almost all the applied voltage drops close to the Ti layer). G/G_0 is a conductance parameter equivalent to the number of filaments at very low voltages: in a very approximate way, a single highly conductive filament can be viewed as a parallel combination of elementary nanowires [82].



(a) (b)

Figure 3.30: Reset I-V characteristics measured after forming and HRS fitting through eq.(3.2) on correctly working cells (a) and hard to disrupt cells (b).

The conductive filament obtained after Reset on correctly working cells and hard to disrupt cells is depicted in Fig. 3.31: in case of good Reset the presence of a potential barrier is assumed on HRS state, hence the average barrier length d and the radius of the constriction r have been calculated according to [52]. On the contrary, assuming the absence of a potential barrier on hard to disrupt cells, the normalized conductance of the filament G/G_0 has been calculated using large negative values for Φ in the above expression, since in such condition the barrier plays no role (neither β nor α affect the results) and a large negative barrier is a trick to eliminate the barrier effect.

Fig. 3.32 shows the HRS curves obtained after Reset I-V operation on *IF* (a) and *IFV* (b) formed cells. It can be observed that *IF* show higher conductance values than *IFV*: only 39% of *IF* formed cells showed HRS curves below $I = G_0V$ whereas 65% of *IFV* formed cells showed HRS curves below the limit. Even if such hard to disrupt cells percentages are very high, incremental step Reset algorithm allowed obtaining a strong reduction on both

IF (around 9%) and *IFV* cells (0%), as shown in Fig. 3.28. $I = 2 * G_0 V$ and $I = 0.8 * G_0 V$ are reported as upper and lower limit, respectively.

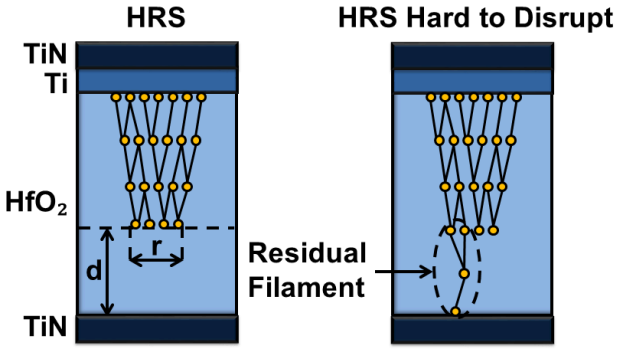
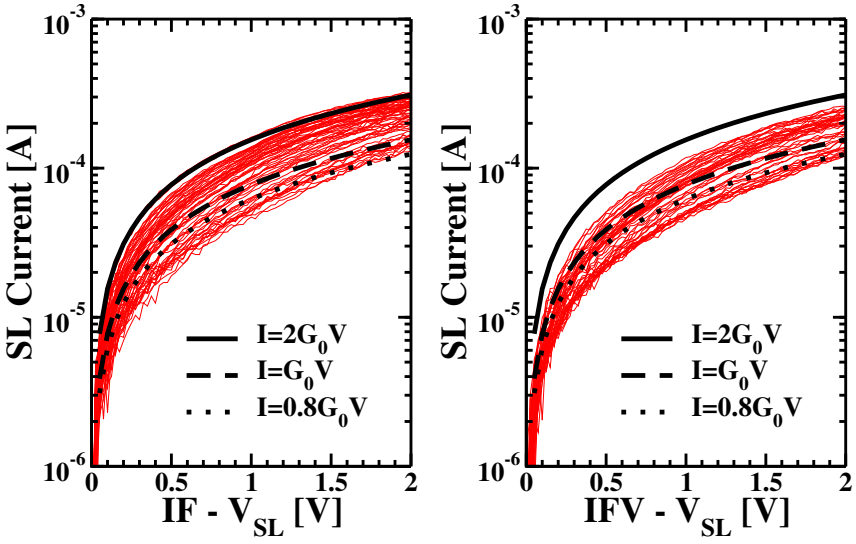


Figure 3.31: Schematic showing the conductive filament shape after Reset (HRS) for correctly working and hard to disrupt cells.

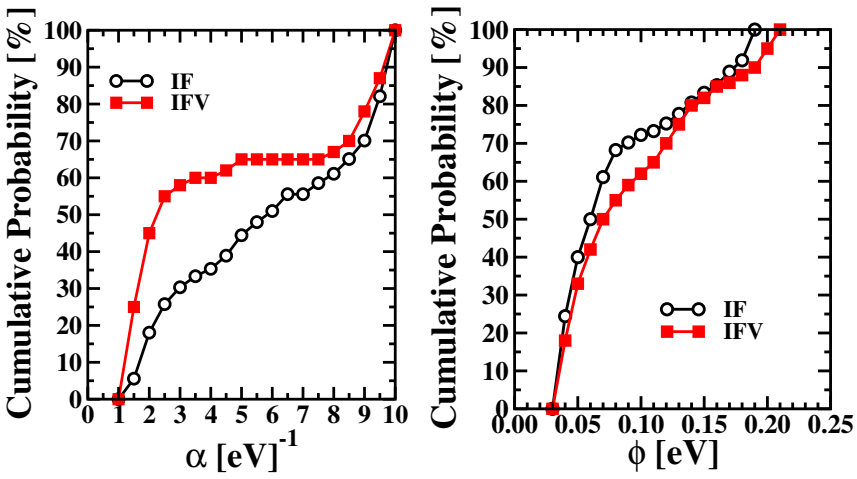


(a) (b)

Figure 3.32: HRS curves obtained after Reset I-V operation on *IF* (a) and *IFV* with $\Delta V_{BL} = 0.01$ V (b) formed cells.

In case of correct Reset operation (i.e. HRS curve below $I = G_0 V$) fitting has been performed assuming $G/G_0 = 1$ and the presence of a potential

barrier. The cumulative distributions of α and Φ fitting parameters used on correctly working cells are reported in Fig. 3.33.

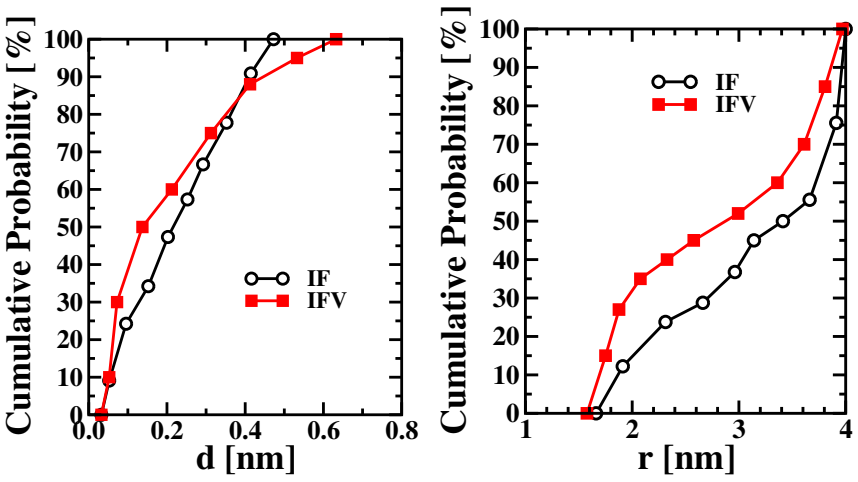


(a) (b)

Figure 3.33: Cumulative distribution of α and Φ fitting parameters used on correctly working cells.

The cumulative distributions of calculated barrier length d and radius r of the constriction for correctly working cells are reported in Fig. 3.34: *IFV* cells show smaller r. These values are sensitive to the effective mass, which is unknown with certainty, so that they should be considered for comparative purposes only. The barrier in HRS is very low for both forming methods so it only affects the low voltage part of the I-V curve, after that a linear behavior can be observed.

In case of hard to disrupt cells fitting has been performed assuming large negative Φ values, α fixed to 1 (even if α and Φ play no role in such condition) and $G/G_0 \geq 1$ due to the presence of the residual filament. Fig. 3.35 shows the cumulative distribution of G/G_0 conductance values fitting parameters used on hard to disrupt cells: it can be observed that *IF* hard to disrupt cells resulted in higher conductance values.



(a) (b)

Figure 3.34: Cumulative distribution of calculated barrier length d (a) and radius of the filament constriction r (b) on correctly working cells.

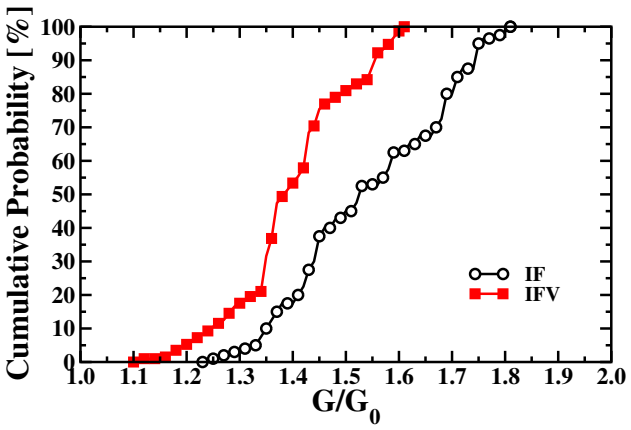


Figure 3.35: Cumulative distribution of G/G_0 fitting parameters used on hard to disrupt cells.

3.5.4 Endurance analysis

To quantify the advantages obtained through *IFV* forming during lifetime, 2k endurance cycles have been executed. Fig. 3.36 plots the *IF* and *IFV* average and minimum read windows, calculated as in [50], as a function of the Set/Reset cycles. The advantages of the *IFV* scheme are even appreciable

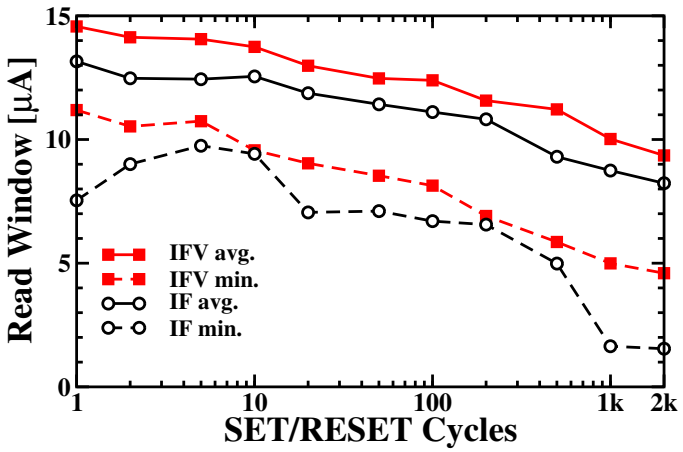


Figure 3.36: Average and minimum read windows as a function of the Set/Reset cycle number for *IF* and *IFV* ($\Delta V_{BL} = 0.01$ V) forming schemes.

during cycling: as far as the average criterion is considered, the average read window gain during cycling of the *IFV* scheme on the *IF* method sets around 7%, representing a marginal yet non negligible advantage. When the minimum criterion is considered, the average read window gain during cycling increases up to 37%. This represents once again a plus for the *IFV* scheme since it demonstrates its enhanced ability in Set/Reset tail bits (i.e., cells harder to be switched) reduction [83]. The read window closure due to endurance degradation [84] could be attributed to the impact of impurities in the metal-organic AVD precursor, in particular to carbon [63].

The ultimate advantage of the *IFV* is shown in Fig. 3.37, which exhibits the cumulative number of Set/Reset pulses applied to the entire memory array during the 2k cycles experiment as a function of the cycle number for the *IF* technique and the relative saving that can be obtained with *IFV*. The saving in terms of the total number of Set/Reset pulses during cycling is mainly due to a lower Reset switching voltage required by cells formed with *IFV*. All the described advantages translate, both during the design stage of larger arrays such as [85, 86] and at a system level, in shorter switching/operating times, less operative energy for the Reset operation, and in lower power consumption for the circuitry driving either BLs or SLs.

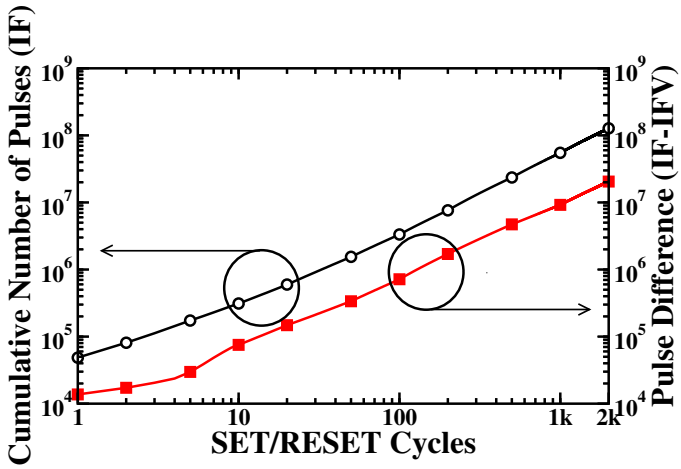


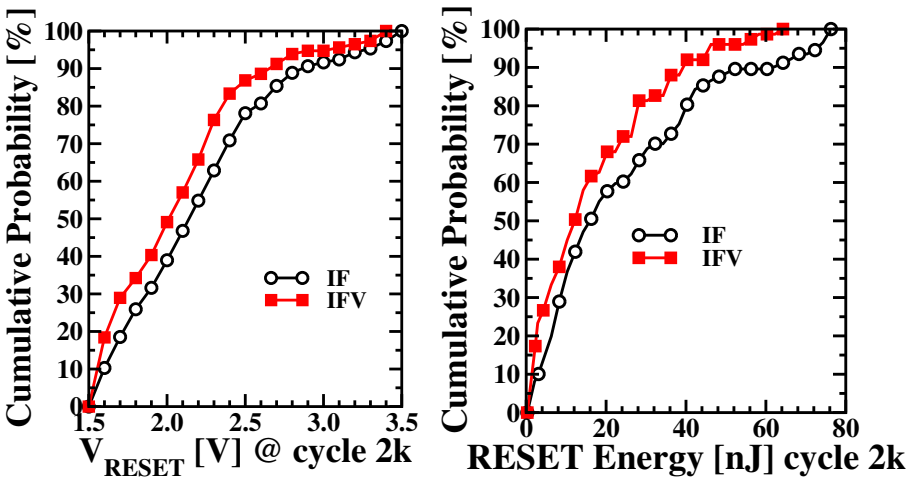
Figure 3.37: Cumulative number of Set/Reset pulses applied to the entire memory as a function of the cycle number for the *IF* forming scheme (left axis) and cumulative pulse number saving for the *IFV* (ΔV_{BL} = 0.01 V) forming scheme with respect to *IF* (right axis).

Experimental results show that even after 2k cycles *IFV* technique exhibits a lower Reset switching voltage and a reduced operation current: Fig. 3.38 (a) show the cumulative distribution of the reset switching voltages for *IF* and *IFV* (ΔV_{BL} = 0.01 V) forming schemes. Fig. 3.38 (b) shows the cumulative density function of the energy required to perform Reset operations: while the advantages after forming are clear, the energy gap is reduced during cycling because of device degradation. Fig. 3.39 shows the difference between *IF* and *IFV* average energy (a) and time (b) required to perform Set and Reset operations during cycling, calculated as:

$$\Delta E_{Reset/Set} = E_{IF,Reset/Set} - E_{IFV,Reset/Set} \tag{3.3}$$

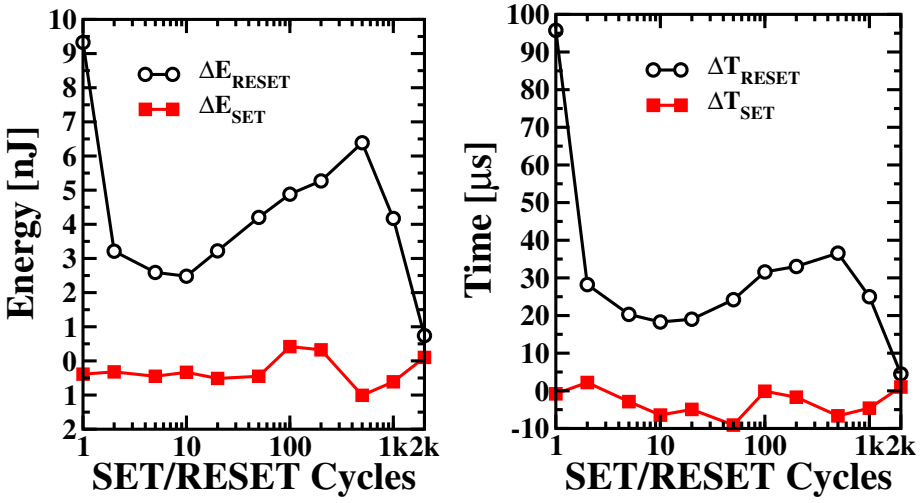
$$\Delta T_{Reset/Set} = T_{IF,Reset/Set} - T_{IFV,Reset/Set} \tag{3.4}$$

Set energy has been calculated as previously reported in Eq. 3.1 for Reset operation. Reset operation requires a higher number of pulses compared to



(a) (b)

Figure 3.38: Cumulative distribution of the reset switching voltages (a) and overall Energy required to perform Reset operation (b) at Set/Reset cycle 2k for *IF* and *IFV* ($\Delta V_{BL} = 0.01$ V) forming schemes.



(a) (b)

Figure 3.39: Average difference between *IF* and *IFV* ($\Delta V_{BL} = 0.01$ V) Energy (a) and Time (b) required to perform Set and Reset operations as a function of the Set/Reset cycle number.

Set, resulting in higher energy and time constraints. Even if *IFV* Set energy and time requirements are shown to be slightly higher, the advantages in terms of Reset energy and timing are clear until 2k cycles. After that, the device degradation reduces *IFV* advantages. It is worth mentioning that the main drawback of the *IFV* forming technique could lie on the longer forming times for cells requiring high forming voltages compared to the *IF* scheme, mainly because of the verify operation between the forming steps. However, it must be pointed out that this operation is performed just once, therefore its latency increase is favourably traded with the time saved during the subsequent Set/Reset operations.

3.5.5 Forming Oscillations analysis

In this study, the incremental pulse forming with verify algorithm is used to monitor the cells behavior during forming. Such technique allows recognising different cells behavior during forming in terms of read-verify current oscillations: the impact of these oscillations, interpreted either as the charging of a trap close to the surface of the conductive filament (CF) or the movement of an atom/defect in the filament [87], has been investigated in terms of reliability and cell-to-cell variability during 1k endurance cycles and 100k stress pulses in different cycling conditions.

Forming operation has been performed using a pulse-verify scheme: a sequence of increasing voltage pulses from 2V to 3.5V with $V_{step} = 0.01V$, $T_{step} = 10\mu s$ is applied on the BL with a WL voltage $V_{WL} = 1.4V$ to set the forming current compliance and after every pulse a read-verify operation with $V_{read} = 0.2V$, $T_{read} = 10\mu s$ is performed. When the read current reaches $I_{target} = 20\mu A$ the forming operation is stopped. Incremental pulse scheme with verify has been also implemented for Set and Reset operations, increasing V_{set} pulses on the BL and V_{reset} pulses on the SL from 1 V to 3.5 V with $V_{step} = 0.05V$, $T_{step} = 10\mu s$, $V_{WL,set} = 1.4V$, $V_{WL,reset} = 2.8V$ and the same read-verify condition used in forming. $T_{rise} = T_{fall} = 1\mu$s have been used on all pulses in order to avoid overshoot issues. Set operation is stopped

on a cell when the read-verify current reaches $I_{target} = 20\mu A$, whereas Reset is stopped when $I_{target} = 10\mu A$ ensuring a minimum resistance ratio of two. Set and Reset BL/SL voltages necessary to reach the requested read-verify current targets are defined as V_{SET} and V_{RES}. I_{LRS} and I_{HRS} are defined as the read currents measured after Set and Reset operations, respectively.

Three different behaviors observed during forming process are reported in Fig. 3.40: while in many cells the read-verify current shows a sudden increase due to the creation of the CF, there are some cells showing read current oscillations with different amplitudes during forming due to the charging of a trap close to the surface of the CF or the movement of an atom/defect in the filament. Oscillations generally appears after reaching the quantum conductance unit $G_0 = 2e^2/h$ corresponding to the creation of a single mode nanowire [52].

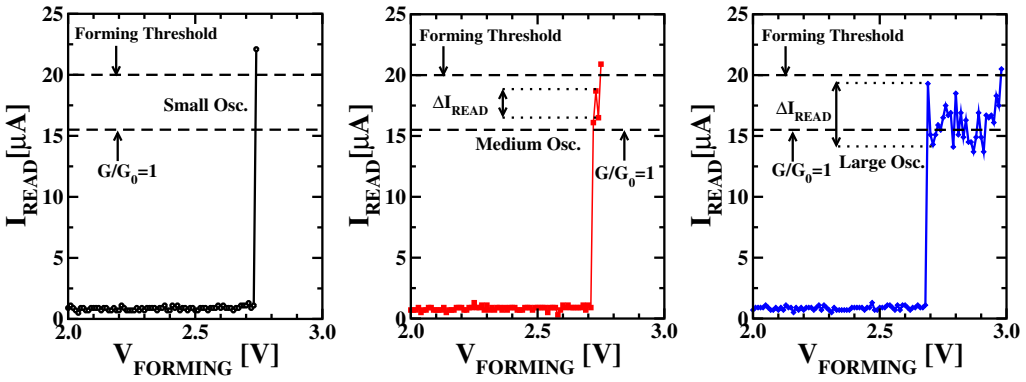
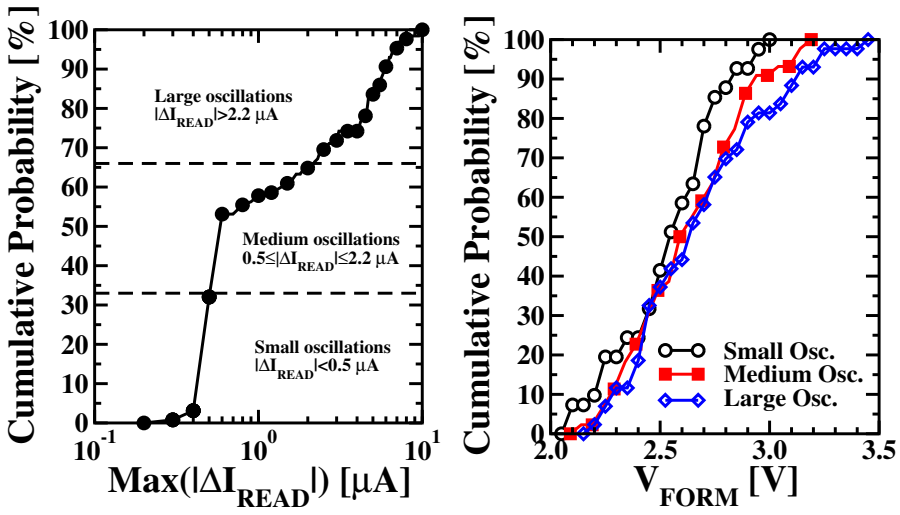


Figure 3.40: Three different behaviors observed during forming process: small (left), medium (centre) and large (right) read-verify current oscillations.

Fig. 3.41(a) shows the cumulative distribution of the maximum $|\Delta I_{read}|$ measured during forming. ΔI_{read} is the difference between two consecutive read verify steps after $G/G_0 = 1$ has been reached, where $G = I_{read}/V_{read}$. The cells have been arbitrarily gathered in three groups with the same amount of cells defined as follows, as a function of the maximum $|\Delta I_{read}|$ oscillation: small ($|\Delta I_{read}| < 0.5\mu A$), medium ($0.5 \leq |\Delta I_{read}| \leq 2.2\mu A$) and



(a) (b)

Figure 3.41: Cumulative distribution of the maximum read current oscillations measured during forming (a). Cumulative distribution of the forming voltage for the three forming oscillations groups (b).

large ($|\Delta I_{read}| > 2.2\mu A$). Fig. 3.41(b) shows the cumulative distributions of the forming voltages, defined as the voltage required to reach the read-verify target $I_{target} = 20\mu A$ during the incremental pulse and verify forming scheme. It can be seen that cells with lower forming voltages exhibit smaller current fluctuations.

To evaluate the endurance properties of the cells, 1k cycles have been performed through an incremental Set/Reset procedure: Fig. 3.42 shows the cumulative distributions of the Resistance Ratio, Set and Reset voltages calculated after cycling. Resistance ratio is calculated as I_{LRS}/I_{HRS} at $V_{read} = 0.2V$. The cells formed with smaller oscillations are shown to require higher V_{SET} and V_{RES} after 1k cycles: that means small oscillations correspond to wider filaments.

The Resistance Ratio, V_{SET}, V_{RES} average values and dispersion coefficients calculated during cycling are reported in Fig. 3.43. To evaluate the cell-to-cell variability the dispersion coefficient of I_{LRS} and I_{HRS} distributions, defined as (σ^2/μ), has been used. Resistance ratio of cells with large

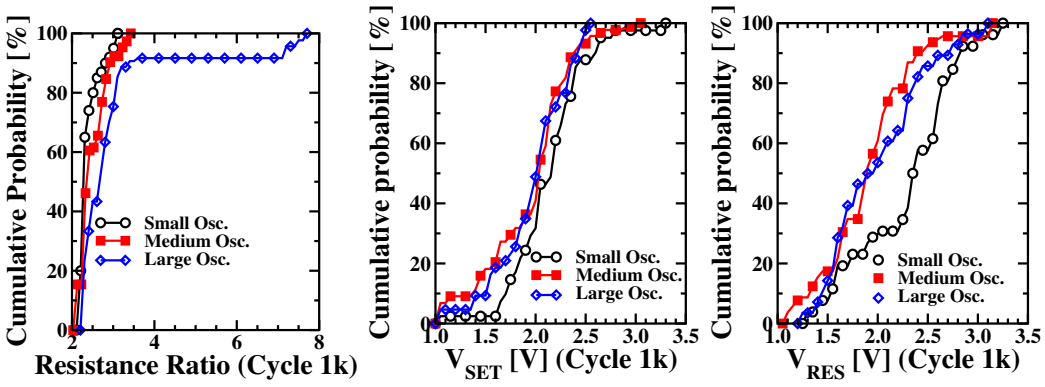


Figure 3.42: Resistance ratio, V_{SET}, V_{RES} cumulative distributions for the different forming oscillations groups calculated on cycled devices.

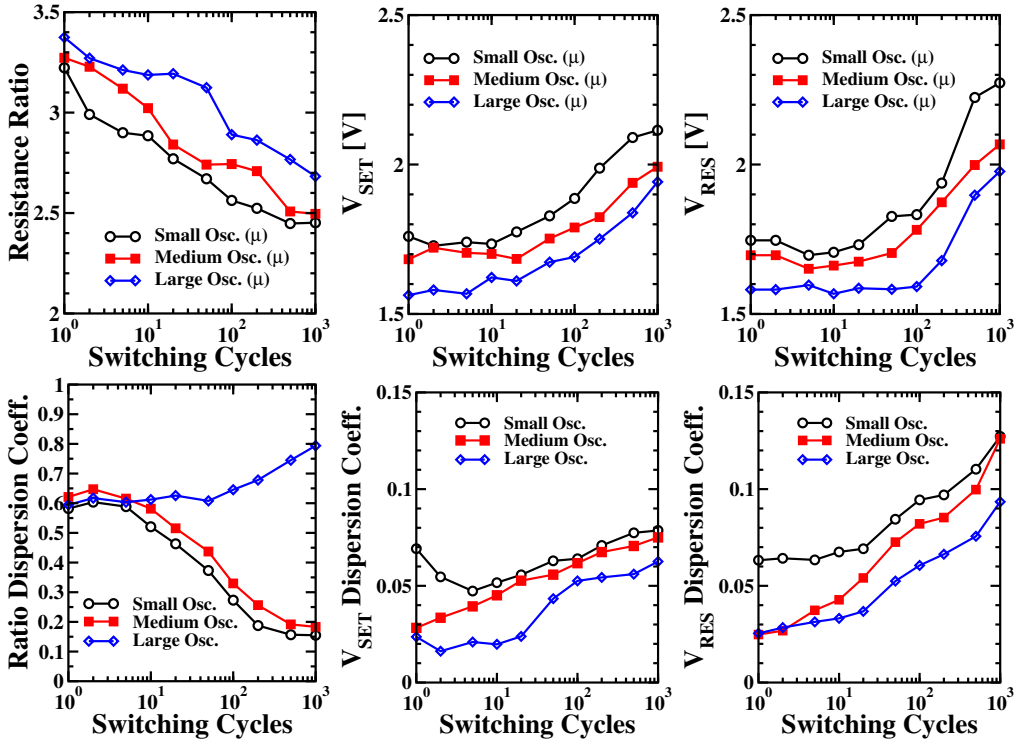


Figure 3.43: Resistance Ratio, V_{SET} and V_{RES} average values and dispersion coefficients calculated during cycling.

forming oscillations show both higher average value and dispersion coefficient in all cycling conditions: that means large fluctuations correspond to nar-

rower filaments. V_{SET}, V_{RES} average values and dispersion coefficients are shown to increase during cycling: switching voltages on cells formed with large oscillations show lower average values and dispersion in all cycling conditions. This indicates cells with lower V_{SET}, V_{RES} have a not fully developed filament: this explains the large fluctuations. One reason of the parameters dispersion could be the root mean square surface roughness of HfO_2 films due to the columnar structure of the TiN bottom metal electrode [57].

To evaluate the disturbs immunity of each cells group, 100k Reset stress pulses have been applied after Set with $V_{stress,res} = 0.8V$, $T_{stress,res} = 10\mu s$ and 100k Set stress pulses after Reset with $V_{stress,set} = 0.8V$,$T_{stress,set} = 10\mu s$ at different cycles. Set/Reset stress voltage pulses with 0.8V have been used since it's almost half of the average Set/Reset voltage measured on fresh devices. Cumulative distributions of the read currents measured after Reset (HRS), Set (LRS) and during Set and Reset stress on fresh devices are reported in Fig. 3.44 and Fig. 3.45, respectively: in both cases cells formed with larger current oscillations show a lower disturb immunity. That reveals larger fluctuations indicate a not so well formed filament thus more prone to exhibit lower immunity.

The average current variation and dispersion coefficient calculated on LRS cells during Reset stress are shown in Fig. 3.46: cells with larger current oscillations during forming show a higher variation of the average read current and dispersion coefficient during Reset stress in both cycling conditions. Similar consideration can be derived on Fig. 3.47, showing the average current variation and dispersion coefficient calculated on HRS cells during Set stress. Even if after 1k cycles cells show lower sensitivity to Set stress due to devices degradation, which means smaller current variation compared to fresh devices, cells with larger current oscillations during forming still show slightly lower disturbs immunity.

To provide a possible physical explanation of the measured phenomenon, Fig. 3.48 shows the distributions of the read-verify currents measured during forming with medium and large oscillations in units of G_0.

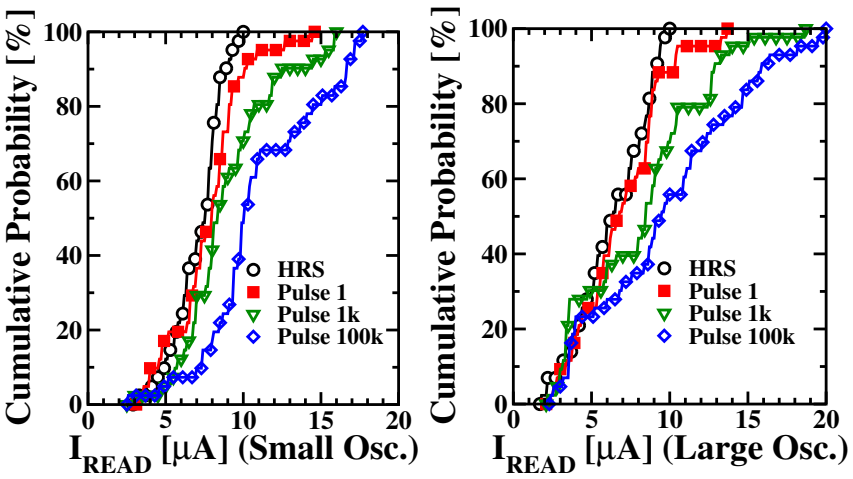


Figure 3.44: Cumulative distributions of the read currents (with $V_{read} = 0.2V$) measured during set stress on HRS after different number of disturb pulses, at endurance cycle 1.

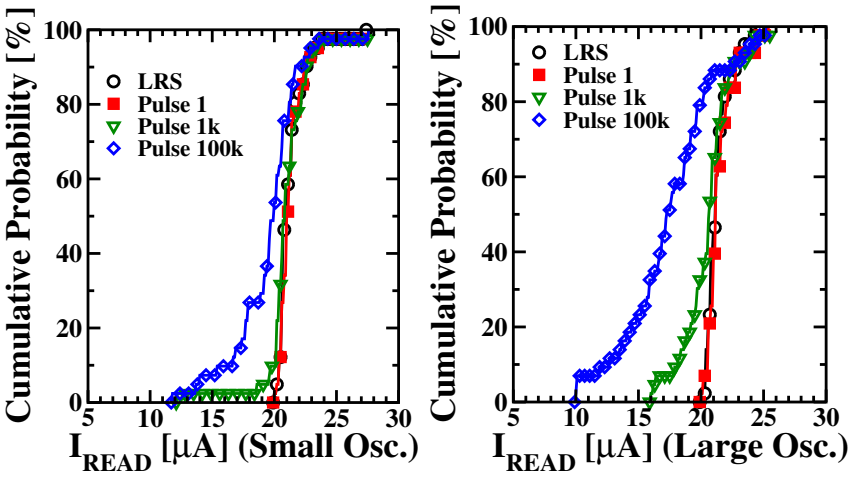


Figure 3.45: Cumulative distributions of the read currents (with $V_{read} = 0.2V$) measured during reset stress on LRS after different number of disturb pulses, at endurance cycle 1.

In order to evaluate only oscillations observed after the creation of the conductive filament, the analysis has been performed considering only the read currents values measured in the three verify steps performed after that

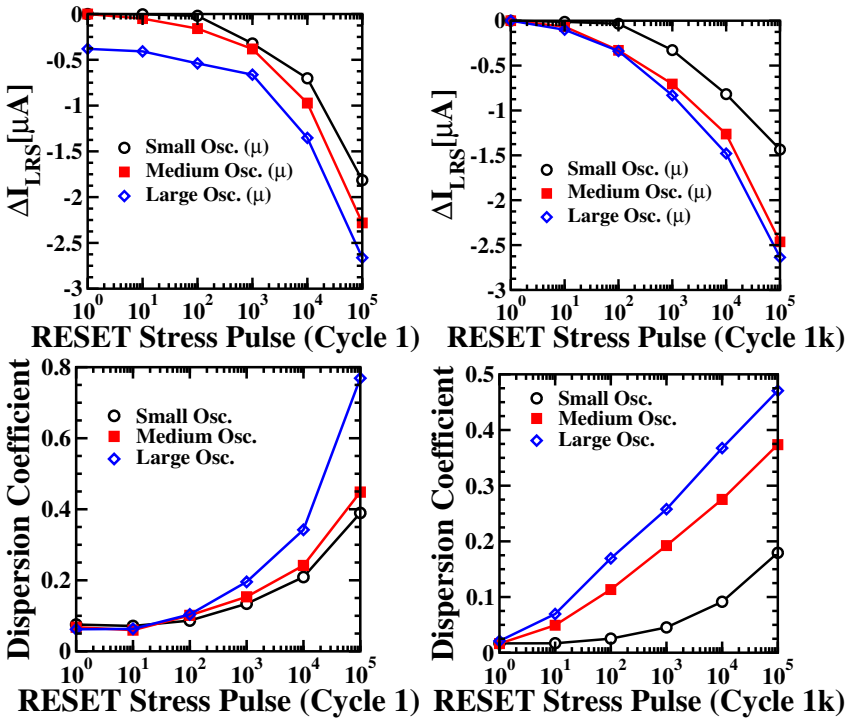


Figure 3.46: Average read current variation ($V_{READ} = 0.2V$) of LRS and dispersion coefficient evolution calculated during 100k reset pulse stress, with $V_{pulse} = 0.8V$ after endurance cycle 1 (left) and 1k (right).

$G/G_0 = 1$ has been reached. It can be observed that the distribution of medium read-verify current oscillations shows the highest peak in $G/G_0 = 1.22$, whereas large oscillation current distribution shows the highest peak in $G/G_0 = 1.02$ and the second highest peak $G/G_0 = 1.22$. This means that in this case we have two quasi-stable states: the one with the lower conductance corresponds to the narrowest constriction. This is consistent with the current magnitude observed in Fig. 3.40: the largest fluctuations corresponds to the lower current level in which the filament is formed by only a few atoms and that is why it is very sensitive. A schematic representation of the CF evolution during medium (1,2,3) and large (4,5,6) current oscillations is shown in Fig. 3.49: a narrow constriction is more prone to exhibit more fluctuations because it is formed by very few atoms/vacancies.

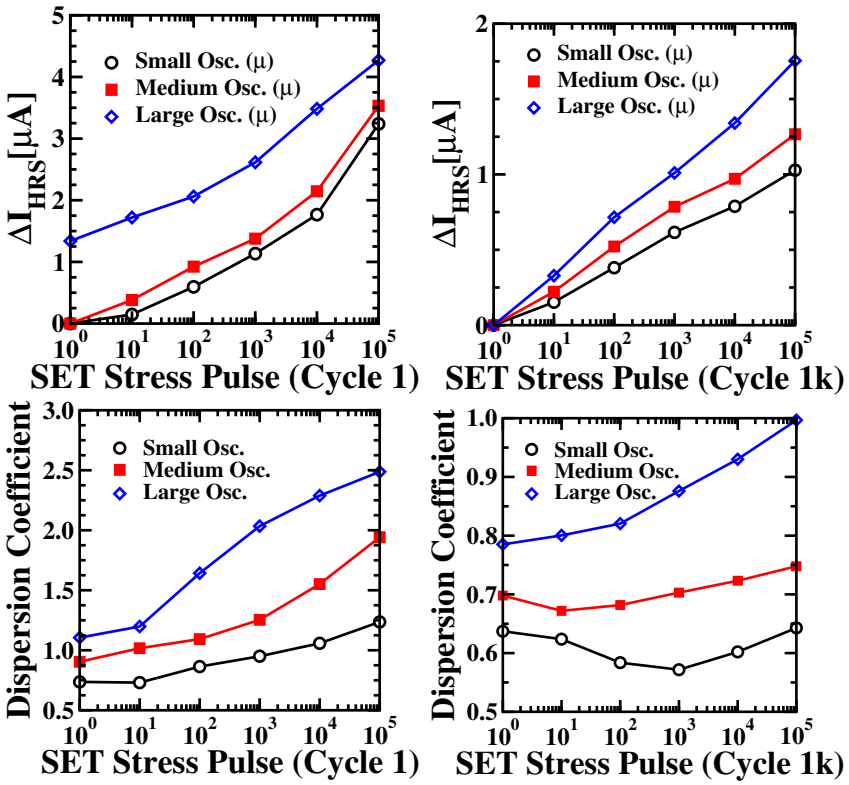


Figure 3.47: Average read current variation ($V_{READ} = 0.2V$) of HRS and dispersion coefficient evolution calculated during 100k set pulse stress, with $V_{pulse} = 0.8V$ after endurance cycle 1 (left) and 1k (right).

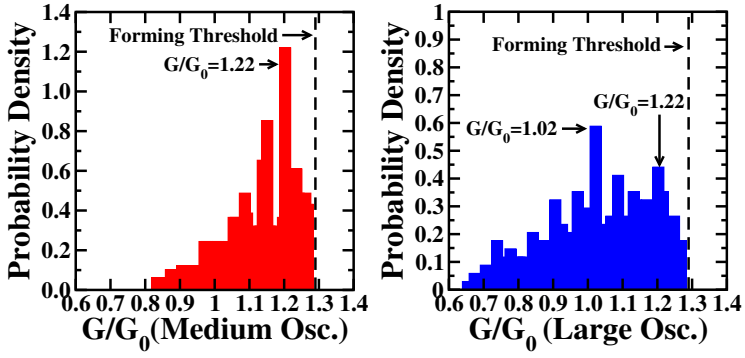


Figure 3.48: Distribution of the read-verify current values measured during forming process in units of G_0.

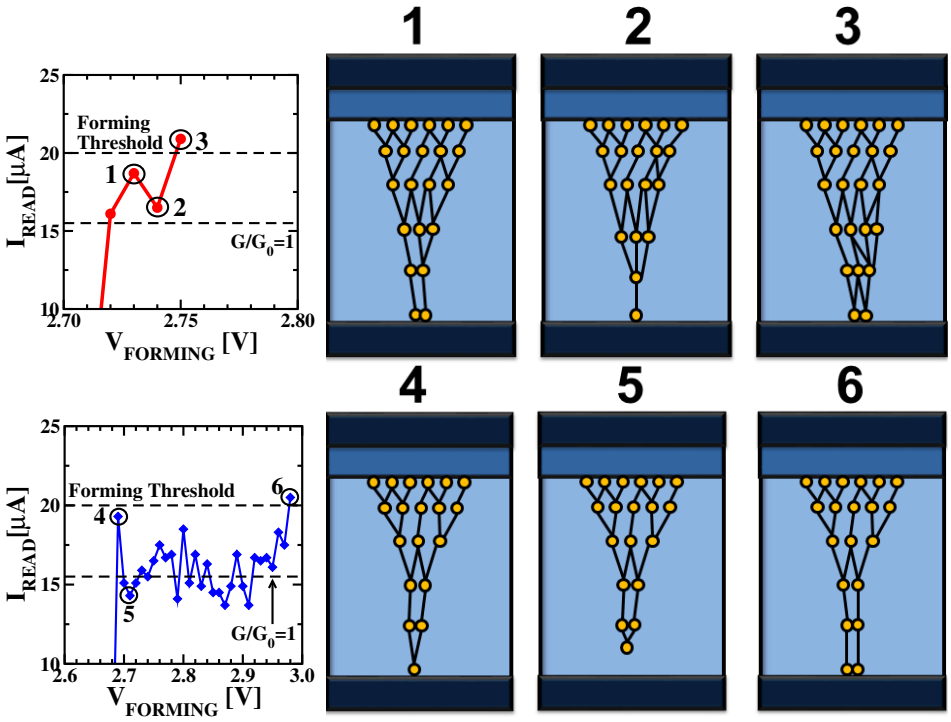


Figure 3.49: Schematic representation of the CF evolution during medium (1,2,3) and large (4,5,6) read-verify current oscillations.

Fig. 3.50 represents the post-forming conductive filaments with different current oscillation properties. Cells with narrowest CF show higher read-verify current oscillations during forming, lower disturbs immunity and higher variability but also better switching properties in terms of Set and Reset voltages and higher resistance ratio.

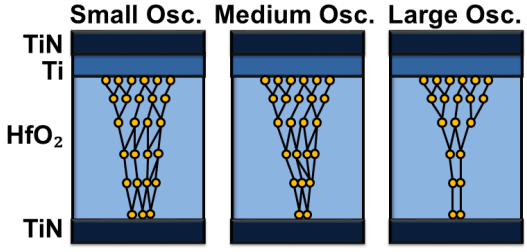


Figure 3.50: Schematic representation of the post-forming conductive filaments showing different current oscillation properties.

3.6 Process parameters impact on variability

The choice of a proper Metal-Insulator-Metal (MIM) technology for RRAM cells, exhibiting good uniformity and low switching voltages, is still a key issue for array structures fabrication and reliable electrical operation [88]. Such a process step is mandatory to bring this technology to a maturity level. In this section, both single cells and array comparison are performed. Firstly, a comparison between 1T-1R RRAM 4 Kbits arrays manufactured either with amorphous [50] or poly-crystalline [89] HfO_2 is reported in order to evaluate with large statistics how the material properties translate into device properties. After that, a comparison of 1T-1R cells obtained with 4 different processes is reported in order to deepen understand the impact of HfO_2 precursor and deposition temperature on cells performance and reliability.

3.6.1 Amorphous vs. Poly-crystalline HfO_2: array comparison

In this subsection, a comparison between amorphous and poly-crystalline HfO_2 RRAM arrays in terms of performance, reliability, Set/Reset operations energy requirements, intra-cell and inter-cell variability during 10k endurance cycles is reported [90, 91]. In amorphous HfO_2 the conduction mainly occurs through a conductive filament with a variable concentration of defects, whereas in poly-crystalline HfO_2 the conduction occurs only through grain boundaries with a very low defect concentration. The differences in terms of conduction properties and defect concentrations translate into different switching properties [88], with several implications on inter-cell variability (variations between cells) and intra-cell variability (cycle-to-cycle variations of any given cell).

To understand the relationship between the reliability properties observed during the endurance and read disturb tests and the conductive filament properties Quantum Point Contact (QPC) modeling [52] was used, since it allows to correctly represent the measured I-V characteristics independently

from the conduction mechanism. Even if the QPC allows to model the conductive filaments properties taking into account the cell-to-cell variability, it offers a technology description that sometimes is complex to be implemented in circuit simulation tools. To this extent, an equivalent circuit model able to offer a simpler description of the devices was applied and validated on both MIM technologies. The memory cells used in this work can be modelled using a diode-resistance equivalent circuit model. The model parameters extracted from the fittings of experimental I-V curves provide additional information about electrical properties of the memory cells to be exploited in the design of RRAM arrays.

The 1T-1R memory cells in the 4kbits arrays are constituted by a select NMOS transistor manufactured with a 0.25 μm BiCMOS technology whose drain is in series to a MIM stack. The wordline (WL) voltage applied to the gate of the NMOS transistor allows setting the cell current compliance. The variable MIM resistor is composed by 150 nm TiN top and bottom electrode layers deposited by magnetron sputtering, a 7 nm Ti layer, and a 8 nm HfO_2 layer deposited with two different Atomic Vapour Deposition (AVD) processes resulting either in amorphous (A) or poly-crystalline (P) HfO_2 films, respectively. The resistor area is equal to 0.4 μm^2. Amorphous films have been integrated also with a resistor area equal to 1 μm^2. This latter process option shows improved reliability and performance [49]. The Forming/Set/Reset operations on the arrays were performed by using an Incremental Pulse and Verify algorithm. The bitline (BL), sourceline (SL) and WL voltages applied during Forming, Set, Reset and Read operations are reported in Tab. 3.5. Reset operations were performed by applying the highest WL voltage available (2.8 V on array A and 2.5 V on array P) to maximize the cells switching yield while avoiding the breakdown of the MIM [77]. Pulses were applied during Forming by increasing V_{BL} with ΔV_{BL}=0.01V, whereas during Set and Reset ΔV_{BL}=0.1V and ΔV_{SL}=0.1V have been used, respectively. Each pulse featured a duration of 10μs, with a rise/fall time of 1μs to avoid overshoot issues. Set operation was stopped on a cell when the

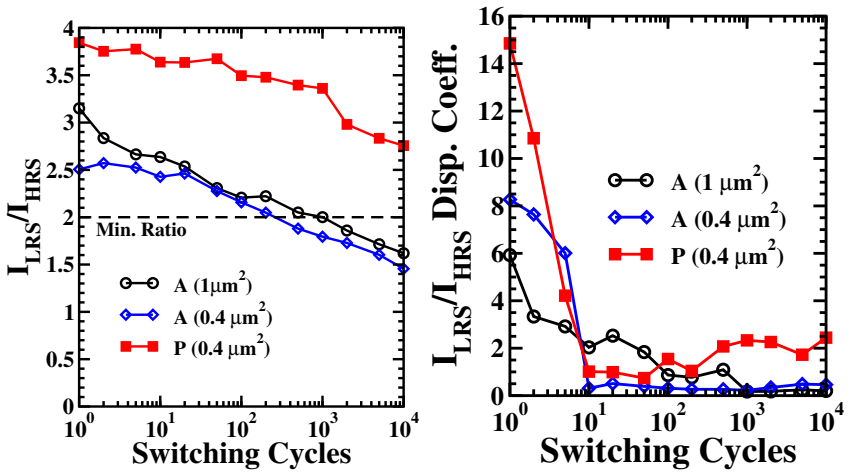
Table 3.5: Forming, Set, Reset and Read Voltage Parameters.

Operation	V_{SL} [V]	V_{BL} [V]	V_{WL} [V]
Forming	0	2-3.2	1.5
Set	0	0.2-3.2	1.5
Reset	0.2-3.2	0	2.5 (A)/ 2.8 (P)
Read	0	0.2	1.5

read-verify current reached 20μA, whereas Reset was stopped when 10μA was reached. Forming, Set and Reset BL/SL voltages necessary to reach the requested read-verify current targets are extracted from the characterization data and labelled as V_{FORM} ,V_{SET} and V_{RES}, respectively.

Arrays using A-HfO_2 (A-array) with resistor area of 0.4 μm^2, 1 μm^2 and P-HfO_2 (P-array) resulted in a Forming Yield (calculated as the cell percentage showing a read verify current after forming $I_{read} \geq 20\mu A$) of 58%, 90% and 95%, respectively. Fig. 3.51 shows the average current ratios between Low Resistive State (LRS) and High Resistive State (HRS) read currents (I_{LRS}/I_{HRS}), calculated on the entire cells population during Set/Reset cycling at V_{read} = 0.2V on A-array and P-array, and their relative dispersion coefficient. The minimum current ratio that allows to correctly discriminate between HRS and LRS, defined as I_{LRS}/I_{HRS} >2, is indicated for comparison [50]. The average ratios of A-arrays with resistor area of 0.4 μm^2 and 1 μm^2 go under the minimum ratio limit after 200 and 1k cycles, respectively. To evaluate the cell-to-cell variability the dispersion coefficient of I_{LRS} and I_{HRS} distributions, defined as (σ^2/μ), has been used. P-array showed higher Ratio ($\approx$ 2.8) even after 10k cycles, but also a higher dispersion coefficient after Forming (i.e., cycle 1). The grain boundaries conduction mechanism in the poly-crystalline HfO_2 structure could be the reason of the higher cell-to-cell variability in P-arrays [92]. A-array with resistor area of 1 μm^2 shows a slightly higher average ratio than A-array with resistor area of 0.4 μm^2.

Fig. 3.52 shows a comparison between I_{LRS} and I_{HRS} cumulative distributions measured at cycle 1 and after the endurance test: A-arrays show more

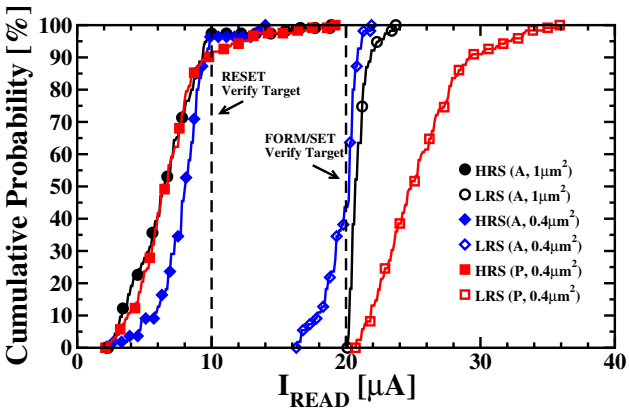


(a) (b)

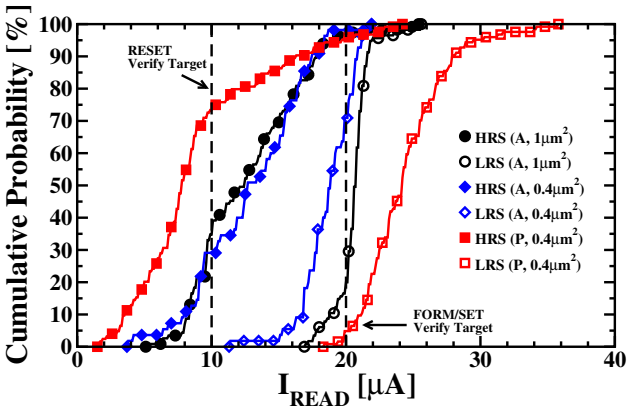
Figure 3.51: I_{LRS}/I_{HRS} current ratio average values (a) and dispersion coefficients (b) calculated during cycling.

compact distributions at cycle 1, however after the endurance test P-array shows a higher percentage of correctly switching cells reaching the Set/Reset verify targets. I_{HRS} cumulative distribution in P-array show a longer tail at cycle 1 compared to A-arrays. After 10k cycles only an increase of the tail in P-array can be observed whereas on A-arrays a strong shift of the distributions towards higher currents occurs, resulting in a higher number of cells not reaching the Reset threshold. I_{HRS} cumulative distribution in A-array with resistor area of 1 μm^2 shows lower currents at cycle 1 than A-arrays with resistor area of 0.4 μm^2, however after 10k cycles I_{HRS} cumulative distributions are very similar. In I_{LRS} cumulative distributions a tail creation of cells not able to reach the Set threshold can be observed on P-arrays after 10k cycles, whereas on A-arrays a strong shift of the distributions towards lower currents occurs, resulting in a higher number of cells not reaching the Set threshold especially when cells with resistor area of 0.4 μm^2 are considered. A-array with resistor area of 0.4 μm^2 shows a high number of cells not reaching the Set threshold even at cycle 1.

Fig. 3.53 shows the average Set and Reset switching voltages (V_{SET},



(a)



(b)

Figure 3.52: I_{HRS} and I_{LRS} cumulative distributions at cycle 1 (a) and at cycle 10k (b).

V_{RES}) and their relative dispersion coefficients: lower V_{SET} and V_{RES} are required on P-array which shows no variations during the endurance test, whereas V_{SET}, V_{RES} increase on A-arrays during cycling. V_{RES} on P-array shows the highest variability. A-arrays show similar behavior of the average V_{SET} and V_{RES} (a lower average V_{SET} is observed on A-array with larger resistor area only up to 500 cycles), while a higher V_{SET} and V_{RES} dispersion can be observed in A-array with smaller resistor area.

Fig. 3.54 shows the cumulative distributions of Forming, Set and Reset switching voltages at cycle 1 and after the endurance test: Forming, Set and

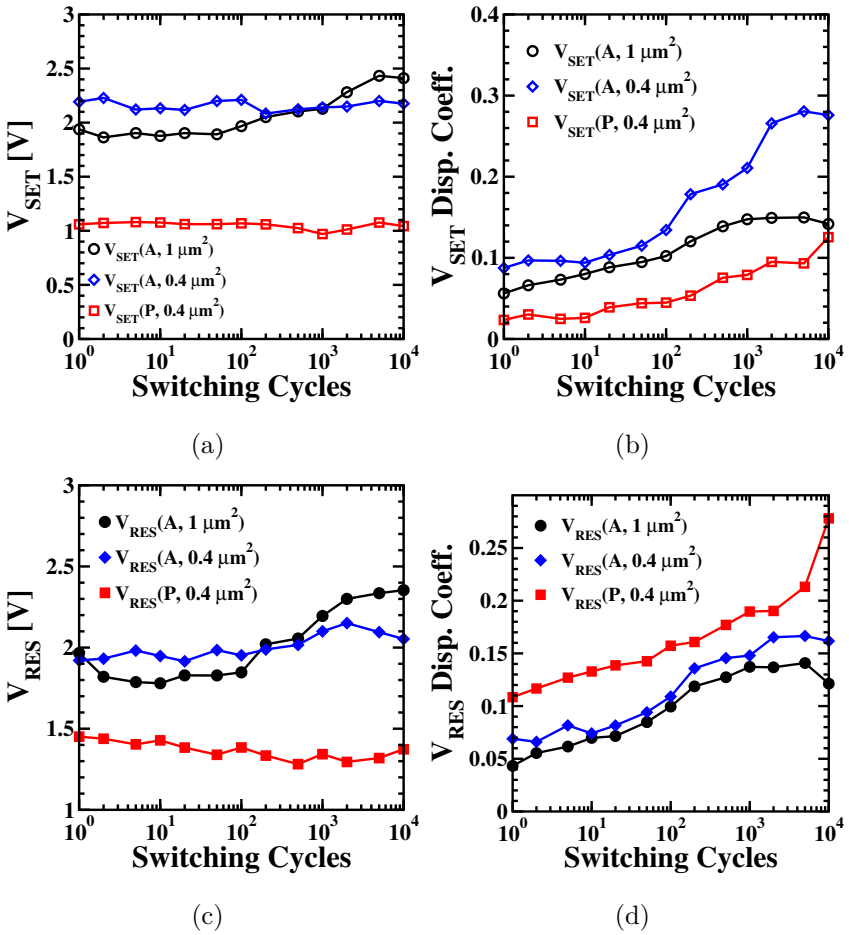
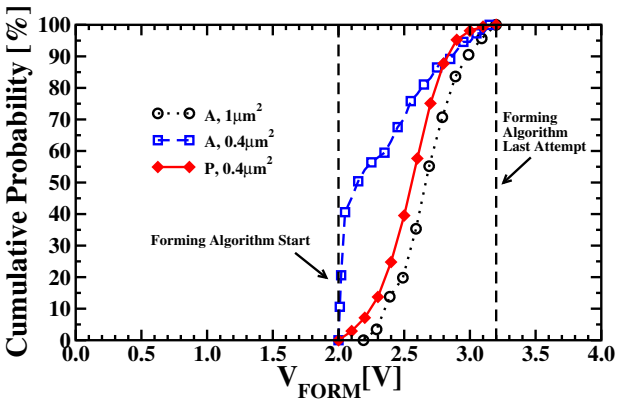


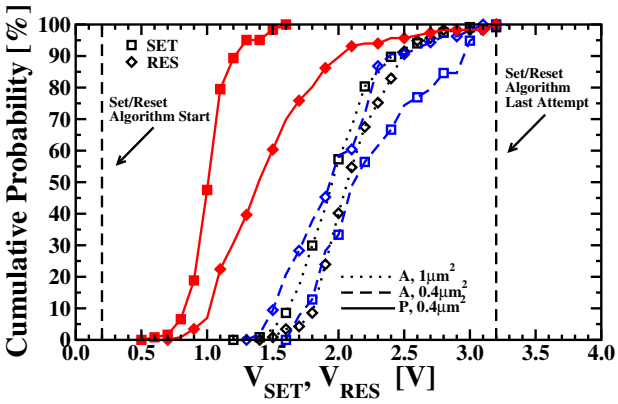
Figure 3.53: V_{SET} and V_{RES} average values (a,b) and dispersion coefficients (c,d) calculated during cycling.

Reset algorithms starting point and last attempt are indicated, corresponding to the first and the last voltage pulse available in the incremental pulse and verify procedure. P-array requires lower V_{SET} and V_{RES} but higher V_{FORM} if compared to A-array with the same resistor area. A-array with larger resistor area requires higher V_{FORM}, moreover it can be observed that $\approx 40\%$ of the devices with smaller resistor area reached the forming threshold at V_{FORM}=2 V, corresponding to the first attempt of the Forming Algorithm.

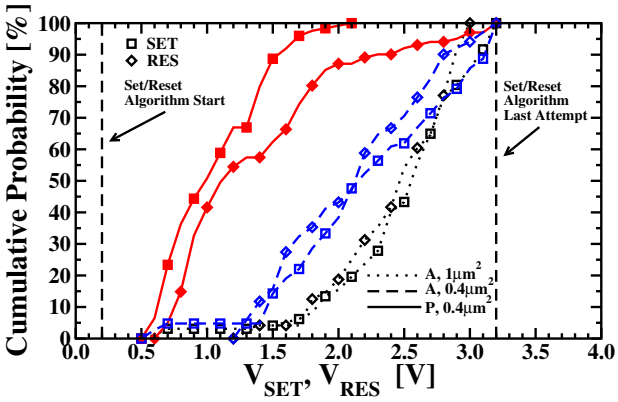
Since P-array shows a more compact distribution on V_{SET} and a larger V_{RES} than A-arrays, faster Set operation could be reliably used on P-array,



(a)



(b)

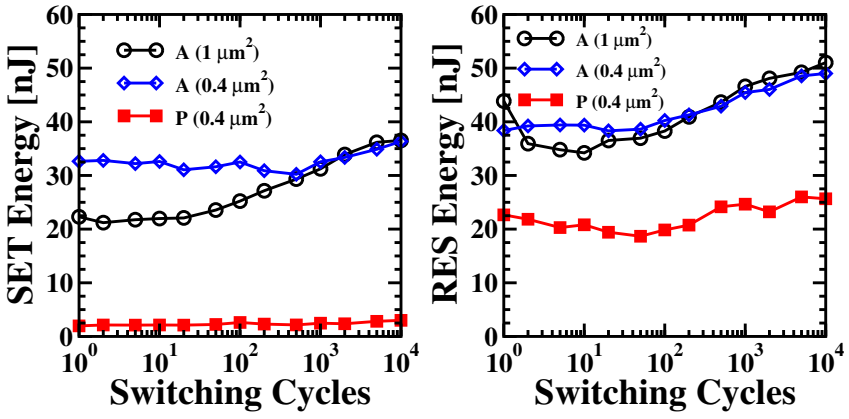


(c)

Figure 3.54: V_{FORM} (a), V_{SET} and V_{RES} cumulative distributions at cycle 1 (b) and at cycle 10k (c).

whereas on Reset an incremental pulse with verify technique is required to ensure good reliability. A-arrays show large distributions on both V_{SET} and V_{RES}, hence the adaptation of incremental pulse with verify techniques is mandatory on such arrays.

Fig. 3.55 shows the average energy required to perform Set and Reset operations on a single cell: P-array shows lower power consumption with a lower increase during cycling. A-arrays with different resistor area show similar power consumption during Reset operation, whereas a lower consumption during Set is observed on A-array with larger resistor area only up to 500 cycles. The overall energy required to create/disrupt the conductive filament during Set/Reset operations has been calculated as in Eq. 3.1.

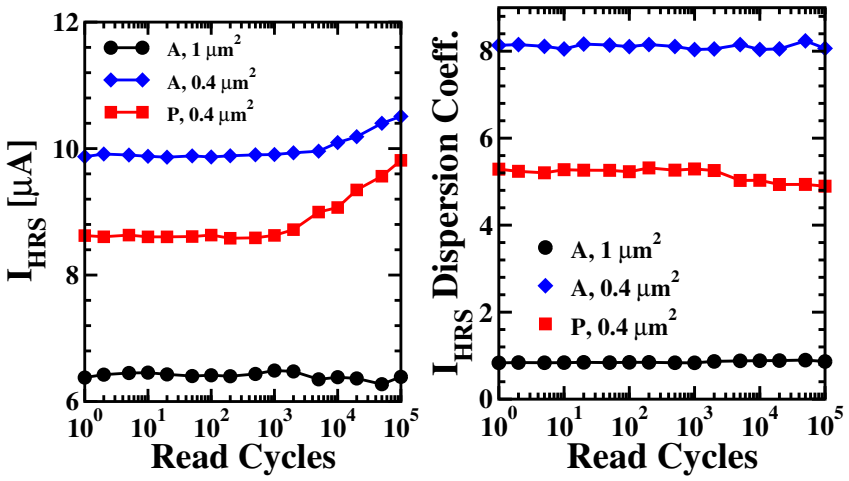


(a) (b)

Figure 3.55: Energy required to perform Set (a) and Reset (b) operations as a function of the Set/Reset cycle number.

In the considered RRAM cells the read signals has the same polarization of the Set operation (both pulses are applied on the BL), hence the read disturb could only be a problem on cells in HRS state since a very long sequence of read pulses could slowly re-create the conductive filament, resulting into an undesired switch from HRS to LRS [77]. Read disturb has been evaluated only on cells in HRS state for each considered technology: Fig. 3.56 shows the average HRS read current and its relative standard deviation measured

during 100k read operations. P-array shows the highest read current variation, confirming that on such technology due to the high leakage currents it is easier to create conductive paths.



(a) (b)

Figure 3.56: Average read current variation (a) and dispersion coefficient evolution (b) of HRS calculated during 100k read disturb pulses, with V_{pulse} = 0.2V.

3.6.2 1T-1R cells modeling

Extracting and modeling suitable parameters for the I-V characteristics is important to gather statistical information for any kind of non-volatile memory [93]. In the RRAM arrays of this work, I-V characteristics have been measured after-forming and modeled with two different approaches: in order to understand the differences on the conductive filament properties and variability QPC modeling has been used as in [94], while an equivalent circuit model [95] was used to obtain a description implementable in circuit simulation tools.

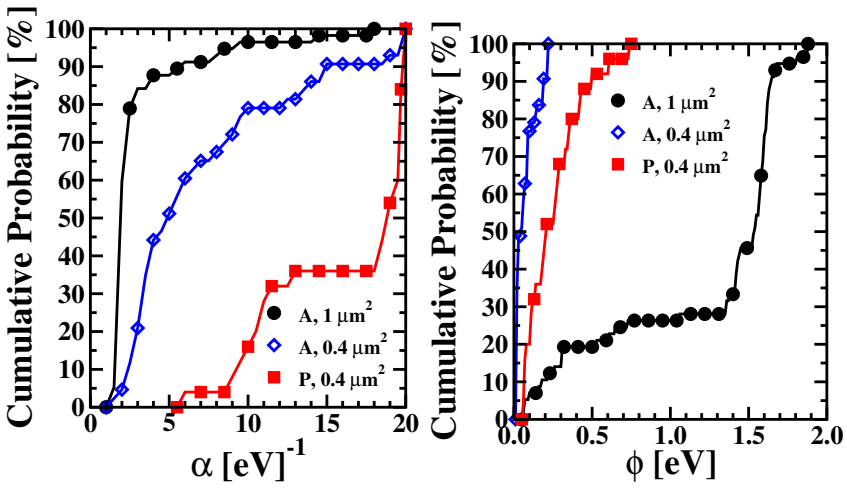
QPC modeling

Reset I-V characteristics measured after-forming were used to analyze the conductive filament properties through QPC model. HRS current is calculated according to the expression 3.2. I-V Reset operation has different impacts from cell-to-cell, resulting either into a break or a modulation of the conductive filament (CF) [70,94]. In the former case the presence of a potential barrier is assumed, hence fitting is performed considering $G/G_0 = 1$ and the average barrier length d and radius of the constriction r are calculated according to [52]. In the latter case, assuming the absence of a potential barrier, the normalized conductance of the filament G/G_0 is calculated. The percentage of cells resulting either into a CF break or modulation are reported in Tab. 3.6: the high leakage current in P-array makes very difficult to completely interrupt the conductive path hence the lowest percentage of CF break is obtained, whereas the highest percentage is obtained on A-array with the larger resistor area. The cumulative distributions of α and Φ fitting parameters calculated on the CF break cells are reported in Fig. 3.57. Average value and standard deviation of the fitting parameters are reported in Tab. 3.7.

Table 3.6: Reset condition comparison.

Technology	C.F. Break [%]	C.F. Modulation [%]
A, 1 μm^2	45	55
A, 0.4 μm^2	34	66
P 0.4 μm^2	20	80

The cumulative distributions of calculated barrier length d and radius r of the CF constriction are reported in Fig. 3.58, while the average value and standard deviation are reported in Tab. 3.8. A-array with the small resistor area shows the largest radius with the lowest barrier length: the presence of a very large constriction with a very low barrier explains the issues in controlling the cells' uniformity during Set and Reset operations. A-array with the larger resistor area shows higher barrier and smaller radius,



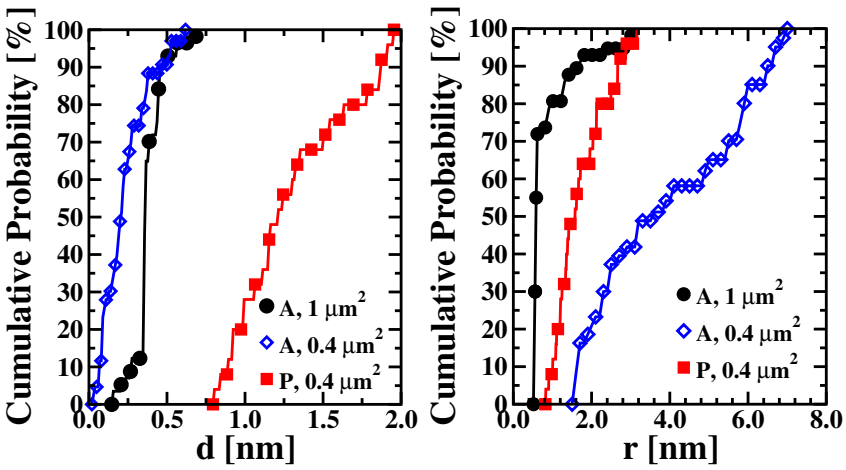
(a) (b)

Figure 3.57: Cumulative distribution of α and Φ fitting parameters used on CF break cells.

Table 3.7: Average and std. dev. of α and ϕ calculated on CF break cells.

Technology	$\alpha[eV]^{-1}$		$\phi[eV]$	
	avg.	std.	avg.	std.
A, 1 μm^2	2.67	3.02	1.21	0.58
A, 0.4 μm^2	6.83	5.48	0.07	0.06
P, 0.4 μm^2	16.08	5.06	0.17	0.25

resulting into a higher controllability during Set and Reset. Moreover, the highest parameters uniformity is observed, which translates into the highest HRS and LRS currents uniformity. A possible reason to explain the difference in the potential barrier between the two amorphous films is related to the defects concentrations in the HfO_2. Indeed, amorphous films integrated with lower area are affected by a higher defect concentration that eases the Reset process and therefore results in a lower potential barrier in the HRS. P-array shows the largest barrier with the highest variability: the highest barrier is the reason of the higher average ratio between HRS and LRS, while the high variability generates the high current variability observed in HRS.



(a) (b)

Figure 3.58: Cumulative distribution of calculated barrier length d (a) and radius of the filament constriction r (b) on CF break cells.

Table 3.8: Avgerage and std. dev. of d and r calculated on CF break cells.

Technology	$d[nm]$		$r[nm]$	
	avg.	std.	avg.	std.
A, 1 μm^2	0.37	0.11	0.85	0.62
A, 0.4 μm^2	0.25	0.24	4.25	2.25
P, 0.4 μm^2	1.29	0.36	1.71	0.65

In case of CF modulation fitting has been performed assuming large negative Φ values, α fixed to 1 (even if α and Φ play no role in such condition) and $G/G_0 \geq 1$ due to the presence of the residual filament. Fig. 3.59 shows the cumulative distribution of G/G_0 conductance values fitting parameters used on hard to disrupt cells: it can be observed that A-array with the larger resistor area shows the lowest variability, which is the reason of the lowest HRS current variability observed during Reset with the Incremental Pulse and Verify algorithm. Average value and standard deviation of the fitting parameter G/G_0 are reported in Tab. 3.9.

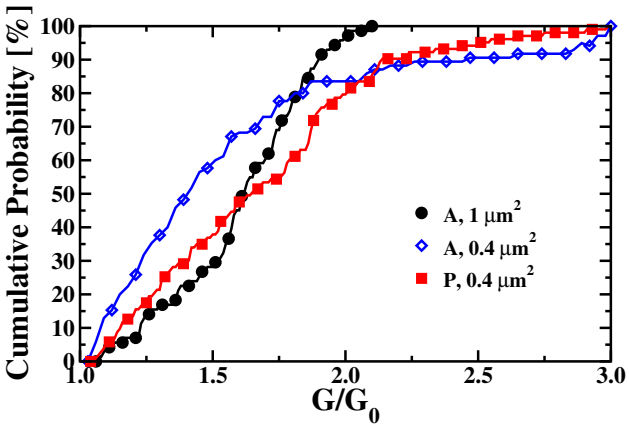


Figure 3.59: Cumulative distribution of G/G_0 fitting parameters used on hard to disrupt cells.

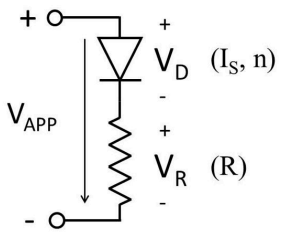
Table 3.9: G/G_0 average and std. dev.

Technology	G/G_0	
	avg.	std.
A, 1 μm^2	1.61	0.26
A, 0.4 μm^2	1.56	0.53
P, 0.4 μm^2	1.67	0.42

Equivalent circuit modeling

Electrical models are a powerful tool to analyze memory cells and circuits based on Resistive Switching (RS) devices allowing evaluating characteristics like power consumption or performance in large RS devices arrays [95, 96]. To model the experimental I-V curves during both RS states (i.e., LRS and HRS) a Diode-Resistor based circuit is used (Fig. 3.60) where the resistance (R), the diode saturation current (I_s) and diode ideality factor (n) are the parameters of the model [95]. V_{APP} represents V_{BL} or V_{SL}, which are the applied voltages to produce the Set and Reset processes respectively.

To fit all the experimental I-V curves an automatized process has been developed to extract the model parameters values (Tab. 3.10 and Tab. 3.11) for each curve. Fig. 3.61 shows some examples of experimental LRS I-V



$$I_R = \frac{V_{APP} - V_D}{R} = I_S \cdot \left(e^{\frac{V_D}{nV_T}} - 1 \right) = I_D$$

Figure 3.60: Equivalent model for the 1T-1R device based on a Diode-Resistor circuit. Resistance (R), saturation current (I_s) and ideality factor (n) of the diode are the parameters used to fit the conduction of the 1T-1R devices at both resistive states, LRS and HRS, and for both types of samples, amorphous and poly-crystalline.

Table 3.10: Model parameters of the amorphous samples I-V fittings for both states, LRS and HRS.

	Set		Reset	
	avg.	std.	avg.	std.
R[KΩ]	25.84	29.6	6.25	4.18
I_S [A]	9.75e-5	4.56e-5	6.39e-8	1.37e-7
n	8.95	7.06	5.76	5.19

curves (circles) before the Reset process and the simulated curves using the circuit model of Fig. 3.60 with suitable parameters (continuous lines). As can be observed, the model fits perfectly with the experimental results for both amorphous and poly-crystalline samples. For each kind of samples, the analyzed voltage range was limited by the Reset voltage that is lower for the poly-crystalline samples.

The same automatic process was also used to fit HRS I-V curves for both samples types. Fig. 3.62 shows experimental HRS curves (circles) before the Set process and the corresponding simulated curves (continuous lines).

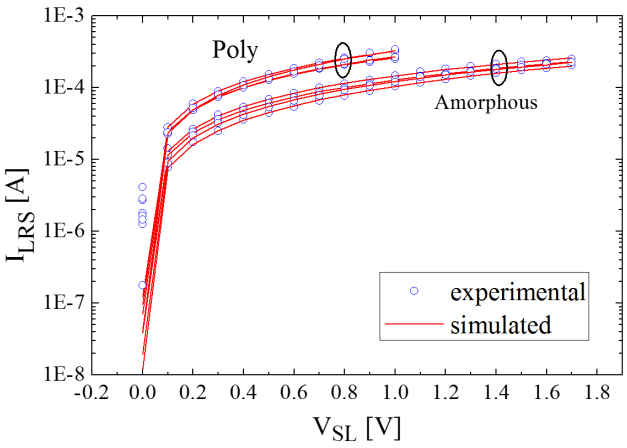


Figure 3.61: Experimental LRS I-V curves (circles) and the simulated curves (continuous lines) obtained using the Diode-Resistor model. With a suitable parameter Set, the model reproduces properly the experimental curves for both poly-crystalline and amorphous samples.

Table 3.11: Model parameters of the poly-crystalline samples I-V fittings for both states, LRS and HRS.

	Set		Reset	
	avg.	std.	avg.	std.
R[KΩ]	1.34	1.61	8.88	12.9
I_S [A]	5.67e-5	8.4e-5	1.52e-5	4.29e-7
n	3.95	5.03	1.18	2.35

Amorphous samples show very noisy I_{HRS} currents at low voltages (< 1 V) that could be caused by the nature of the memory cell and the array structure where the drive transistor effect on the electrical characteristics of the memory must be analyzed in detail. This noisy current must be neglected to avoid errors during the fitting process. For this reason, I_{HRS} values for V_{BL} below 1 V are not considered to force better fittings for voltages larger than 1 V, where the I-V curves are not affected by the noise. This consideration affects the obtained model parameters values (Tab. 3.10 and Tab. 3.11). Thus, the best fittings for the amorphous HRS I-V curves require a mean

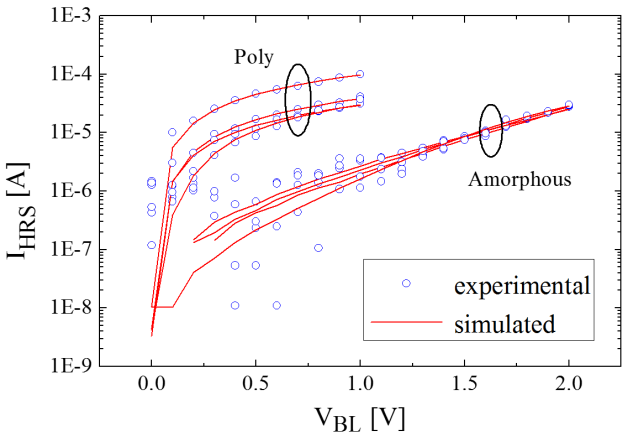


Figure 3.62: Experimental HRS I-V curves (circles) and simulated HRS curves (continuous lines) using the Diode-Resistor model for both amorphous and poly-crystalline samples. Noisy currents at low voltages cannot be fitted by the model, especially for the amorphous samples where current values for V_{BL} below 1 V are not considered.

R parameter value (6.25KΩ) which is lower than the one obtained for the LRS (25.84KΩ). This low value of the R parameter at HRS combined to the very low value of I_S at HRS (6.39e-8A) provides the best fitting between the experimental and the modelled curves.

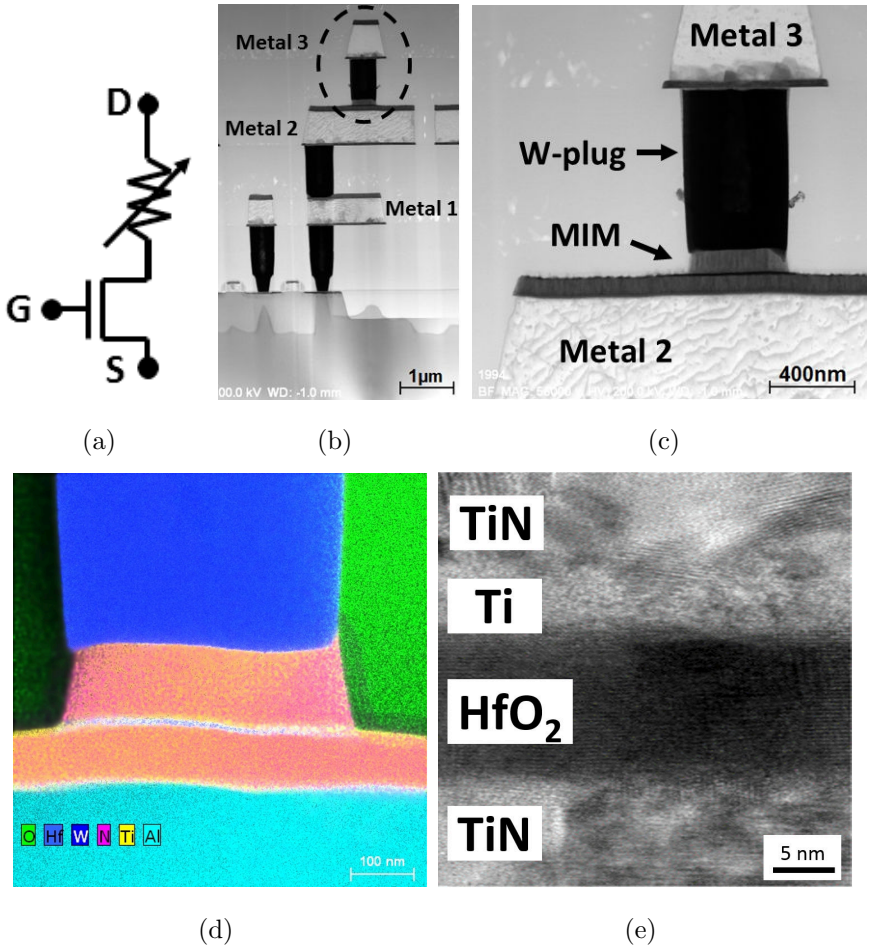
In conclusions, 1T-1R RRAM arrays manufactured with P-HfO_2 shows several advantages compared to A-HfO_2 even considering their improved process: higher current Ratio, lower switching voltages, lower power consumption, minor endurance degradation and higher overall yield. Moreover, P-array show very low V_{SET} variability, hence faster Set operation could be reliably performed. P-array disadvantages are represented by the larger HRS distribution after Forming, the higher Reset voltage dispersion, the lower read disturb immunity and the higher V_{FORM} if compared to A-array with the same resistor area, however it must be pointed out that such operation is performed only once. The grain boundaries conduction mechanism in the poly-crystalline HfO_2 structure could be the reason of the higher cell-to-cell variability observed in P-arrays. QPC modeling allowed showing that the

higher uniformity observed on A-array with the large resistor area can be ascribed to a lower conductive filament shape variability in terms of radius of the constriction and barrier height, whereas the P-array shows the highest variability in terms of conductive filament shape: the reason could be ascribed again to the different conduction mechanism and the higher leakage currents observed on such technology. A diode-resistor equivalent circuit model correctly fits the experimental RS I-V characteristics of poly-crystalline and amorphous samples for both LRS and HRS. However, noisy current levels at low voltages, especially for amorphous samples, could lead to a non-well fitted curve. Thus, it is needed to remove them for a suitable current fitting at larger voltages.

3.6.3 Impact of HfO_2 process on variability

In this work a comparison in terms of cell-to-cell variability and reliability of different HfO_2 Atomic Layer Deposition (ALD) processes on 1T-1R cells is performed. Cells behavior during Forming, Set and Reset is monitored through an incremental pulse and verify algorithm [66,77] in order to analyze the peculiarity of each cell in terms of the switching behavior activation and the process-induced inter-cell variability of the threshold voltages, on 100 cells for each process. To evaluate endurance properties, 100 endurance cycles have been performed on a cells subset, analyzing the impact on the threshold voltages and Resistance Ratio. Modeling of the HRS obtained after Forming has been performed through the Quantum-Point Contact (QPC) model [52, 67, 94] to link technology process characteristics with cells performance and reliability.

The 1T-1R memory cells are constituted by a select nMOS transistor manufactured in BiCMOS technology (width of 1.14 μm and length of 0.24 μm), which also sets the current compliance, whose drain is in series to a MIM stack. The schematic and cross-sectional TEM images of the integrated RRAM cell including the metal lines, the MIM materials and the W-based Via connections are shown in Fig. 3.63. Metal 1 as well as Metal 2 are



(a) (b) (c)

(d) (e)

Figure 3.63: Schematic (a), cross-sectional TEM image of the 1T-1R cell (b), and MIM stack insight (c). Energy-dispersive X-ray spectroscopy (EDX) of the MIM stack, tungsten plug, and Metal 2 (d), and high magnification TEM image of the MIM (e).

metallic layer stacks, consisting of Ti/TiN/Al/TiN/Ti. The MIM integrated on the metal line 2 of the BiCMOS process is composed by 150 nm TiN top and bottom electrode layers deposited by magnetron sputtering, a 7 nm Ti layer, and a 8 nm HfO_2 layer deposited through thermal ALD with the four different processes described in Table 3.12. The Hf/O ratio is analyzed via X-ray photoelectron spectroscopy (XPS) for all processes. The MIM area is equal to 0.4 μm^2. There are two reasons why a MIM area of 0.4 μm^2 is used:

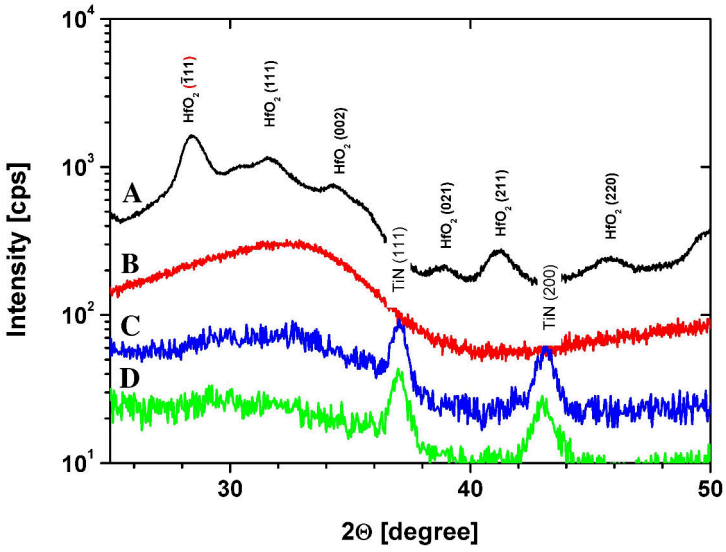
- the data retention is strongly degraded by reducing MIM area due to oxygen out-diffusion thus potentially affecting the measurements precision [97];
- the MIM cell is completely integrated into a standard CMOS process flow. Caused by technological issues of the 250 nm CMOS technology, the area of the tungsten plug, which is required to contact the top electrode of the MIM cell to Metal 3, cannot be reduced further.

Table 3.12: Processes Description

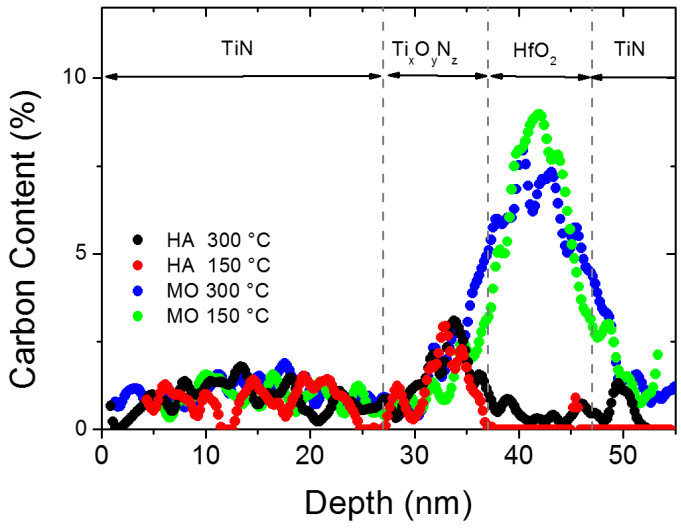
Proc.	Precursor	T_{dep}	Carbon	Oxygen	Phase	Hf/O ratio
A	HA	300°C	0.3%	58%	Poly-crystalline	0.53
B	HA	150°C	0.4%	56%	Amorphous	0.58
C	MO	300°C	7%	49%	Amorphous	0.58
D	MO	150°C	9%	41%	Amorphous	0.51

A halide (HA) Hf precursor ($HfCl_4$) was used for processes A and B, whereas a metal-organic (MO) Hf precursor ($[(CH_3)(C_2H_5)N]_4Hf$) was used for processes C and D. According to X-Ray Diffraction (XRD), the HfO_2 films in processes B, C and D are deposited in the amorphous state, while film A is deposited in the poly-crystalline monoclinic phase as shown in Fig. 3.64(a). XRD measurements on A and B were done after deposition on Si substrates instead of TiN substrates, hence TiN peaks are not present in those samples. The deposition of the HfO_2 films in the 1T-1R cell structures occurs at a temperature within the thermal budget of the BiCMOS process ($T_{dep} < 400°C$), leading to the presence of hydrogen, nitrogen, and carbon based defects caused by the nature of the precursors [98]. Among them, the carbon seems to play a major detrimental role [99]. These defects typically act as trap levels positioned 0.8 eV below the HfO_2 conduction band edge, thus impacting the switching properties of the MIM stack, and consequently the performance and reliability of the 1T-1R cell. The carbon content in the four considered HfO_2 ALD processes is determined through X-Ray Photoelectron

Spectroscopy (XPS), and is reported in Fig. 3.64(b). The carbon percentage refers to the peak concentration measured via XPS in the HfO_2 material. Processes C and D present the highest percentage of carbon contamination due the molecular structure of the Metal-organic precursor [100].



(a)



(b)

Figure 3.64: XRD (a) and XPS (b) analysis on the four considered processes. The XPS reveals the carbon content in the HfO_2.

In Fig. 3.65 it is reported the pristine current measured at 1 Volt after annealing the TiN/Ti/HfO_2/TiN based cells at 400 °C for 30 min in forming gas ambient. During the annealing process, the nitrogen and oxygen gettering activity at the Ti/HfO_2 interface is activated. XPS studies revealed an increased nitrogen and oxygen gettering activity at the Ti top adlayer/HfO_2 interface (see Fig. 3.66). Process C is used as example in the figure, however similar results were obtained on all processes. This reflects the important role of the deposition conditions on the interface chemistry to achieve resistive switching in HfO_2-based MIM cells [101]. The gettering activity is strongly impacted by the morphology of the HfO_2 film. The large fluctuations of the poly-crystalline device A are caused by high affinity of the grain boundaries to charged oxygen vacancies [102]. The leakage current through the dielectric layer can be described by trap-assisted tunneling (TAT) conduction through the oxygen vacancy defects in their positive charge state [81].

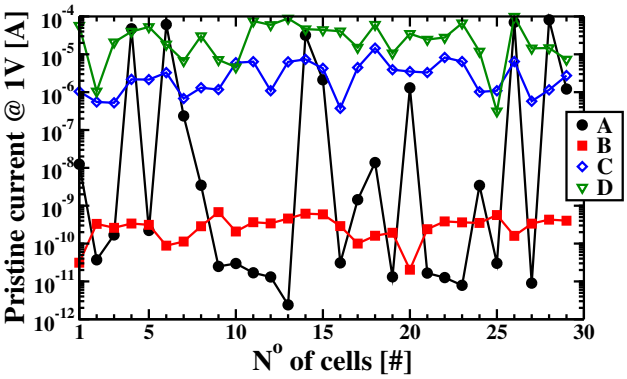


Figure 3.65: Pristine current measured at 1 V for the four considered processes.

The test environment for cells characterization consists in a Keithley 4200-SCS wafer-level tester. The Forming/Set/Reset operations were performed by using an incremental step pulse (V_{STEP}=0.1 V) and verify algorithm as in [94]. A sequence of increasing voltage pulses is applied on the drain of the cell during Forming and Set, with a transistor gate voltage V_G = 1.5V to set the Forming/Set current compliance, whereas the sequence of

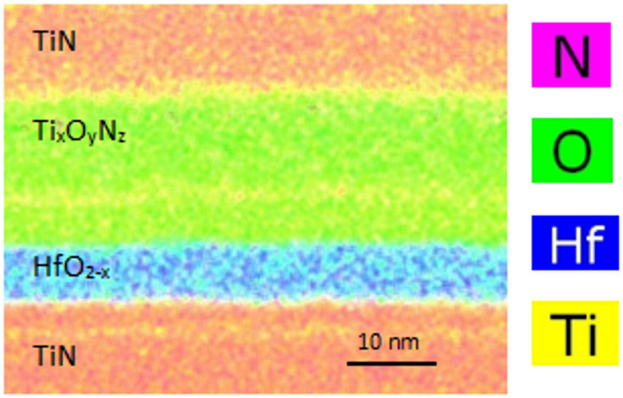


Figure 3.66: Mapping of the elements N, O, Hf and Ti by EDX of a process C MIM cell (similar results were obtained on all processes). The nitrogen and oxygen gettering activity at the Ti/HfO_2 interface after annealing at 400°C in forming gas atmosphere is represented by the thick $Ti_xO_yN_z$ layer.

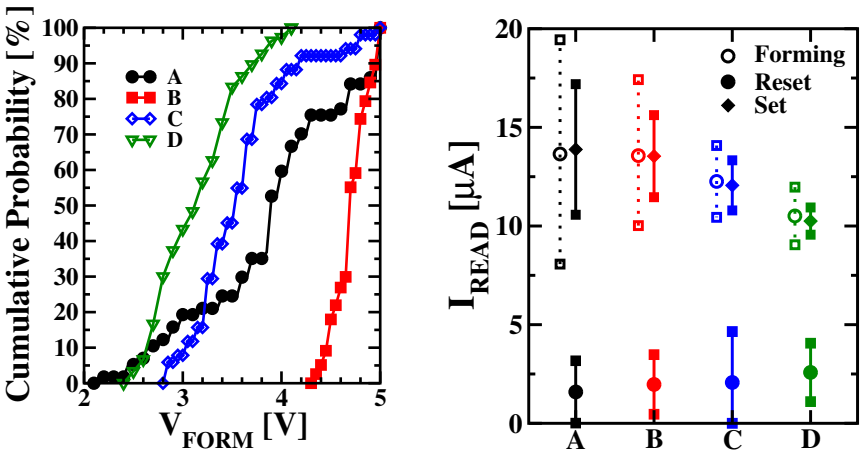
increasing voltage pulses is applied on the source of the cell during Reset, with a transistor gate voltage $V_G = 2.8$ V which leads to a 120 μA compliance current. All pulses feature a duration of 10 μs in order to maximize the switching yield [50]. After every pulse a read-verify operation is performed, where the cell current I_{read} was measured by applying 0.2 V on the drain of the cell with $V_G = 1.5$ V and a read time $T_{read} = 10$ μs. When the read current reaches $I_{target} = 10\mu$A the Forming and Set operations are stopped, whereas during Reset the operation is stopped when the read current reaches $I_{target} = 2\mu$A. V_{FORM}, V_{SET} and V_{RES} denote the voltages at which the targets are reached during Forming, Set and Reset operations, respectively. These parameters reflect the operation of the memory when a Set/Reset algorithm is considered, since they guarantee that a sufficiently high read margin is obtained.

To assess the process quality, we calculated the process yield as the percentage of non-leaky cells showing a read current $I_{read} < 1\mu$A before Forming. To qualify the Forming algorithm, we calculated the Forming yield as the cell percentage showing a read verify current after forming $I_{read} > 10\mu$A, thus ensuring the creation of a conductive filament. Process and Forming yield

obtained are reported in Tab. 3.13 with B and C showing the best results. 100 cells were measured for each yield calculation. Each process had its own wafer resulting in a total testing campaign of 400 cells. Cumulative distribution of the measured Forming voltages are shown in Fig. 3.67 (a). Read current average values and standard deviations calculated on 100 cells for each process after Forming, Set and Reset are reported in Fig. 3.67 (b).

Table 3.13: Process and Forming Yield.

Process	Process Yield	Forming Yield
A	79 %	70 %
B	92 %	88 %
C	100 %	97 %
D	61 %	58 %



(a) (b)

Figure 3.67: Cumulative distribution of the Forming voltages measured on processes A, B, C and D (a). Read current average values and standard deviations (in μA) after Forming, Set and Reset (b).

The process yield is an indicator for the quality of the deposition process in terms of observed current leakage before the forming operation used to evidence potential failing spots in the dielectric films. The forming yield

instead, is an indicator for the quantity of cells in which a conductive path can be efficiently created with the voltage and timings parameters adopted. Both process and forming yield do not necessary translate into good switching properties. Indeed, even if a cell can be successfully formed, its subsequent Set and Reset operations can be largely affected by the physical nature of the cell in terms of dielectric deposition phase and carbon content. Process D is the one where the pristine current (see Fig. 3.65) is the highest due to several factors like the high carbon content in the HfO_2. In this case, it is easier to form a cell within this process since the high number of carbon defects helps the creation of a conductive path. On the other hand, the subsequent Set/Reset operations will be difficult (i.e., the re-creation/rupture kinetics of the filament is threatened by the large carbon defects concentration, thus being less controllable) with the result of increasing the required voltages for switching and their dispersion as well. Process D has a higher carbon concentration than process C since it is deposited with a low process temperature, hence its process yield is impacted by a high number of failing spots in the cells dielectric. Process B features a lower carbon defects concentration in the dielectric that translates in a slightly difficult forming operation (i.e., larger voltage requested as shown in Fig.3.67 (a)), but with a higher degree of uniformity as shown by its pristine current. However, when the conductive filament is correctly created, for the cells within this process, the subsequent Set and Reset operations will not be impacted by carbon defects as for process C, resulting in lower average and dispersion of the switching voltages. To sum up, the precursor choice strongly influence performance and reliability of the devices and the ALD temperature plays an important role, with effects which are different depending on the precursor. The use of halide precursor instead of metal-organic allows reducing roughness, contaminants and leakage current independently from temperature [103]. When Halide precursor is considered, a temperature increase change the HfO_2 phase from amorphous to poly-crystalline. This means that in this case too high temperatures should be avoided otherwise the conduction will occur through

grain boundaries and the leakage current will increase, reducing the controllability of the Reset operation. When metal-organic precursor is considered, the temperature increase is beneficial since allows reducing roughness, carbon content and leakage current without causing the HfO_2 crystallization.

The Resistance Ratio (i.e., the ratio between the RRAM cells' resistances measured after Set and Reset operations), V_{SET}, and V_{RES} average values and normalized variances (i.e., the ratio between variance and average value) evolution during 100 Set/Reset cycles with Incremental Step Pulse and verify algorithm, calculated on 20 cells for each process, are reported in Fig. 3.68.

Set and Reset targets are reached at lower voltages in halide precursor processes (A, B) with respect than metal organic processes (C, D). This is due to the carbon content which is the major contributor in controlling the filament creation/rupture speed. Indeed, a higher carbon content in the HfO_2 corresponds to a higher number of trap sites in the dielectric resulting in a less controllable filament creation/rupture kinetics [98, 99]. As shown in Fig. 3.68 (a), process A shows the largest Resistance Ratio after forming but also a fast reduction during the endurance test, whereas a higher stability of the Resistance Ratio during cycling is observed for the amorphous films. Moreover, for the latter processes the performance seems to be related to the carbon content: film B, corresponding to the amorphous process with the lowest carbon content, shows the highest Resistance Ratio with the lowest normalized variance (see in Fig. 3.68 (b)) after 100 cycles and the lowest V_{SET} and V_{RES} with the highest stability in cycling. As shown in Fig. 3.68 (c,e), V_{SET} and V_{RES} increase as a function of the carbon content in amorphous films. If normalized variances reported in Fig. 3.68 (d,f) are considered, process B still shows the lowest values and the highest stability during cycling, confirming that the carbon content plays a fundamental role on cells' performance and reliability [99]. Since process B gave the best performance after 100 cycles, the endurance test has been extended to 1000 cycles: no relevant variation of the parameters has been measured (see Fig. 3.68). In order to identify the causes of the Resistance Ratio reduction,

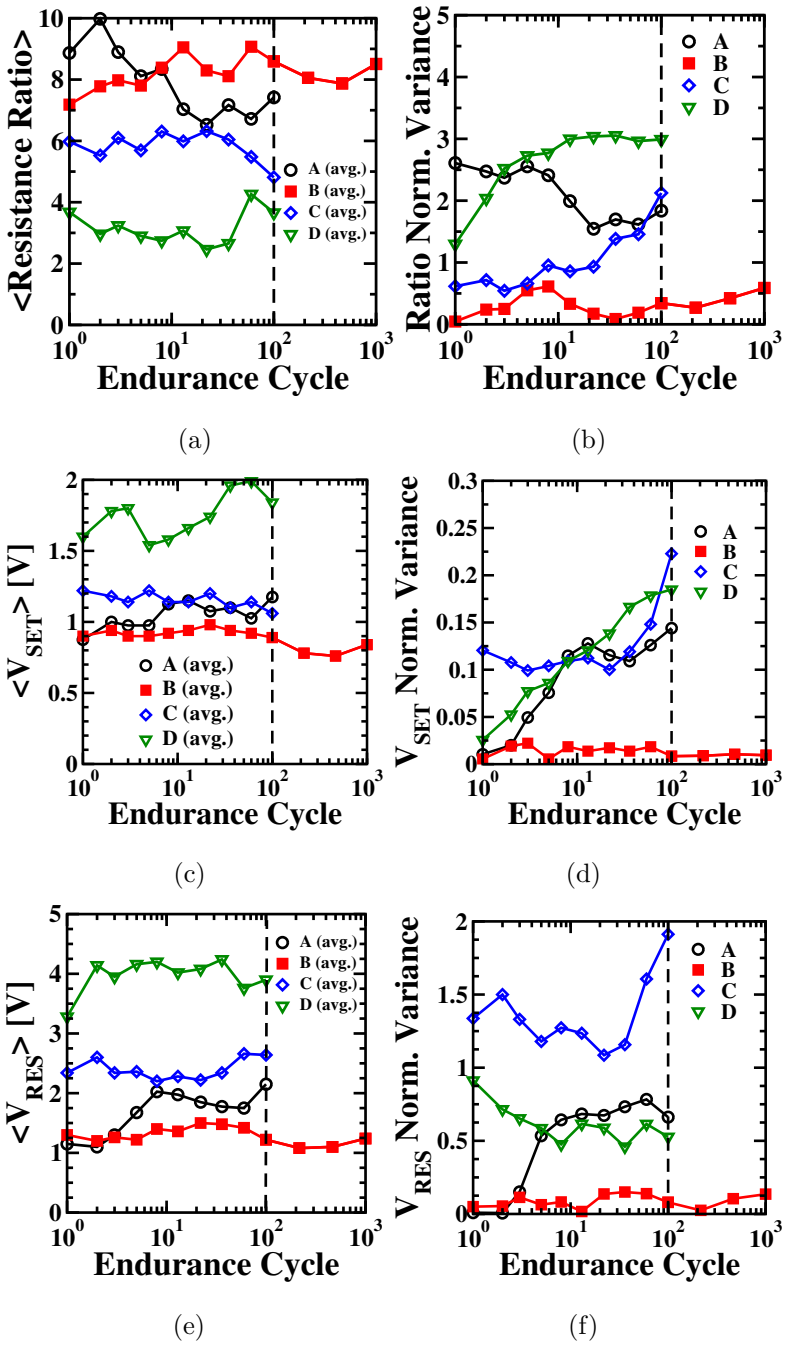


Figure 3.68: Resistance Ratio, V_{SET} and V_{RES} average values and normalized variances calculated on 20 cells for each process during cycling.

HRS and LRS average resistances and their standard deviation calculated during cycling are reported in Fig. 3.69. The ratio reduction observed in processes A and C is attributed to the HRS state degradation, while LRS shows good stability in all processes. All these electrical characterization results are resumed in Tab. 3.14.

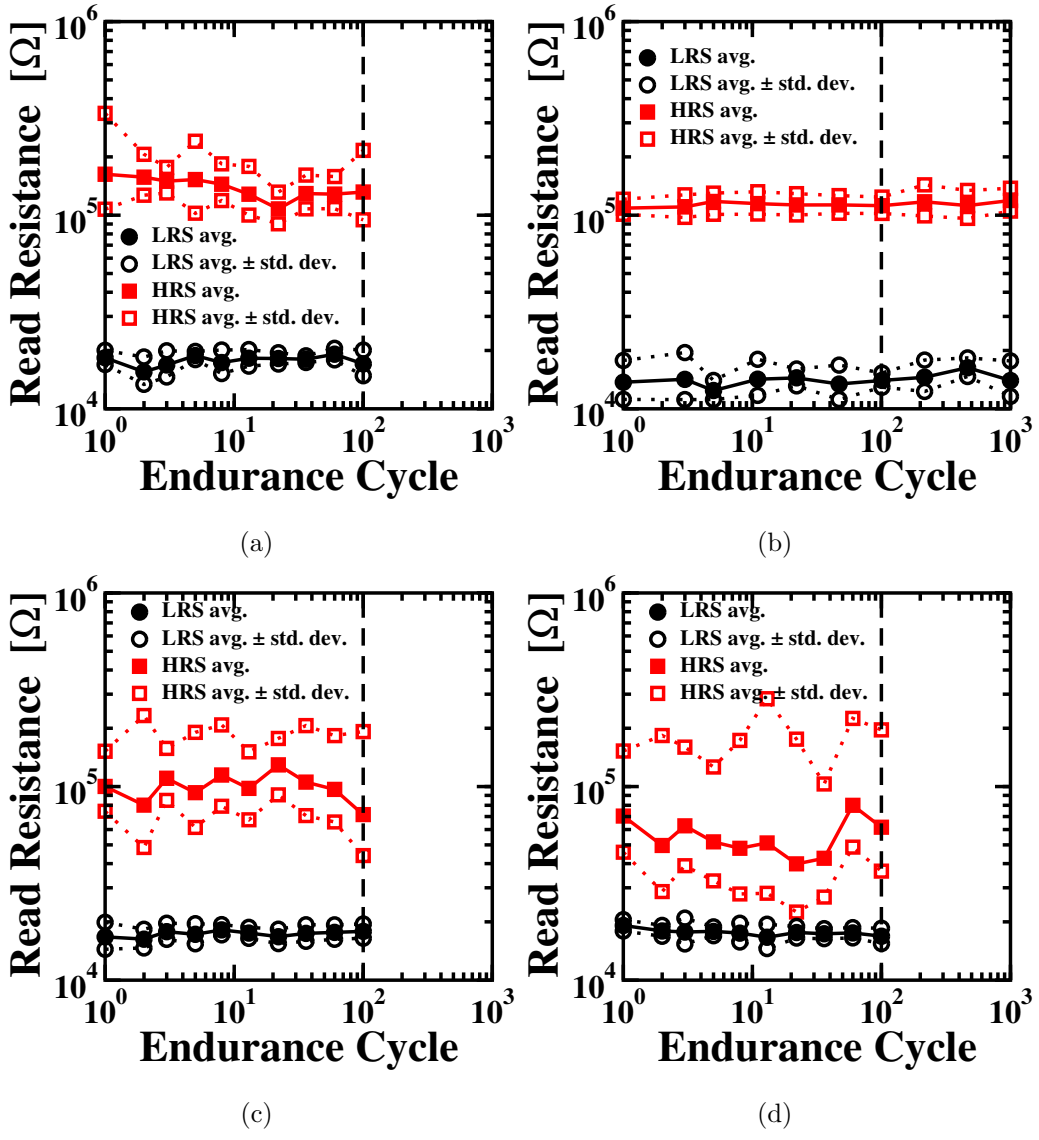


Figure 3.69: HRS and LRS average resistances and their standard deviation calculated on 20 cells for each process during cycling.

Table 3.14: RRAM switching parameters and variability resume extracted from electrical characterization.

Process	A	B	C	D
Initial Ratio	9	7	6	4
Final Ratio (100 cycles)	7	8	5	4
V_{FORM}	Medium	High	Low	Low
V_{FORM} variability	High	Low	Medium	Medium
$V_{SET,RESET}$	Medium	Low	Medium	High
$V_{SET,RESET}$ variability	High	Low	High	High
HRS variability	High	Low	High	High

The use of a physical model taking into account the localized nature of conduction is required to understand the impact of the carbon within the HfO_2 stack on the cells' conduction properties after Reset. In this regard, the Quantum Point Contact (QPC) model has been used [52] since it allows to correctly represent the I-V characteristics measured after Reset independently of the atomic species involved in the conduction mechanism (the charge transient occurs through oxygen vacancies in the amorphous HfO_2, whereas it is dominated by grain boundaries conduction in the poly-crystalline structure) according to equation 3.2, whose parameters were already described in the previous subsection.

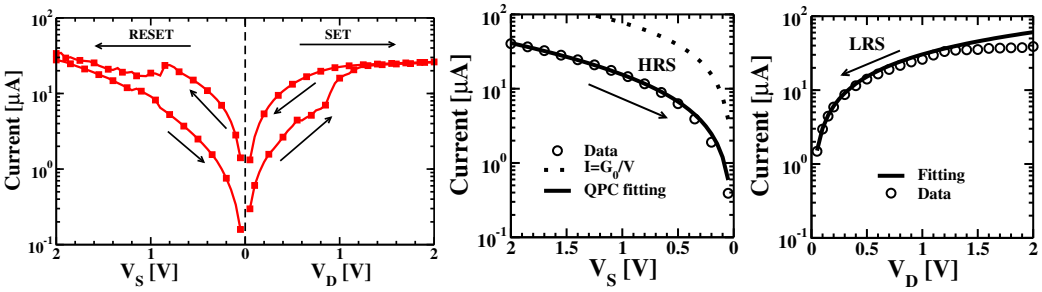
The series transistor integrated in the cells considered in this work has a significantly large area with a good current driving capability (i.e., W=1.14 μm, L=0.24 μm, μ_0=1000 cm^2/Vs, t_{ox} = 5nm). During the read operation the transistor constantly works in the linear region with a fixed resistance that is negligible compared to that of the 1R element, therefore we could rule it out in the application of the QPC model for the HRS state on the 1T-1R cells. Estimated r_{ds} values during read are in the range of 16 mΩ. However, the effect of the series resistance of the transistor in the LRS is considered by changing the QPC fitting equation only for this state.

Within this framework, if the series resistance R external to the constric-

tion is negligible [52], $I = G_0V$ sets a limit: in case of $I > G_0V$ the presence of more than a single conductive filament or multimode conduction should be taken into account. However, both the presence of interfacial layers in the MIM stack and the series resistance of the select transistor lower this limit further and yield to the calculation of the LRS current as:

$$I = G_0 * (V - RI) \tag{3.5}$$

where RI is the cumulative voltage drop attributed to the two former effects. To correlate the results obtained so far with the physical nature of the conduction in the four different processes, we performed a DC-sweep characterization after Forming by increasing V_D from 0 to 2V during Set with V_G = 1.5V, and by increasing V_S from 0 to 2V during Reset with V_G = 2.8V, respectively. In both Set and Reset the applied voltage has been increased with steps of 0.05 V and a sweep ramp of 1 V/s. Since the same degree of fit was obtained on all processes, as example Fig. 3.70 shows the perfect compatibility of the model with respect to the measured Reset (b) and Set (c) I-V curves on a random sampled cell in process B.



(a) (b) (c)

Figure 3.70: Typical I-V characteristic measured on a process B random sampled cell (a), HRS fitting (b) and LRS fitting (c) through QPC.

Since the current measured after Reset is below the G_0 limit we can assume that the conductive filament is interrupted or extremely narrow and in this latter case we can calculate the barrier length and radius of the constriction. If the filament is completely interrupted we have a material barrier

rather than a confinement barrier. The model is valid as long as we have a filament: what we consider an interruption is an extremely narrow constriction thus a high barrier, but lower than the material barrier. Set I-V curve fitting was obtained by using Eq.3.5 with R=20kΩ, which is consistent with literature [52]. A possible explanation of R high value is the presence of a residual potential barrier after Set [52, 104]. Fig. 3.71 shows average values and standard deviations of the fitting parameters for the correctly formed cells considered in the present work. The range of α parameter in Fig. 3.71 (a) is similar for the amorphous HfO_2 processes B, C and D, while the width of the parabolic barrier is larger for poly-crystalline film A. This could be ascribed to the shape of the constriction, which may be changed due to the HfO_2 crystallization. Process B shows the lowest barrier height ϕ variability in Fig. 3.71 (b), confirming the highest cell-to-cell uniformity. Moreover, barrier height ϕ average value is shown to decrease as a function of the carbon content, while its standard deviation is shown to increase. The highest variability of the barrier height is calculated for process A, which is ascribed to the poly-crystalline structure of the film. Process B shows the highest average β values in Fig. 3.71 (c), which means the most asymmetric constrictions, and the lowest fitting parameters variability, resulting in the highest cell-to-cell uniformity.

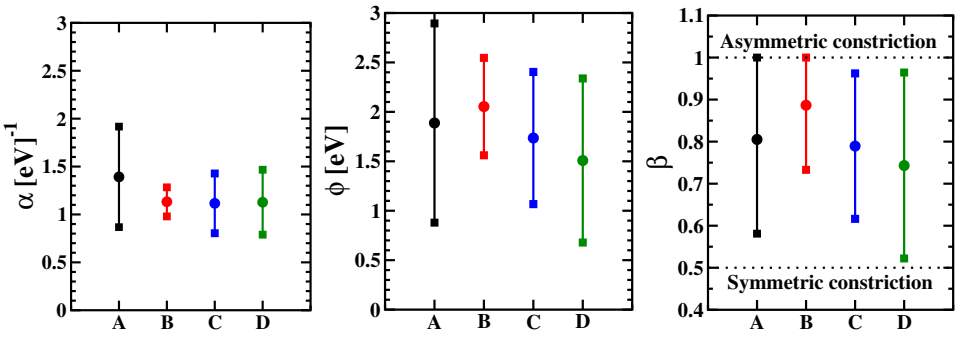


Figure 3.71: Distributions of the QPC model fitting parameters α (a), ϕ (b) and β (c).

In our case β is typically close to 1, since our MIM structure tends to create an asymmetric constriction due to the presence of the Ti layer [94]. However, the presence of carbon defects and the Hf phase have an impact on the filament shape, causing a reduction of the asymmetry. Since the spatial gap d in the conductive filament after Reset acts as a tunneling potential barrier of about his thickness [82], considering the relationship between α and the potential barrier thickness reported in [52] d can be calculated as:

$$d = \frac{h\alpha\sqrt{\phi}}{\pi^2\sqrt{2m^*}} \tag{3.6}$$

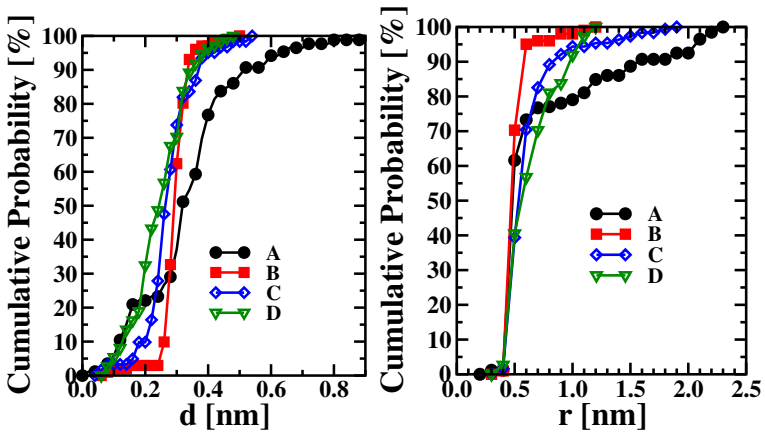
where m^* is the electron effective mass in the constriction. The radius r of the constriction was calculated as:

$$r = hz_0/2\pi\sqrt{2m^*\phi} \tag{3.7}$$

where $z_0 = 2.404$ is the first zero of the Bessel function J_0 [52]. The cumulative distributions of calculated barrier width d and radius r of the constriction are reported in Fig. 3.72 (a) and (b): process B shows the lowest filament radius variability. The conductive filament obtained after Reset is depicted in Fig. 3.73.

The average values of d and r as a function of the carbon content are reported in Fig. 3.74 (a) and (b). The average barrier width is shown to decrease as the carbon content increase, whereas the average radius is shown to increase. These two observations mean that the narrowmost point along the filament tends to become more transparent for the electron flow.

The effect of carbon on barrier width d and conductive filament constriction r is depicted in Fig. 3.75: the carbon atoms form trap levels inside the HfO_2 band gap [99], generating permanently conductive paths that can be modeled as a radius increase when atoms places next to the filament constriction, or as a barrier length reduction. These permanently conductive paths cause large leakage currents as observed in process D, reducing the process yield and the cells' performance since it is more difficult to obtain



(a) (b)

Figure 3.72: Cumulative distribution of calculated barrier length d (a) and radius of the filament constriction r (b).

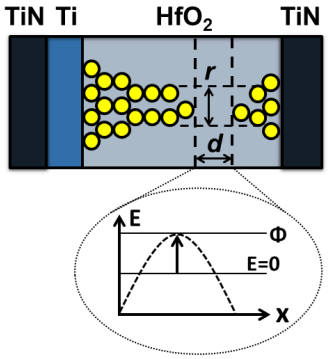
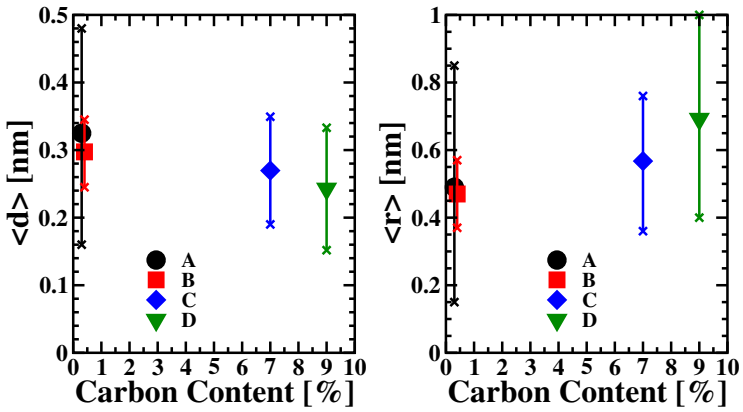


Figure 3.73: Schematic showing the conductive filament shape after Reset.

low resistances during Reset and to control the cell-to-cell variability with Set and Reset operations. To sum up, the cell-to-cell variability sources that can be enlightened through the QPC model applied for HRS fitting of the different processes are two: the phase of the deposited dielectric, and its carbon content. Process A has a low carbon content, but it is deposited as poly-crystalline, whereas processes C and D are deposited as amorphous but with a higher carbon content. Since process B features both an amorphous phase after deposition and a low carbon content, it is the one where the variability sources play a minor role.



(a) (b)

Figure 3.74: Average values of calculated barrier length d (a) and radius of the filament constriction r (b). Error bars show the standard deviation.

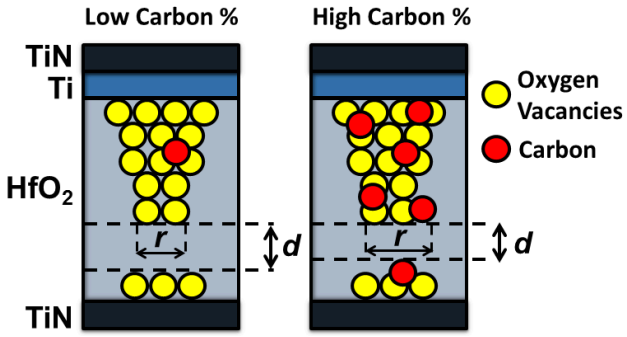


Figure 3.75: Schematic showing the effect of carbon on barrier length d and conductive filament constriction r.

The different performance of several HfO_2 deposition processes for RRAM applications involving halide and metal-organic precursors have been investigated. QPC modeling allowed understanding the physical properties of each process by analyzing the conductive filament properties. The grain boundaries conduction in the poly-crystalline HfO_2 structures could be the reason of the high cell-to-cell variability. The use of tuned deposition parameters allowed obtaining amorphous HfO_2 instead of poly-crystalline with halide precursor, resulting in the highest inter-cell and intra-cell uniformity, as evidenced by electrical characterization and model fitting parameters. Metal-

organic precursors-based processes result in amorphous HfO_2 films as well, although featuring a higher carbon content than other processes. The inter-cell uniformity seems to be affected by carbon: processes with high carbon content show reduced Resistance Ratio and increased variability of the Set and Reset parameters. In conclusion RRAM stacks manufactured with halide precursor and low temperature HfO_2 deposition show the most promising results in terms of cell-to-cell variability, reliability and yield.

3.7 Fundamental variability limits in RRAM

Even if RRAM is seen as a possible replacement for Flash memories due to manufacturing process simplicity, easy integration with logic, lower voltage operation, and good cycling with sufficient retention capability [2], one of the last technical roadblock for adoption is the variability understanding. In the previous sections several approaches capable of reducing the variability were presented, such as employment of enhanced algorithms and process optimization. In this section, the fundamental variability limits of RRAM are investigated through a full characterization of Forming, Set and Reset, where the variability of each operation is extracted from a 4kbits OxRAM array measurement. After introducing the technology used, variability after Forming is presented before finding its relationship with subsequent Set and Reset operations. Cell-to-cell and cycle-to-cycle variabilities are then analyzed and compared to previous literature works confirming their intrinsic origin. Finally, the variability evolution as function of cycling is explained giving clear predictions and guidelines for variability aware operation of RRAM arrays [105].

3.7.1 Experimental Setup

OxRAM technology has been integrated on 130nm CMOS logic. On top of Cu Metal 4, a TiN bottom electrode is defined. Then a CMP touch is done and an HfO_2 10nm/Ti 10nm/TiN stack is deposited. Main integration

steps and cross section of a 300nm diameter integrated device are described in Fig. 3.76. The integration platform is shared across multiple resistive-based technologies, potentially hosting: OxRAM, CBRAM, and MRAM as well. The current pristine state of the cell has been measured around 1pA from 1R structure. 4 kbits 1T-1R array measurement shows that the high resistive state (HRS) and low resistive state (LRS) can be separated down to 3σ with no extrinsic bits, thus showing good process maturity (Fig. 3.77).

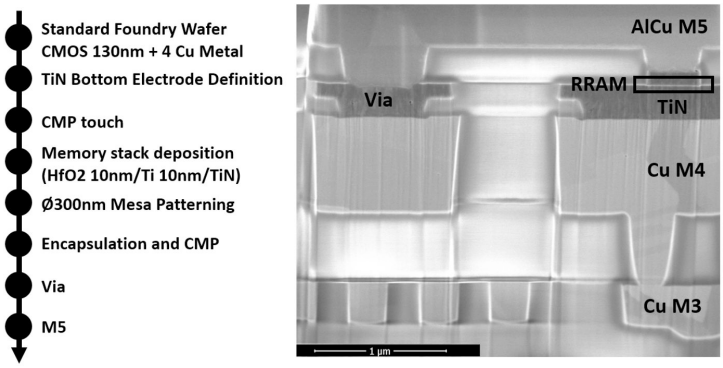
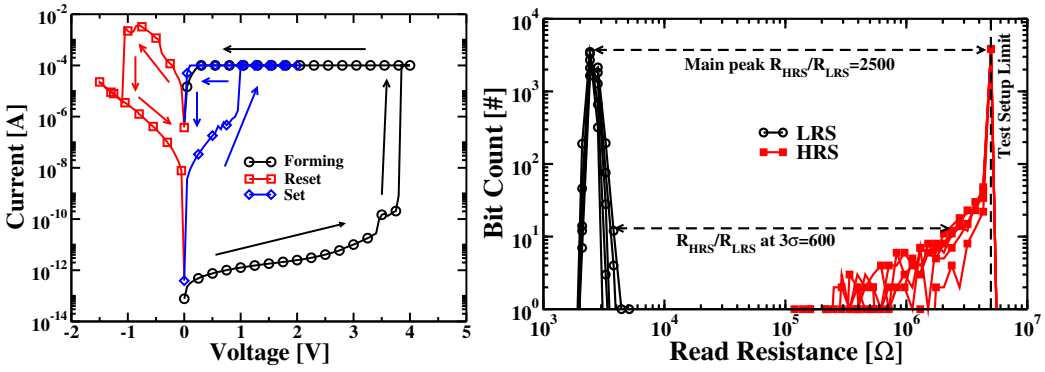


Figure 3.76: Description of the integration flow and TEM cross section of the integrated TiN/HfO_2/Ti/TiN OxRAM.



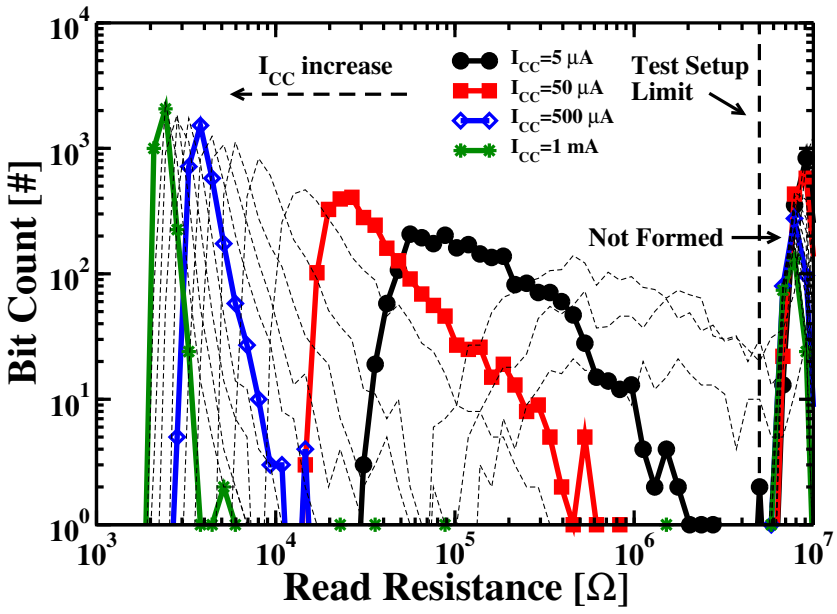
(a) (b)

Figure 3.77: 1R Forming, Set and Reset I-V curves (a). LRS and HRS distributions measured on 4 kbits array with Set and Reset conditions allowing to obtain large resistance window (b).

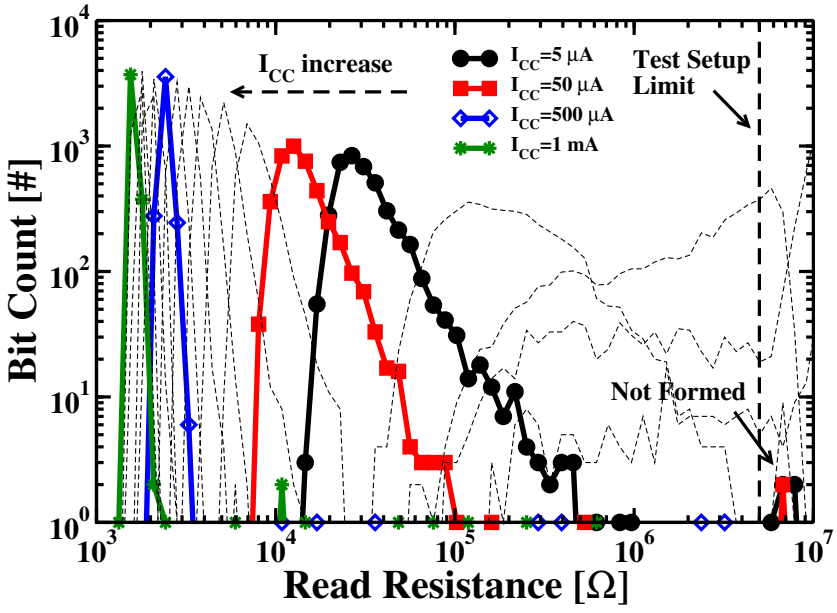
The measurements in this work have been performed on 4kbits 1T-1R array developed within this platform. The wordline (WL) is connected to the NMOS gate, setting the current compliance I_{CC}. Forming and Set operations are performed by applying a positive voltage pulse on the bit line (BL) that is also the OxRAM top electrode, whereas Reset is performed by applying a voltage pulse on the source line (SL) whose contact is on the transistor side (bottom electrode). The select transistor features large width in order to assess cell operation on a large current range (1μA - 5mA) and to ensure minimal dispersion impact (σ_T at read = 4 Ω). As a result, all measured resistance dispersion on the array can be always considered as intrinsic to the memory element. The standard deviation σ used for this analysis is calculated on the array after excluding the extreme outliers point due to extrinsic causes. The 2σ bounds are plotted for clarity and consistency.

3.7.2 Forming for low variability

Incremental current compliance pulse operations were performed to investigate the impact of I_{CC}, V_{BL} and pulse length (T_{PULSE}) on Forming. The use of longer pulses allows obtaining more compact distributions at same I_{CC} (Fig. 3.78). The filament creation process is found to be purely field driven and happens around V_{BL}=3.8V using 100ns pulses for this stack. Further, we confirmed that the median resistance final value is directly related to I_{CC} (Fig. 3.79). Fig. 3.80 shows that time impact on the formed cells creation follows a Weibull law, hence a single pulse has the same impact of a train pulse with equivalent duration. While 50% of the cells are formed within 200 ns, it required 40μs to form about a 3σ range of the cells population. This very large time dispersion between fast and slow formed cells can be detrimental if a too large I_{CC} is used. In this latter case, the fast formed cells are stuck to low resistance state (LRS) before the slow formed cells are formed. Tab. 3.15 shows that a low compliance current of 5μA allows forming a cells population within 3σ range without degrading any cell. This remove the need of time/energy consuming verification algorithms during RRAM operations.

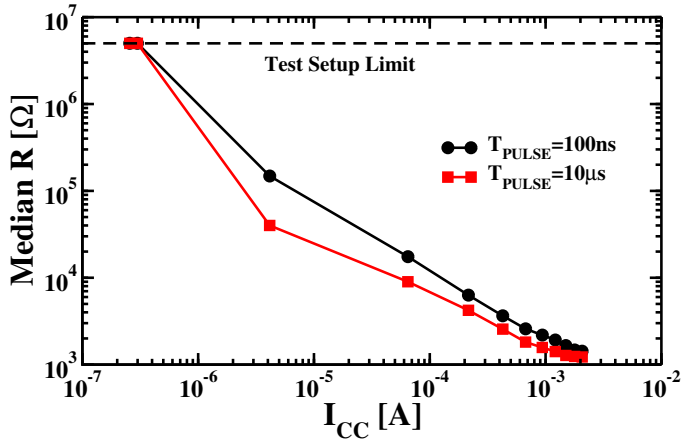


(a)

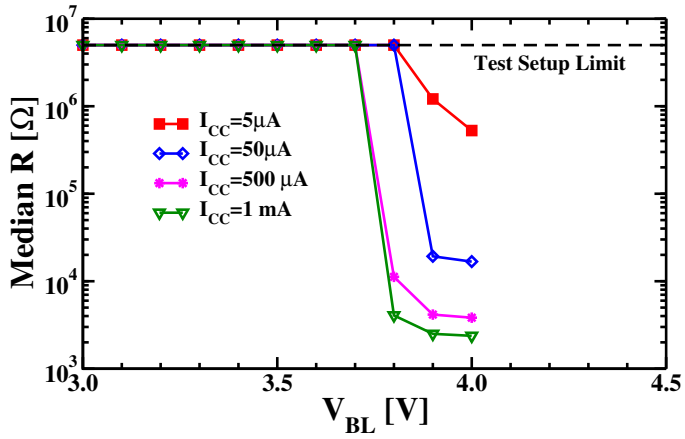


(b)

Figure 3.78: Forming with increasing I_{CC} and V_{BL}=4V: read resistance distributions evolution with T_{PULSE}=100ns (a) and 10μs (b). Dotted lines show the distribution evolution.



(a)

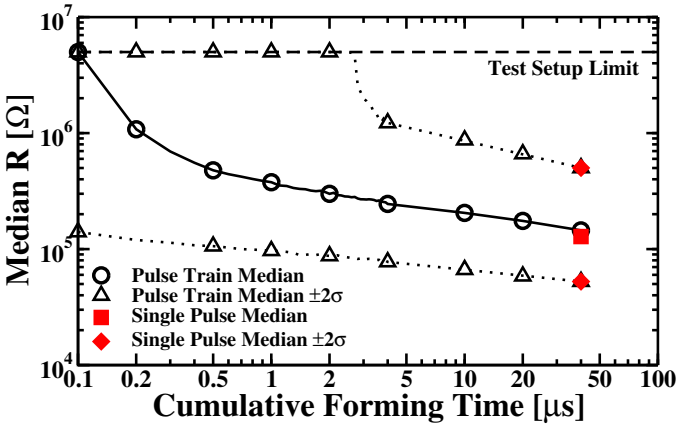


(b)

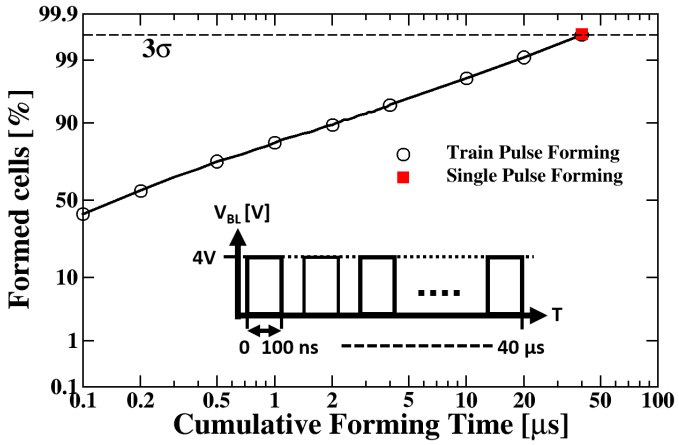
Figure 3.79: Median resistance as a function of increasing I_{CC} with V_{BL}=4V, T_{PULSE}=100ns and 10μs (a). Forming with V_{BL} increasing, T_{PULSE}=100ns and different I_{CC} (b).

Table 3.15: RRAM switching parameters and variability resume extracted from electrical characterization.

Forming I_{CC}	Switching Cells	3σ Forming Time	3σ Forming Energy
5 μA	99.6 %	40	0.8 nJ
50 μA	96.5 %	20	4 nJ
500 μA	86.5 %	10	20 nJ



(a)



(b)

Figure 3.80: Median resistance $\pm 2\sigma$ during pulse train Forming procedure with I_{CC}=5μA, V_{BL}=4V, T_{PULSE}=100ns and after Forming with a single pulse of the same cumulative time T_{PULSE}= 40μs (a). Formed cells percentage during pulse train Forming and after single pulse Forming (b).

We observe that the cell-to-cell variability after Forming is directly related to the median resistance (Fig. 3.81). Note that in this case no Reset has been applied yet. Fig. 3.82 shows that for median resistances higher than $1/G_0$ the dispersion follows a lognormal distribution with median parameter T_{50}=172kΩ and form factor λ=0.78. G_0=(12.9k Ω)$^{-1}$ is the quantum conductance unit corresponding to the creation of a single conductive nanowire [52].

In this case, the current flowing through the cell can be calculated as:

$$I = G_0 e^{-\alpha\phi} V \tag{3.8}$$

where ϕ is the barrier height and α is related to the width of the barrier, assuming a parabolic longitudinal potential as sketched in Fig. 3.83. Those parameters refer to the QPC model [52,67]. Since the conduction mechanism in OxRAM cells with resistance higher than $1/G_0$ is due to a number of vacancies following a Poisson distribution [106], it is possible to extract the standard deviation from the cells population as a function of the median resistance:

$$\sigma = \sqrt{G_0 e^{-\alpha\phi}} R^{1.5} \tag{3.9}$$

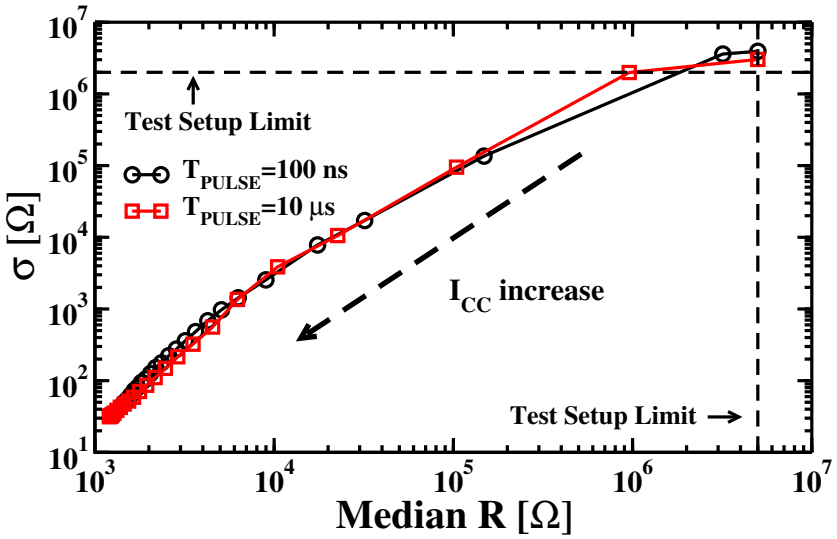
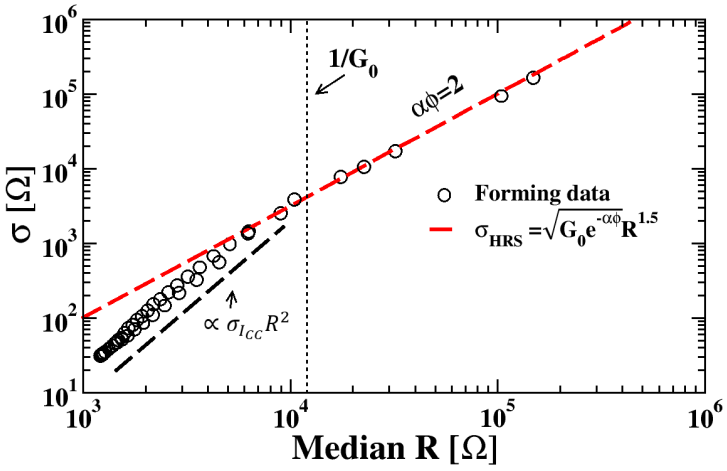
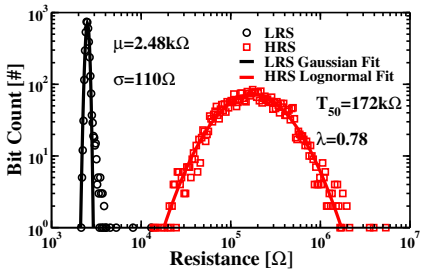


Figure 3.81: Median resistance and σ evolution during Forming with increasing I_{CC}, V_{BL}= 4 V, T_{PULSE}= 100 ns and 10 μs.

For median resistances lower than $1/G_0$, the dispersion is found to be gaussian with a dispersion lower than Eq.3.9 and proportional to $\sigma_{I_{CC}} R^2$, with $\sigma_{I_{CC}}$ being the dispersion of I_{CC} at Set condition. The LRS and HRS



(a)



(b)

Figure 3.82: σ vs. median R fit above $1/G_0$ (a). LRS and HRS distributions fitting after Forming with I_{CC}= 5 μA, V_{BL}= 4 V and T_{PULSE}= 40 μs (b).

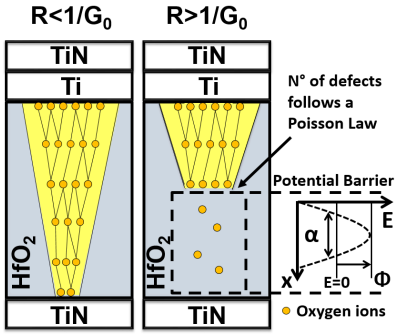
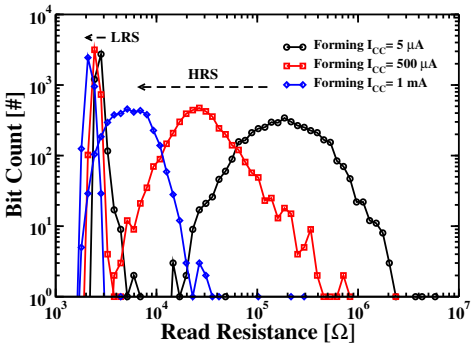
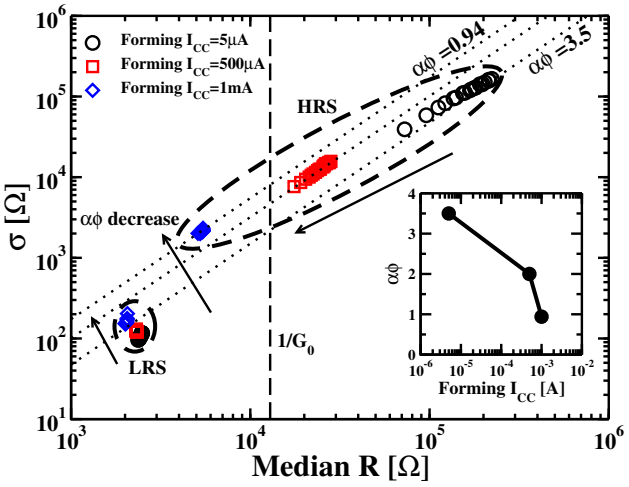


Figure 3.83: CF schematic representation in case of R below and above $1/G_0$. QPC model parameters α and ϕ are depicted.

after Forming distribution dependence on I_{CC} with identical Set and Reset pulse is shown in Fig. 3.84. In this case, Set is performed with $I_{CC} = 0.4$mA, $V_{BL} = 2$V, $T_{PULSE} = 100$ ns whereas Reset is performed with $I_{CC} = 2.2$mA, $V_{SL} = 2.5$V, $T_{PULSE} = 100$ ns. We see that the σ/R relationship follows Eq. 3.9 for I_{CC} values of 5μA with $\alpha\phi$=3.2 while $\alpha\phi$ decreases for higher I_{CC}. These data confirm that large compliance currents induce higher variability and lower $\alpha\phi$ even at the first Set/Reset operation.



(a)



(b)

Figure 3.84: LRS and HRS distributions after Forming with V_{BL}=4V, T_{PULSE} allowing to form up to 3σ cells and different I_{CC} (a). LRS and HRS σ vs. median R after Forming with different I_{CC} (b). Inset show $\alpha\phi$ decrease when Forming I_{CC} increase.

3.7.3 Variability in Set/Reset

The maximum R_{HRS}/R_{LRS} ratio reachable is higher than 10^3 (limited by our array measurement setup). Fig. 3.85 shows the impact of I_{CC}, V_{BL} and T_{PULSE} on the median and dispersion. After Forming, V_{BL} for Set has a threshold-like behavior requiring 1V to set at least 50% of the cells and 1.6V for 2σ. For Reset, as stated earlier, the I_{CC} is the most important, and V_{SL} has no impact beyond 2.5V. The time impact is small from 100ns to 10μs demonstrating a fully operational OxRAM array using 100ns pulses. All these data are plotted on Fig. 3.86 in order to show the σ/R relationship in different operating condition. Independently of the operation, all the experimental points follows Eq. 3.9 for median resistances higher than $1/G_0$ with $\alpha\phi$=2 whereas in case of extreme condition such as high I_{CC} or long pulses the points deviate from the line due to an increase of the standard deviation, with $\alpha\phi$ decreasing to 0.32.

3.7.4 Cell to cell VS. cycle to cycle distribution

For the first time a comparison between cell-to cell and cycle-to-cycle variability in different operating condition is reported (Fig. 3.87). For resistances below $1/G_0$, cell-to-cell variability is higher than cycle-to-cycle variability since the dispersion is limited by the cell selector variability and not by the memory element. When the resistance is above $1/G_0$, the opposite situation is observed: cycle-to-cycle dispersion is higher as the median value typically evolves with cycles. When literature data of other filament-based RRAM technologies are compared (Fig. 3.88), we can see that all the points are in a limited range of $\alpha\phi$ confirming that the minimal dispersion achievable for median resistance above $1/G_0$ follows Eq.3.9. This fundamental limit is linked to the discreteness of the number of elements (vacancies/atoms) participating to the conduction. The points may deviate below $1/G_0$ as the main variation source is not related to the memory element, whereas all points are aligned above $1/G_0$ when considering Eq. 3.9 with different $\alpha\phi$ values that

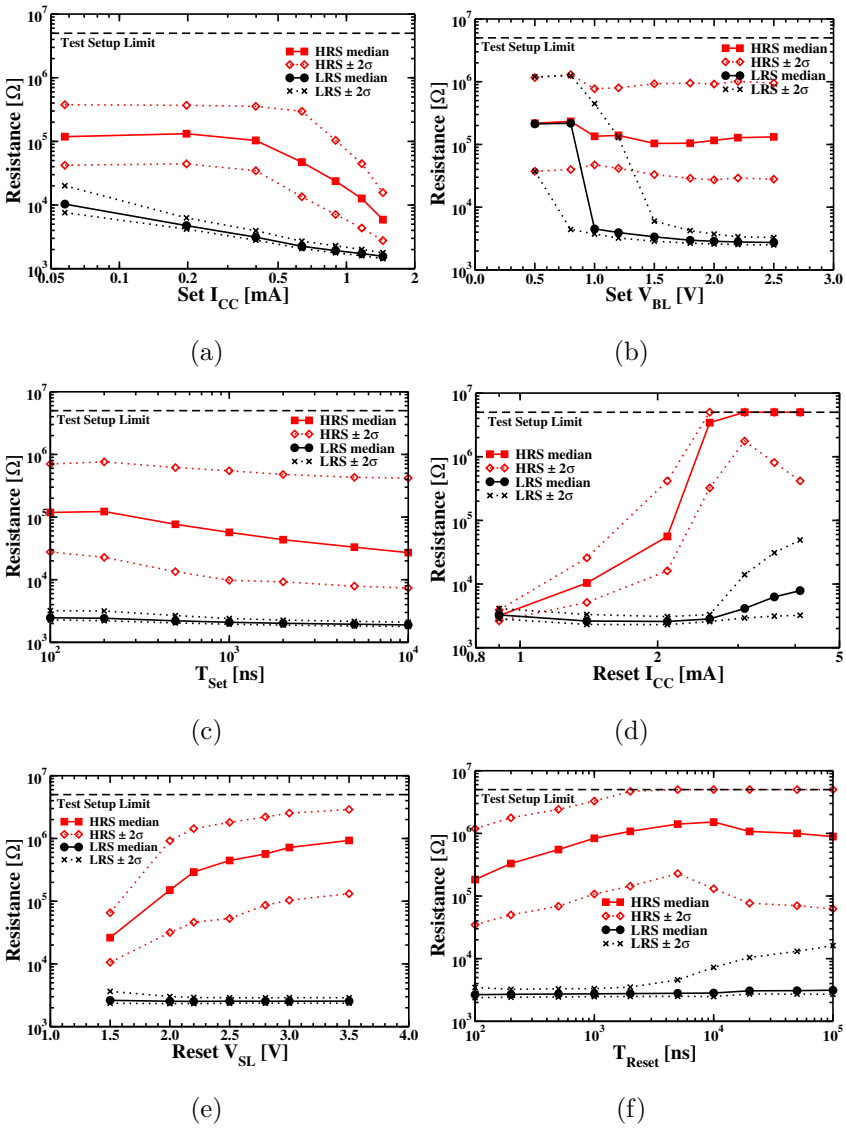
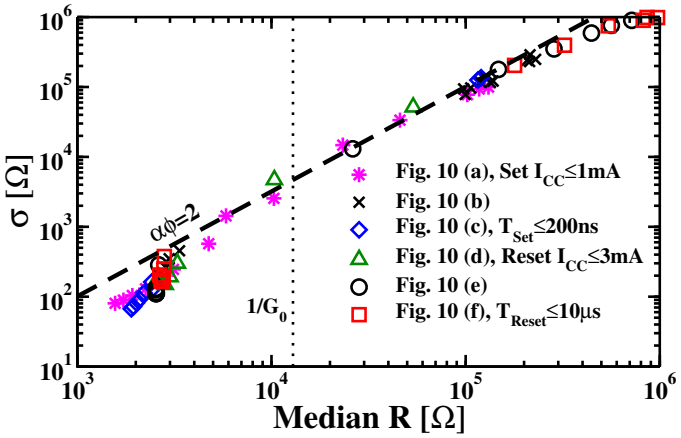
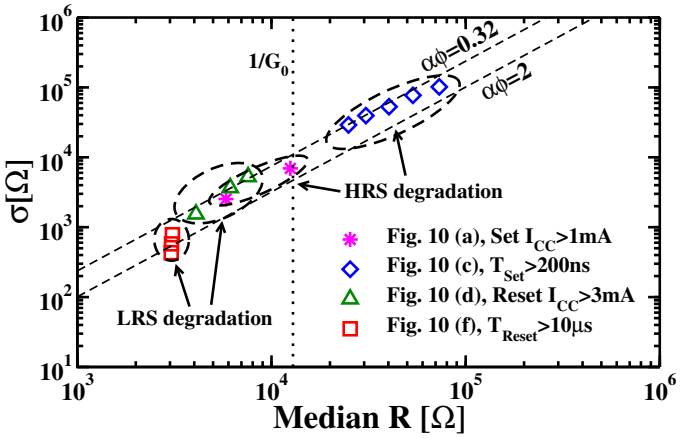


Figure 3.85: Impact of Set I_{CC} (a), V_{BL} (b) and T_{PULSE} (c) on Set and Reset (median $\pm 2\sigma$), with $V_{SL,Reset}$=2.5V, $I_{CC,Reset}$=2.2mA and $T_{PULSE,Reset}$=100 ns. Set I_{CC}=0.4 mA, V_{BL}=2V and T_{PULSE} = 100ns were used when varying other parameters. Impact of Reset I_{CC} (d), V_{SL} (e) and T_{PULSE} (f) on Set and Reset (median $\pm 2\sigma$), with $V_{BL,Set}$=2V, $I_{CC,Set}$=0.4mA and $T_{PULSE,Set}$=100 ns. Reset I_{CC}=2.2 mA, V_{SL}=2.5V and T_{PULSE} = 100ns were used when varying other parameters.



(a)



(b)

Figure 3.86: σ vs. median R in not degrading (a) and degrading (b) conditions extracted from measurements in Fig. 3.85. Degradation causes $\alpha\phi$ decrease.

depends on materials, thicknesses of the stacks or degradation effects. For median resistances below $1/G_0$ the distribution is depending on the previous operations (Fig. 3.89). After a Reset pulse, the distribution will follow the Poisson limited distribution whereas after a Set pulse the distribution will be lower than the Poisson limited distribution, as the number of element composing the filament is no more random but dictated by I_{CC}. Fig. 3.89 shows that this transition is happening precisely at $1/G_0$. As a result the

single-pulse resistance distribution can be minimized during operation only after a Set with a resistance below $1/G_0$. Hence, distribution after a Set pulse will be smaller than after a Reset pulse at same median resistance value.

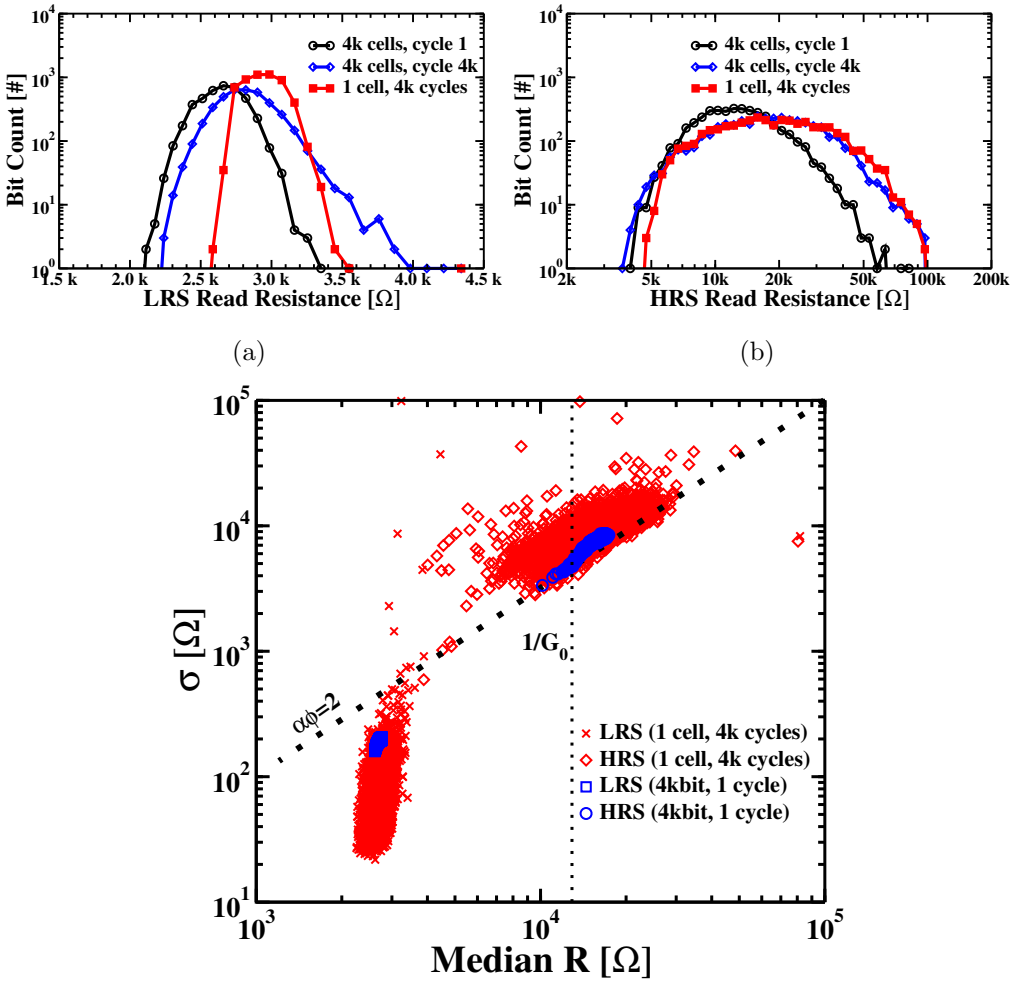


Figure 3.87: Comparison of LRS (a) and HRS (b) cell-to-cell and cycle-to-cycle distributions during 4k cycles with Set I_{CC}=0.4 mA, V_{BL}=2V, T_{Set} = 100ns and Reset I_{CC}=2 mA, V_{SL}=2.5V, T_{Reset} = 100ns. σ vs. median resistance comparison of cell-to-cell and cycle-to-cycle distributions (c).

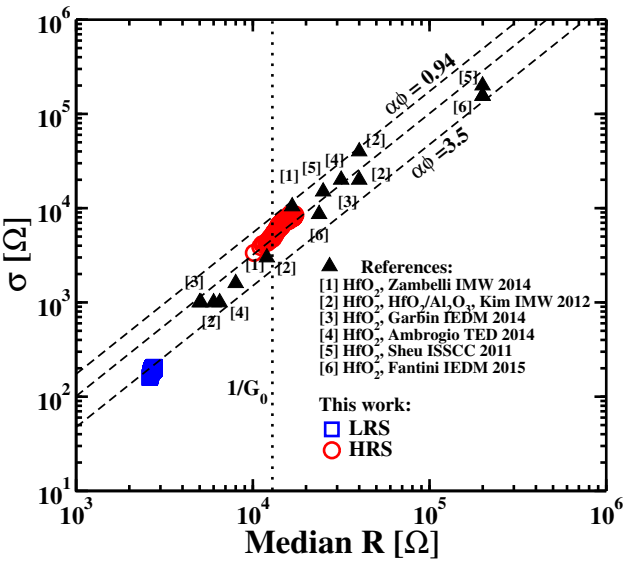


Figure 3.88: σ vs. median resistance: comparison of cycle-to-cycle distributions results with data from literature.

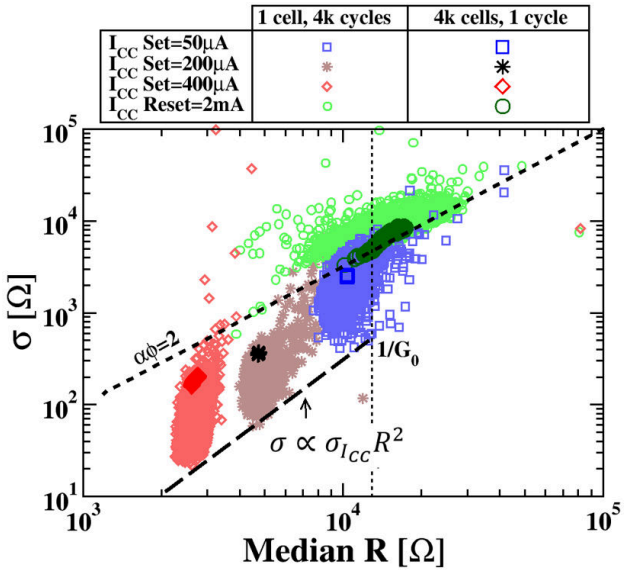
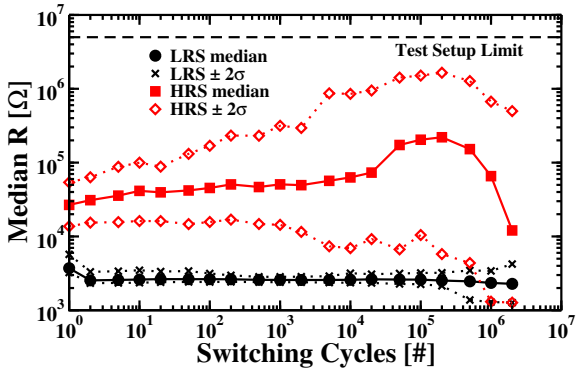


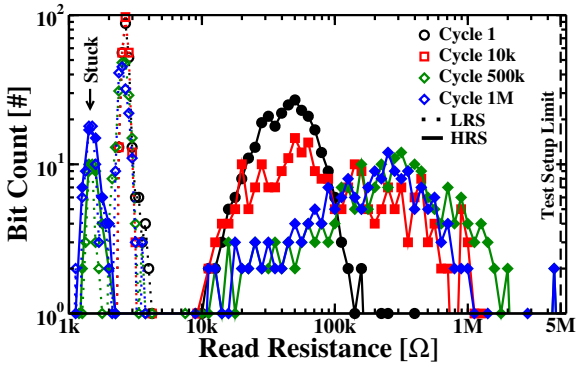
Figure 3.89: Comparison of cell-to-cell and cycle-to-cycle σ vs. median resistances around $1/G_0$ with different Set I_{CC}, V_{BL}=2V, T_{Set} = 100ns and Reset I_{CC}=2 mA, V_{SL}=2.5V, T_{Reset}=100ns.

3.7.5 Variability during cycling

2 million cycles test has been performed on all cells of a 4kbits array with no-verify condition for Set and Reset operations (Fig. 3.90). A separation of HRS and LRS states can be obtained at 2σ up to 500k cycles. Distributions show that the failure is due to stuck cells at low resistance. However, it is observed that starting at the very first cycle, the median resistance and the dispersion increase continuously before reaching a maximum: afterwards the operating windows decrease rapidly.



(a)



(b)

Figure 3.90: LRS and HRS median $\pm 2\sigma$ resistances measured during 2M cycles performed with Set I_{CC}=0.4mA, V_{BL}=2V, T_{Set}=100ns and Reset I_{CC}=2.2mA, V_{SL}=2.5V, T_{Reset}=100ns (a). Distributions measured in different cycling conditions (b).

This behavior, very difficult to be observed on single cells, is observed for all cycling conditions in the array (Fig. 3.91). The stronger the Reset condition, the higher R_{HRS} and the shorter the cycling performance can be obtained. The tradeoff R_{HRS}/R_{LRS} vs. cycle is comparable to state of the art published filamentary RRAM data (Fig. 3.92).

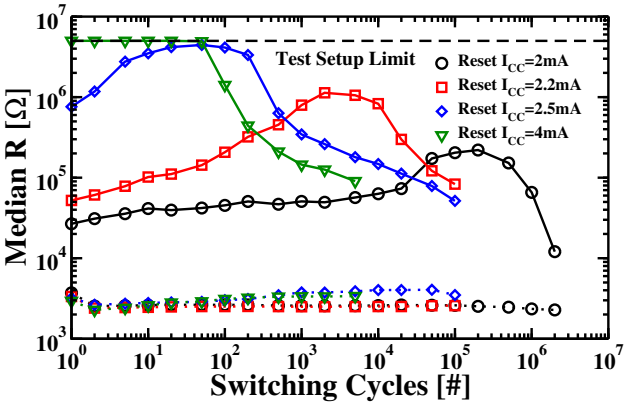


Figure 3.91: Comparison of LRS and HRS median resistances in cycling with Set I_{CC}=0.4 mA, V_{BL}=2V, T_{Set}=100ns, V_{SL}=2.5V, T_{Reset}=100ns and different Reset I_{CC}.

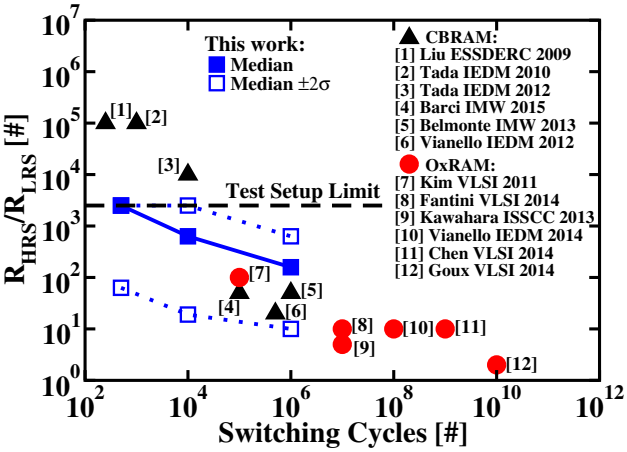


Figure 3.92: R_{HRS}/R_{LRS} vs. switching cycles. Median resistance $\pm 2\sigma$ results are reported and compared with literature. Limits in the test setup bound R_{HRS}/R_{LRS} ratios measurement to 2500.

The increase of dispersion vs. resistance during cycling (Fig. 3.93) follows Eq. 3.9 universal law before showing a higher dispersion. The evolution of the ratio $\sigma/R^{1.5}$ shows that even if initially the dispersion increases, this is not related to a degradation mechanism. After cycling, similarly with too strong Forming conditions or too strong Reset conditions, the $\sigma/R^{1.5}$ ratio increases. This is due to a decrease of $\alpha\phi$, which can be explained as a degradation of the barrier with cycling (Fig.3.94).

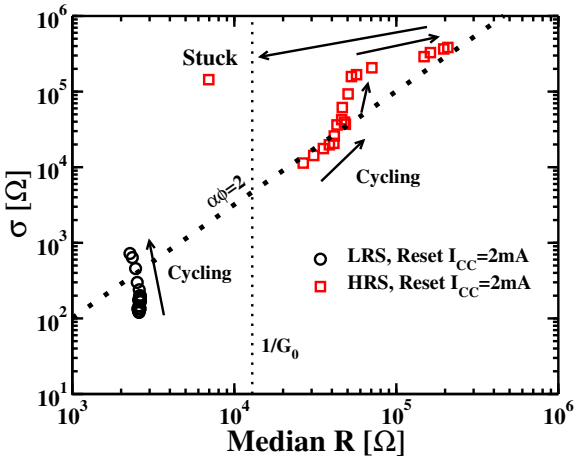


Figure 3.93: σ vs. median resistances during 2M cycles with conditions of Fig. 3.90. A gradual σ increase is observed, causing deviations from $\alpha\phi$ initial value.

In conclusion, an extensive study of variability of OxRAM on multi-kbits RRAM arrays over the full operation range has been presented. In contrast to all papers based on single cell analysis with extensive cycle-to-cycle analysis, OxRAM array distributions are extremely stable and predictable. A resume of fundamental variability limits in case of R below and above $1/G_0$ is reported in Tab. 3.16. The minimal variability of filamentary-based RRAM achievable for median resistances above $1/G_0$ or after Reset is bounded to Eq. 3.9. This minimal variability law is universal for all filament-based RRAM. In this condition, the cell-to-cell variability is smaller than the cycle-to-cycle variability. As a result, the single pulse resistance distribution can be minimized only for resistance below $1/G_0$ and after Set. In this condition, the

dispersion is due to the selector variability at Set and proportional to $\sigma_{I_{CC}} R^2$. The cycle-to-cycle dispersion is in this case smaller than the cell-to-cell dispersion and defined by the operating conditions. Furthermore, this study gives a metric to quantify the variability degradation of filamentary based RRAM during cycling.

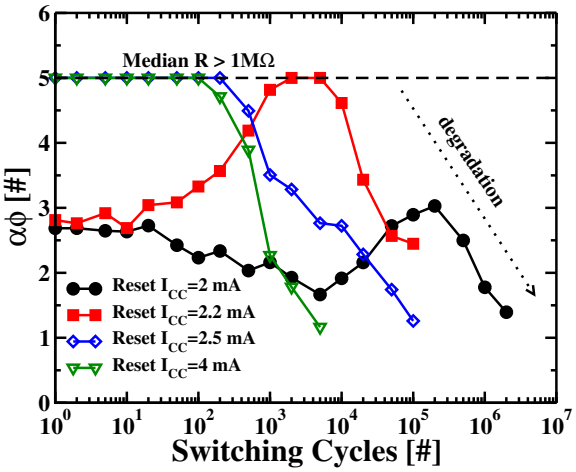


Figure 3.94: $\alpha\phi$ evolution during cycling: Higher R_{HRS}/R_{LRS} ratio corresponds to higher initial $\alpha\phi$ and faster degradation.

Table 3.16: Resume of fundamental variability limits of filamentary based RRAM.

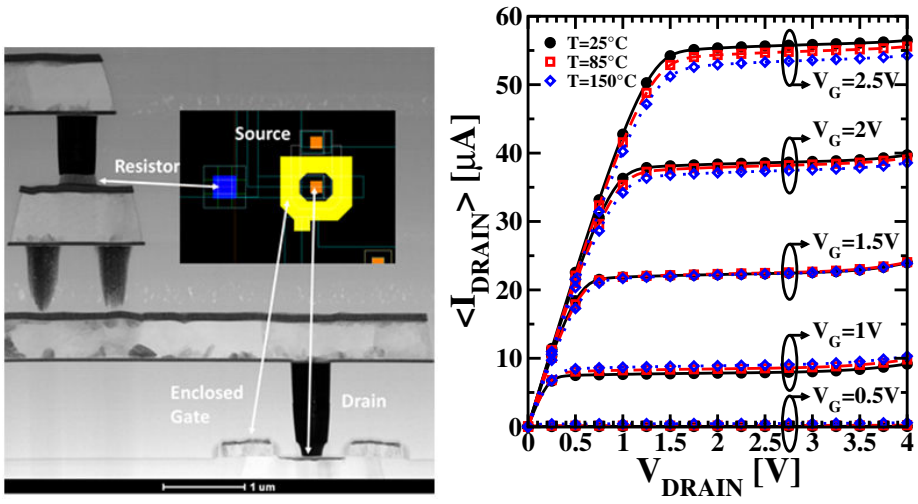
	R< $1/G_0$ AND after Set	R> $1/G_0$ OR after Reset
Variability limiting factor	I_{CC} variability at Set condition	N° of vacancy/atoms (Poisson Law)
Resistance distribution	Gaussian	Lognormal
Limiting component	Selector (1T)	Resistor (1R)
σ (cell-to-cell)	$\sigma \propto \sigma_{I_{CC}} R^2$	$\sigma = \sqrt{G_0 e^{-\alpha\phi}} R^{1.5}$
Cell-to-cell vs. cycle-to-cycle	$\sigma_{cycle-to-cycle} < \sigma_{cell-to-cell}$	$\sigma_{cycle-to-cycle} > \sigma_{cell-to-cell}$

3.8 Radiation hard application perspectives

Semiconductor memories, among rad-hard integrated circuit scenario, are one of the most critical topics for space applications. Actually both volatile and non-volatile memories are integrated using standard processes and standard architectures. This means that the final device is typically at least Rad-tolerant and not Rad-Hard and failure during mission is avoided using Error Correcting Code techniques including redundancy at the board level. Standard silicon memories such as flash memories tend to fail under irradiation, whereas RRAM are intrinsically radiation tolerant: the use of RRAM could significantly reduce the radiation-induced bit upsets [107]. Nevertheless, the 1T-1R structure of the memory array consists of NMOS access transistors, which are sensitive to radiation [108]. In standard NMOS devices, ionizing radiation may generate holes trapped in the gate oxide, and the trapped holes could induce leakage paths from the drain to the source region. A suitable approach to eliminate the leakage path in NMOS transistors is to adopt a gate-enclosed layout [109]. In this section a Rad-Hard designed Enclosed Layout Transistor (ELT) to be integrated with a TiN/HfO_2/Ti/TiN based resistor is characterized and a memory architecture suitable for Rad-hard RRAM arrays is characterized [110].

The TEM cross view of the Rad Hard designed 1T-1R cell and the ELT Current-Voltage characteristics measured at different temperatures are reported in Fig. 3.95. The drain current versus gate voltage characteristics after irradiation of the ELT is illustrated in Fig. 3.96. The leakages are almost unaffected during the total ion dose (TID) irradiation up to 750 kRad.

In order to achieve an increased resistance against radiation an architectural solution is proposed where the single bit is the result of the contribution of two RRAM cell located in different array locations; this guarantees an internal redundancy (no reference cells are required for read mode operations) and a wider margin window in a sensing module very similar to SRAMs. In Fig. 3.97 is shown the architectural approach where two arrays (left and right) contain left cells and rights cells: a logic 1 is the result of a low re-



(a) (b)

Figure 3.95: TEM cross view of the Rad Hard designed 1T-1R cell (a). Current-Voltage characteristics of the access ELT (b).

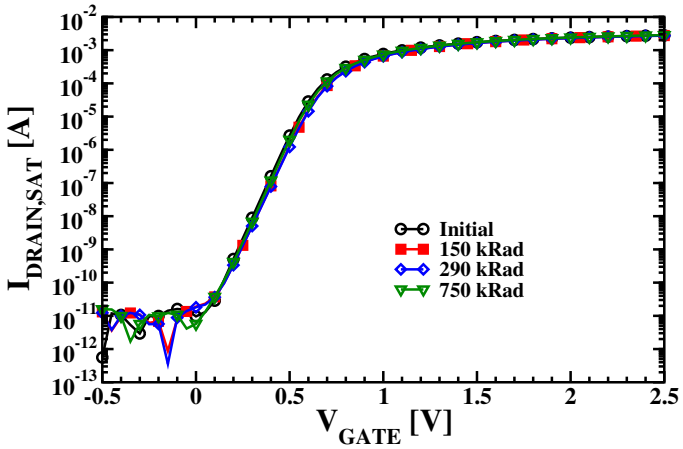


Figure 3.96: Post-irradiation Current-Voltage characteristics of the access ELT with varied dose rate.

sistance in the left cell (red circle) and a high resistance in the right cell (green circle), vice-versa for a logic 0. The differential cell approach, thanks also to the independence of row decoding final stages and column decoding switches, guarantees resistance against Single Event Effects (SEEs) disturbs in all conditions (Read, Set and Reset modes).

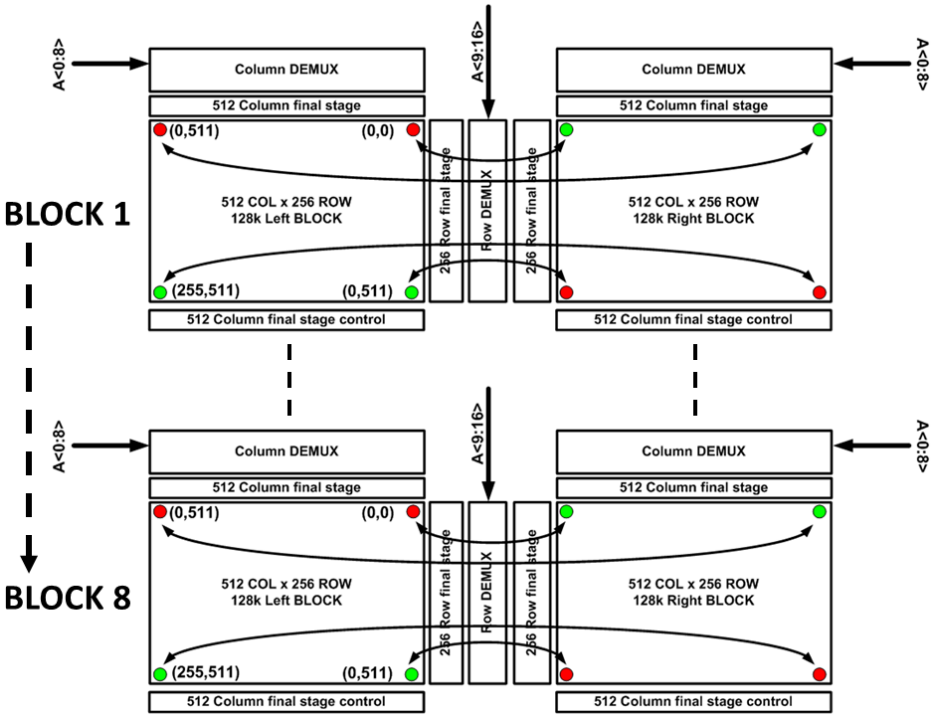


Figure 3.97: The Architecture of the 1Mbit (2Mcell) test vehicle implementing RRAM array.

In the proposed design, the 1Mbit device will be the result of the ensemble of eight 128kbit modules (see Fig. 3.97) each having its own decoding scheme, ATD (Address Transition Detection) and sense amplifier. This architectural approach protect the test vehicle against Multiple Bit Upset (MBU) reducing every contribute from charged particle to at least only one Single Event Upset (SEU) as already demonstrated in previous SRAM devices [111, 112]. Even if complete of all decoding schemes the 1Mbit (2Mcells) device should also contain Direct Memory Access (DMA) in order to provide the possibility to characterize the behavior of RRAM cells independently from the sensing scheme and Set/Reset circuitry. This degree of freedom enables the access via ATE equipment for all analysis related to LRS and HRS state distributions, allowing the characterization of the resistive state of the cell in different programming conditions.

Chapter 4

TAS-MRAM

The discovery of tunnel magnetoresistance (TMR) at room temperature [113, 114] in 1995 has triggered a lot of interest in both fundamental and applied research [115]. This phenomenon is of particular interest for data storage technological application such as hard drive disks or magnetoresistive random-access memories (MRAMs). For that latter application, a magnetic tunnel junction (MTJ) can be used to store binary data. Indeed, the two resistance states of a MTJ are a way to encode logic '1' or '0'. Magnetic Random Access Memories (MRAM) are nowadays one of the most promising candidates to replace traditional Flash in future non-volatile memories generations due to its high speed, endurance and scalability [116, 117]. Among the MRAM paradigms that are under investigation, the Thermally Assisted Switching represents a good candidate for a replacement of the standard flash memories in embedded environments [118–120]. In this chapter, the performance of 1 kbit TAS-MRAM arrays will be characterized with the goal of defining the optimal working conditions and evaluating its performance. After that, a novel TAS-MRAM array with optimized read procedure (Self Referenced) will be tested and compared with the previous one to highlight its advantages in terms of reliability.

4.1 Basics

A basic MRAM cell consists of a single MTJ connected in series with a selection transistor that can be viewed as a switch that allows or not a current flowing through the MTJ. The MTJ consists in two ferromagnetic layers separated by an insulating layer. For MRAM applications, one ferromagnetic layer is pinned in one direction and cannot be switched under functioning conditions (the reference layer), while the second ferromagnetic layer stores the binary information. Its magnetization has to remain stable during standby and reading, while it has to be easily switchable under the writing conditions. During reading, the selection transistor is turned on so that a current flows through the junction. A reference current is also injected in a reference resistance, with its value being between the high and the low resistance states of the MTJ. The MTJ current and the reference current are compared thanks to a sense amplifier that converts these two currents into a logical '1' or '0', depending on the MTJ resistance state. This reading scheme works only when the two MTJ resistance states are well separated among the memory cells.

A new writing scheme able to improve the downsize scalability of MRAM has been demonstrated at Spintec laboratory in 2002 [121] and then developed and industrialized by Crocus Technology. This new writing approach relies on thermally assisted switching (TAS) and allows to improve write selectivity, power consumption, and thermal stability. The general idea consists in temporarily heating the storage layer during write in order to ease the switching of its magnetization while ensuring a strong thermal stability of this magnetization orientation in standby. In thermally assisted MRAM (TAS-MRAM), the storage layer is a ferromagnetic layer pinned with a low blocking temperature (TB) antiferromagnet, such as FeMn (90-160 °C) or IrMn (120-260 °C). The reference layer is a SyF pinned with a high TB antiferromagnet, such as PtMn (350 °C). In standby mode, the storage layer presents a very high thermal stability because it is pinned by the low TB antiferromagnet. The TAS writing procedure is depicted in Fig. 4.1.

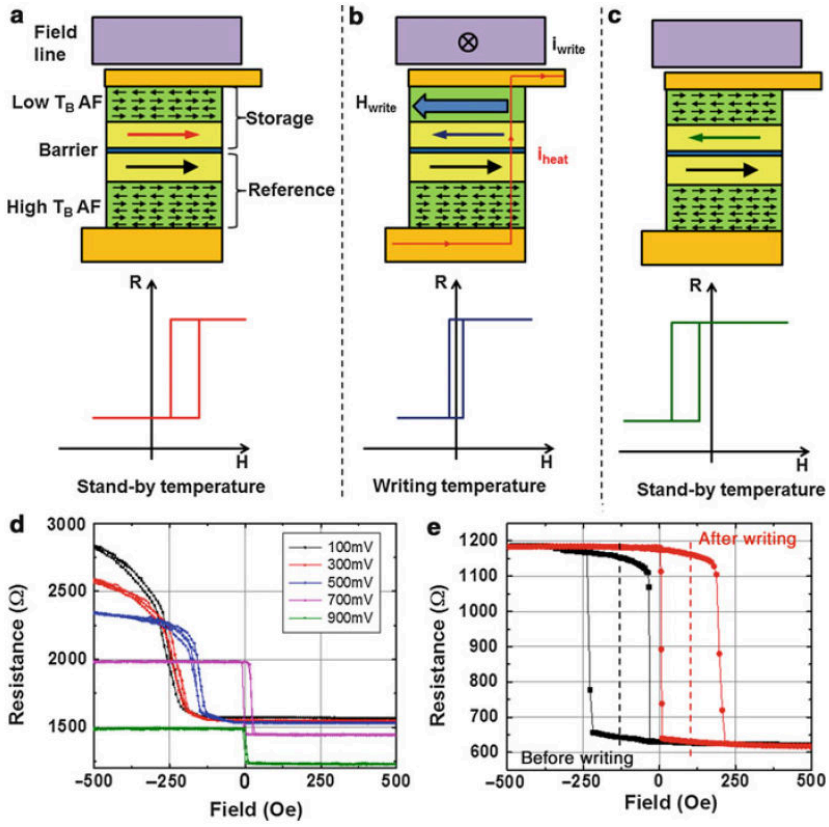


Figure 4.1: Principle of TAS: (a) the storage layer is pinned by an antiferromagnet with low TB at the standby temperature. (b) During writing, a current pulse is injected in the junction that heats the storage layer above TB. In that situation, the storage layer is unpinned and can be easily switched by a magnetic field generated by a single field line. (c) When the heating current is stopped, the junction cools down below TB and the storage layer is pinned by the antiferromagnet in the opposite direction. (d) MTJ with an exchange biased storage layer. Illustration of the decrease in the storage layer loop shift as a function of the heating voltage through the tunnel barrier. Note that when the heating voltage is large enough (700 mV), a small field of a few tens of Oe is sufficient to switch the storage layer magnetization. (e) Illustration of the inversion of the storage layer loop shift in standby between the initial state ('1') and the final state ('0') after writing.

In order to switch the storage layer, a combination of field and heating current is required. In the initial state, the storage layer magnetization is pinned in one direction by exchange bias, so that only one resistance state is possible at zero field. During writing, a current pulse is injected in the MTJ. Due to the inelastic relaxation of electrons tunneling through the barrier, heating occurs in the junction. The heating current is tuned so that the temperature of the storage layer slightly exceeds the blocking temperature of the storage antiferromagnet, but remains below the reference antiferromagnet one. In this situation, the storage layer magnetization gets unpinned, thus enabling its switching. The advantage of the thermal assistance is this possibility to lower the energy barrier for switching thanks to the temporary heating of the cell. Once heated, the storage layer magnetization is then switched thanks to a field generated by a single field line. The heating current is then stopped. The MTJ cools down under the applied field that saturates the storage layer magnetization in the appropriate direction [122]. During cooling to the standby temperature, the storage layer recovers its pinning and correlatively its high thermal stability. Meanwhile, the reference layer is not affected by the writing magnetic field because the temperature in the MTJ remains below the reference antiferromagnet TB. This writing scheme allows circumventing the classical dilemma between the write consumption and retention in memory technology. In the standard writing schemes, increasing the thermal stability factor for improved scalability also increases the write energy consumption, since the writing field is directly proportional to the thermal stability factor. This issue is solved in the TAS approach because heating allows reducing the barrier of energy during writing while the thermal stability factor can be as high as required at the standby temperature. In order to prevent read disturbance, the reading current is about three times lower than the writing current, the injected power density being thus almost one order of magnitude lower than during the writing cycle.

TAS approach also presents many advantages compared to conventional MRAM architectures:

- Since the write selection is temperature driven, a combination of magnetic field with heating current is required to switch the junction. There is no more parasite writing due to half-selected bits.

- The magnetic anisotropy is now provided by the antiferromagnet. The use of circular elements instead of ellipses is then possible since shape anisotropy is not necessary in contrast to toggle MRAM. This simplifies the lithography and etching steps and gives the maximum areal bit density.

- Since the system is not bistable anymore at zero field, thanks to the exchange bias provided by the antiferromagnet, TAS provides good reliability under field disturbance. Indeed, in standby conditions, even if the resistance state of a bit is changed by external parasitic fields, the resistance state after the field perturbation goes back to its initial value.

- TA-MRAMs present a good scalability since the power density required to heat the junction is proportional to the square of the current density. This technology thus scales with the junction area.

- Only one field line is required to write a bit, contrary to conventional technologies. This leads to reduced power consumption, especially because heating the junction requires less current than generating a field (typically, the heating current is in the order of 0.1-0.5 mA, while the writing current is in the order of 5-20 mA at the 130 nm technological node). Moreover, using circular elements allows decreasing the field required to switch the storage layer magnetization under write conditions. Indeed, in that case, the shape anisotropy term is cancelled, and the barrier of energy is reduced to the magnetocrystalline anisotropy energy of the ferromagnet that is engineered to be as low as possible.

- Finally, it becomes possible to use a single field line to switch simultaneously several bits, further reducing the write consumption. The

writing procedure consists in injecting a current with a first polarity in the field line, the generated field being exerted on an assembly of MTJ. A heating current is injected in all the junctions, all bits being thus written in state '1'. The current injected in the junctions that have to be written in the '1' state is stopped, while the current is maintained in the junctions that have to be written in the '0' state. Then, a current is injected in the field line with the opposite polarity, switching the bits that have to be written in the '0' state. An example of such a writing procedure is depicted on Fig. 4.2. N bits can be thus written with only two pulses of magnetic field (for comparison, at least N pulses of magnetic field would be required in toggle writing).

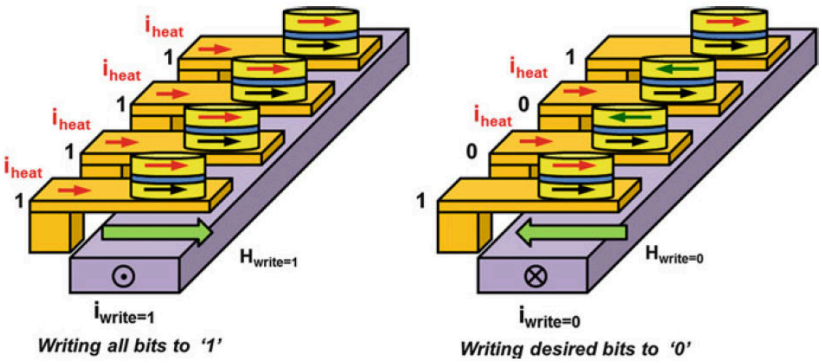


Figure 4.2: Multiple bits writing scheme. First, all bits are written in state '1' by injecting a heating pulse in all junctions while applying a magnetic field in the '1' direction. The heating current is then switched off for the junctions that have to be written in the '1' state and maintained in the junctions that have to be written in the '0' state. Then, the magnetic field is reversed in the '0' direction, switching all junctions that have to be written in the '0' state. This writing procedure allows writing N-bit-long words with only two magnetic fields and thus reducing the overall writing power consumption.

4.2 Experimental Setup

The 1kbits memory device integrated into a CMOS process is made of a 32x32 array. All measurements were performed with RIFLE ATE. The cell and the test array architecture are depicted in Fig. 4.3, where:

- MTJ is the Magnetic Tunnel Junction device, composed of two ferromagnetic layers separated by an insulating layer;
- SP_1, SP_2 and SP_3 are sense pads used during read. SP_1 is on the top of the MTJ, SP_2 is connected right below the MTJ and SP_3 is between a poly 500 Ω resistance and a select transistor.

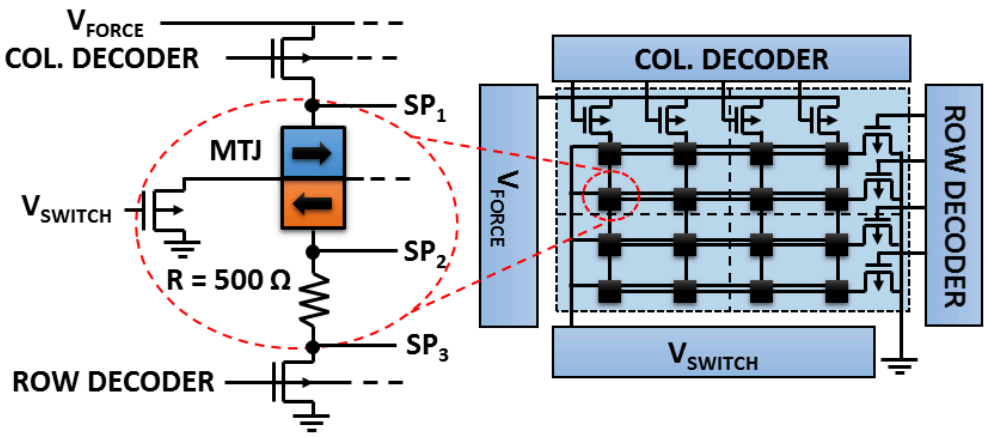


Figure 4.3: Structure of a TAS-MRAM cell and its integration into the array tested in this work.

In order to change the state of a memory cell, two different writing operations are available: Write '0' (W0) and Write '1' (W1). Both operations require two voltages: V_{FORCE} is required to locally heat the magnetic material, whereas V_{SWITCH} allows changing the magnetic field polarization after heating. All write operations have been performed with $T_{FORCE} = 500ns$, $T_{SWITCH} = 600ns$ and $T_{rise/fall} = 500ns$ for both voltages in order to avoid overshoot issues. All read operations have been performed with $V_{SWITCH} = 0V$, $V_{FORCE} = 0.3V$, $T_{FORCE} = 10\mu s$ and $T_{rise/fall} = 1\mu s$.

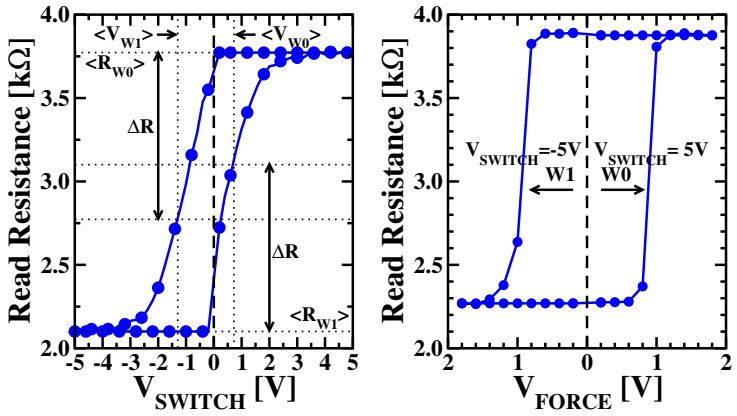
4.3 TAS-MRAM arrays characterization

In this section the reliability and the cell-to-cell variability during 500k endurance cycles have been evaluated by extracting a set of characteristic parameters from measurements performed on 1kbits arrays, following the guidelines provided in [93]. After a preliminary optimization of the writing parameters on fresh devices, the effectiveness of the selected parameters has been verified during cycling by evaluating their impact on cell-to-cell variability and on the reliability lowering due to the cell breakdown [123, 124].

To evaluate the impact of heating and switching voltages on write operations, a preliminary hysteresis analysis of both parameters has been performed on fresh devices. $|V_{SWITCH}|$ has been increased from $0.2V$ to $5V$ with $|\Delta V_{SWITCH}| = 0.2V$ and $V_{FORCE} = 1.4V$. The same procedure has been applied to evaluate V_{FORCE} hysteresis by increasing V_{FORCE} from $0.2V$ to $1.8V$ with $\Delta V_{FORCE} = 0.2V$ and $|V_{SWITCH}| = 5V$. A read operation has been performed after every step in both hysteresis analysis. Fig. 4.4 (a) shows the average resistances measured during switching voltage hysteresis and the switching parameters extracted for further analysis of W0 and W1 operations with $V_{FORCE} = 1.4V$:

- $\overline{R_{W0}}$ and $\overline{R_{W1}}$ are the average values of resistance R_{W0} and R_{W1}, respectively measured at $V_{SWITCH} = 5V$ and $V_{SWITCH} = -5V$.
- $\overline{V_{W0}}$ and $\overline{V_{W1}}$ are the average switching voltages V_{W0} and V_{W1}, respectively, that allow obtaining a variation $\Delta R = 1k\Omega$ of the average measured resistance values.

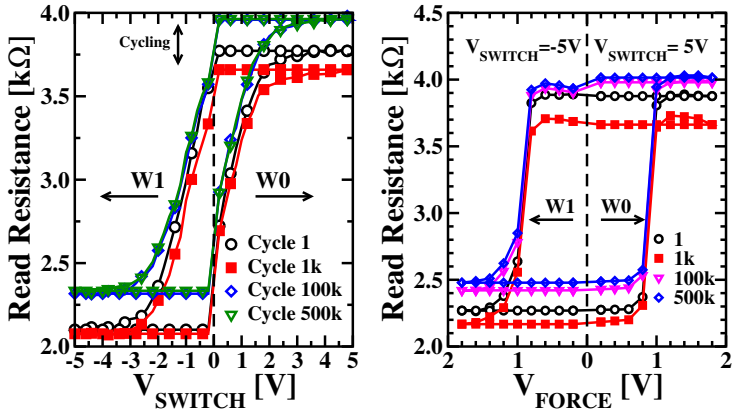
Fig. 4.4 (b) shows the average resistances of the array cells measured during heating voltage hysteresis in W0 and W1 operations with $|V_{SWITCH}| = 5V$. It can be observed that $V_{FORCE} \geq 1V$ is required in order to successfully switch the magnetic field, whereas using $V_{FORCE} > 1.6V$ shows no advantages in terms of average resistance for both W0 and W1 operations. $V_{FORCE} = 1.4V$ and $|V_{SWITCH}| = 5V$ are shown to be the optimal write conditions, ensuring the highest resistance difference in W1 and W0 states.



(a) (b)

Figure 4.4: V_{SWITCH} (a) and V_{FORCE} (b) hysteresis on fresh devices.

To evaluate the cells performance and reliability during cycling and the effect of the cell degradation, 500k W0 and W1 operations have been performed with $V_{FORCE} = 1.4V$ and $|V_{SWITCH}| = 5V$. Fig. 4.5 shows the switching voltage (a) and the heating voltage (b) hysteresis evolution during cycling: an equal R_{W0} and R_{W1} variation can be observed, thus keeping the resistance difference constant during cycling.



(a) (b)

Figure 4.5: V_{SWITCH} hysteresis measured during W0 and W1 operations at different cycles with $V_{FORCE} = 1.4V$ (a). V_{FORCE} hysteresis measured during W0 and W1 operations at different cycles with $|V_{SWITCH}| = 5V$ (b).

Fig. 4.6 shows the cumulative distributions of the read resistances measured after W0 and W1 operations at different cycles. Left tails on the distributions appear during cycles due to the cell degradation for a limited percentage of cells (below 3% after 500k cycles).

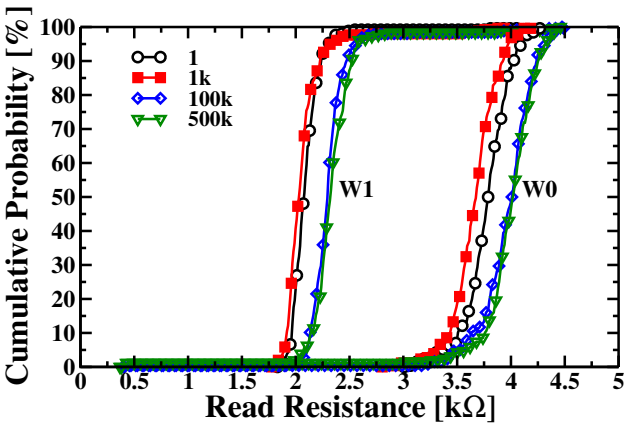


Figure 4.6: R_{W0} and R_{W1} cumulative distributions measured during 500k cycling with $V_{FORCE} = 1.4V$, $|V_{SWITCH}| = 5V$.

In order to evaluate the impact of V_{FORCE} during cycling, 500k cycles have been performed with different V_{FORCE} values and $|V_{SWITCH}| = 5V$, measuring R_{W0} and R_{W1} at different cycles. $\overline{R_{W0}}$ and $\overline{R_{W1}}$ measured during cycling are reported in Fig. 4.7, showing a sudden cell degradation after 100k cycles with $V_{FORCE} = 1.6V$ and after 1k cycles with $V_{FORCE} = 1.8V$, whereas the average resistances do not show any relevant change during the endurance tests with $V_{FORCE} = 1.2V$ and $1.4V$.

The dispersion coefficients (i.e. standard deviation over mean value) for R_{W1} and R_{W0}, evaluated during cycling with different V_{FORCE} conditions, are reported in Fig. 4.8: a rapid increase of their values can be observed before the breakdown with $V_{FORCE} = 1.6V$ and $V_{FORCE} = 1.8V$ at cycles 50k and 500, respectively. The use of $V_{FORCE} = 1.4V$ induces the lowest cell-to-cell variability of R_{W0} and R_{W1} during cycling.

The cumulative distributions of the V_{W0} and V_{W1} parameters measured during cycling with different V_{FORCE} conditions are reported in Fig. 4.9.

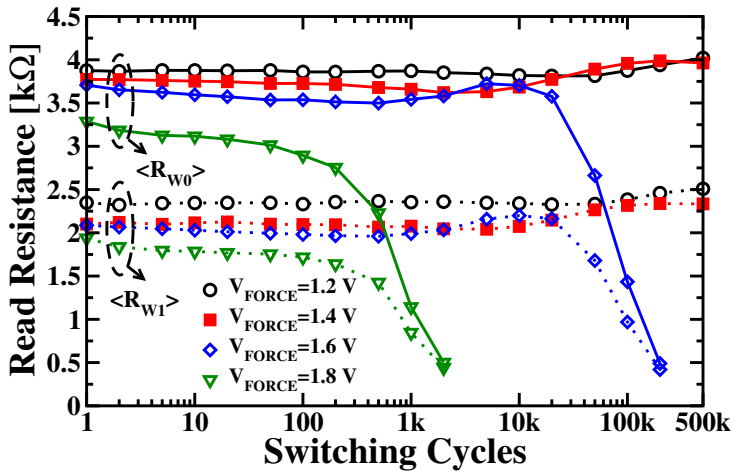
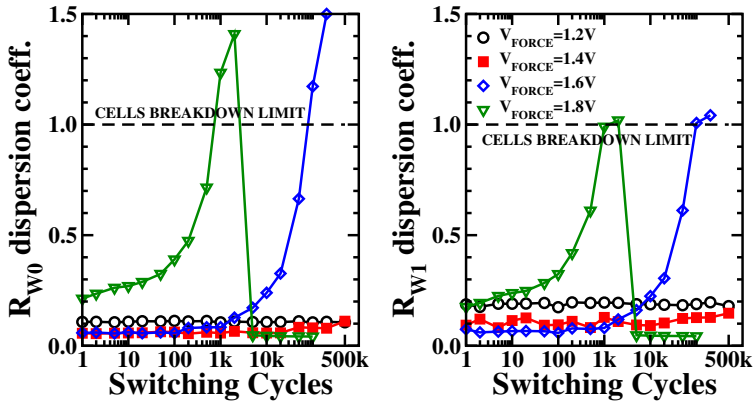


Figure 4.7: $\overline{R_{W0}}$ (full lines) and $\overline{R_{W1}}$ (dotted lines) measured during 500k cycling with different V_{FORCE} conditions.



(a) (b)

Figure 4.8: R_{W0} (a) and R_{W1} (b) dispersion coefficients measured during 500k cycling with different V_{FORCE} conditions.

The cumulative number of cells do not reach 1k because a limited number of cells do not reach, in switching, the assumed $\Delta R = 1k\Omega$. The minimum cell-to-cell variability during the endurance test is obtained by using $V_{FORCE} = 1.4V$. Moreover, $V_{FORCE} = 1.4V$ shows the highest cumulative number of switched cells: this means that a higher cells percentage reached the requested resistance variation $\Delta R = 1k\Omega$. According to the obtained

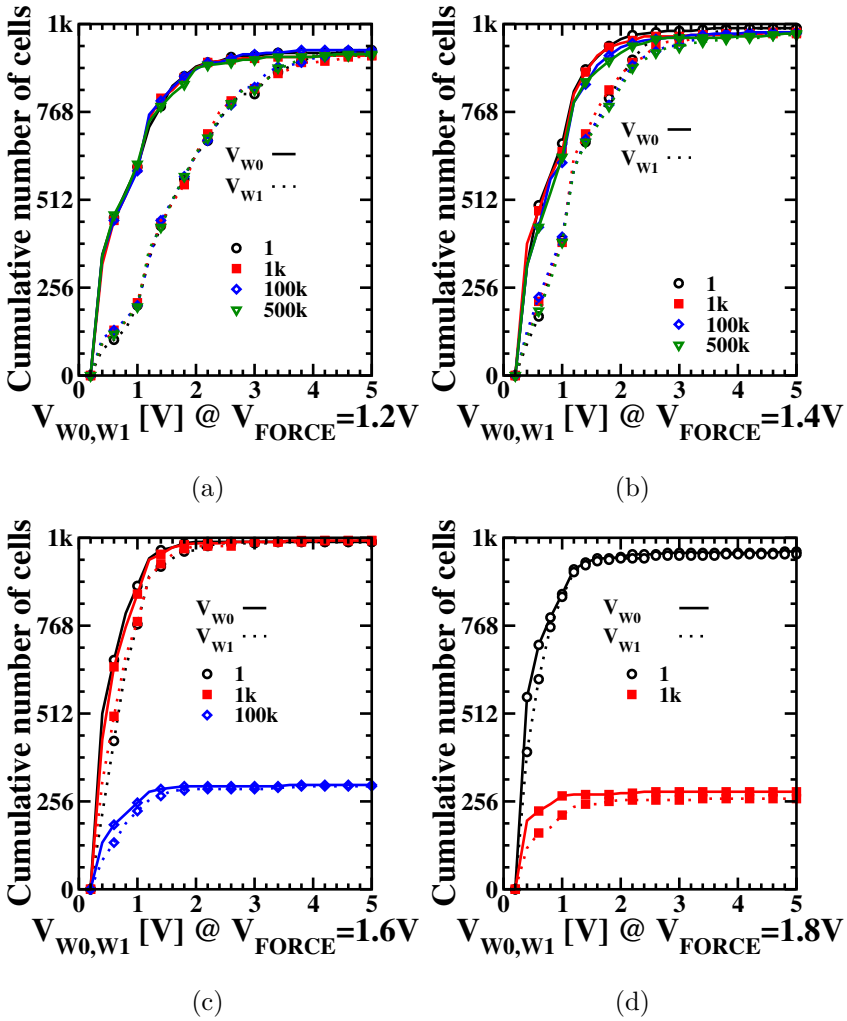


Figure 4.9: Cumulative number of switching cells reaching the assumed variation $\Delta R = 1k\Omega$ with different V_{FORCE} during cycling.

results, the use of $V_{FORCE} = 1.4V$ has to be preferred since it guarantees the best reliability in cycling, denoted as the percentage of cells correctly switching. Using higher heating voltages results in a reduced yield (lower percentages of cells reaching the requested ΔR) and in a faster breakdown. According to the reported analysis, R_{W0} and R_{W1} are shown to depend on V_{FORCE}, V_{SWITCH} and cycling: the relationship between these parameters has been analyzed through 3D plots for different V_{FORCE} conditions for both W0 (Fig. 4.10) and W1 (Fig. 4.11).

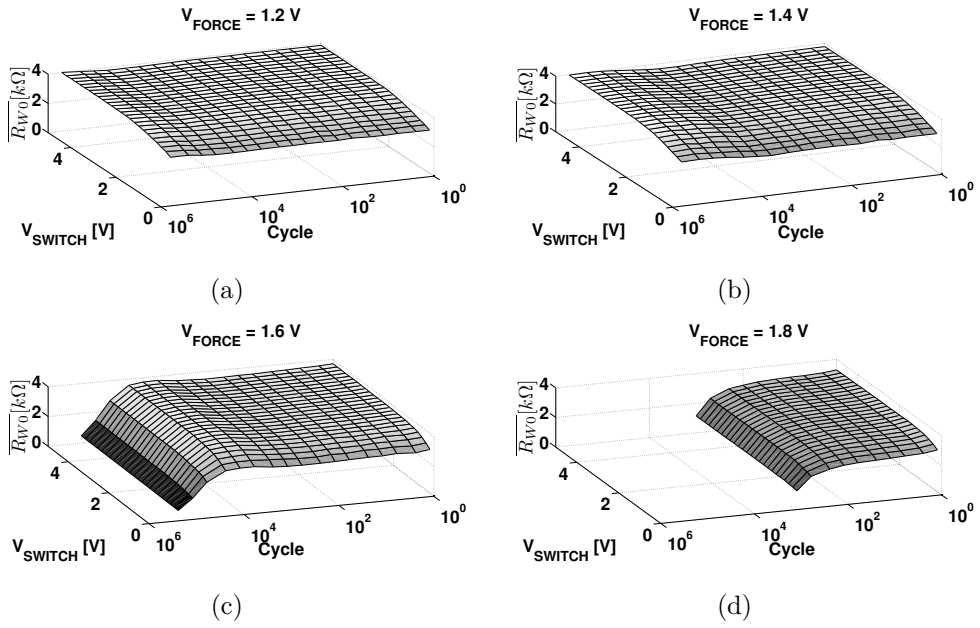


Figure 4.10: 3D plot of $\overline{R_{W0}}$ as a function of V_{SWITCH} and cycle number for different V_{FORCE} conditions.

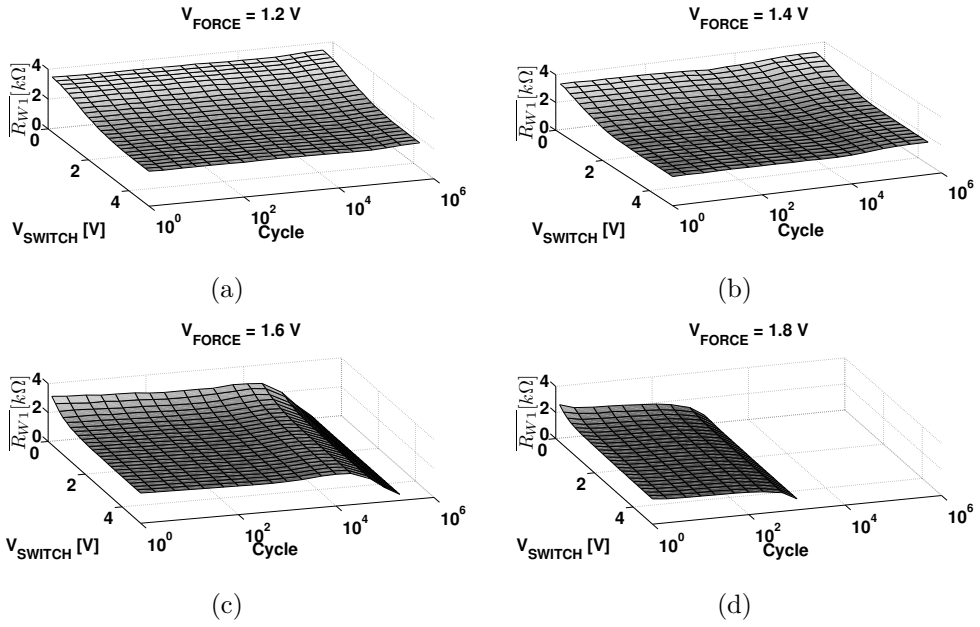


Figure 4.11: 3D plot of $\overline{R_{W1}}$ as a function of V_{SWITCH} and cycle number for different V_{FORCE} conditions.

$\overline{R_{W0}}$ and $\overline{R_{W1}}$ depend on $|V_{SWITCH}|$ for any cycling and V_{FORCE} conditions: $|V_{SWITCH}| = 5V$ allows obtaining the highest $\overline{R_{W0}}$ values and lowest $\overline{R_{W1}}$ values. Using too high heating voltages (i.e. $V_{FORCE} \geq 1.6V$) results in a shorter lifetime, independently from V_{SWITCH}. The most important parameter used to evaluate the switching capabilities on MRAM is the tunnel magnetoresistance (TMR) [125], calculated as:

$$TMR = \frac{R_{W0} - R_{W1}}{R_{W1}} \tag{4.1}$$

TMR cumulative distributions measured during endurance test at different cycles are reported in Fig. 4.12 for each V_{FORCE} condition. Since TMR depends on the difference between R_{W0} and R_{W1} and the resistance shift due to cell degradation is the same on both resistive states, no relevant variations can be observed on TMR until the cell breakdown is reached. $V_{FORCE} = 1.4$ allows obtaining the highest TMR in each cycling condition.

4.4 Self Referenced TAS-MRAM

Field-induced TAS has been also implemented in a second generation of MRAM cells with a self-reference reading scheme [115, 126]. In that configuration, the reference layer is replaced by a soft ferromagnetic layer called "sense layer", while the storage layer is similar to standard in-plane magnetized TAS-MRAM (i.e. a ferromagnetic layer pinned by an antiferromagnet at the standby temperature). A standard self-reference stack thus consists in a F/MgO/F/AF or F/MgO/F/Ru/F/AF multilayer, in which the first F layer is the sense layer, and the second F/AF layers (or F/Ru/F/AF) are the storage layer. The reading scheme is performed in two steps depicted in Fig. 4.13: first, the sense layer magnetization is set in one direction by the magnetic field generated by the field line, and the MTJ resistance is measured. The sense layer is then switched in the opposite direction and the new MTJ resistance is measured. The resistance variation (either positive or

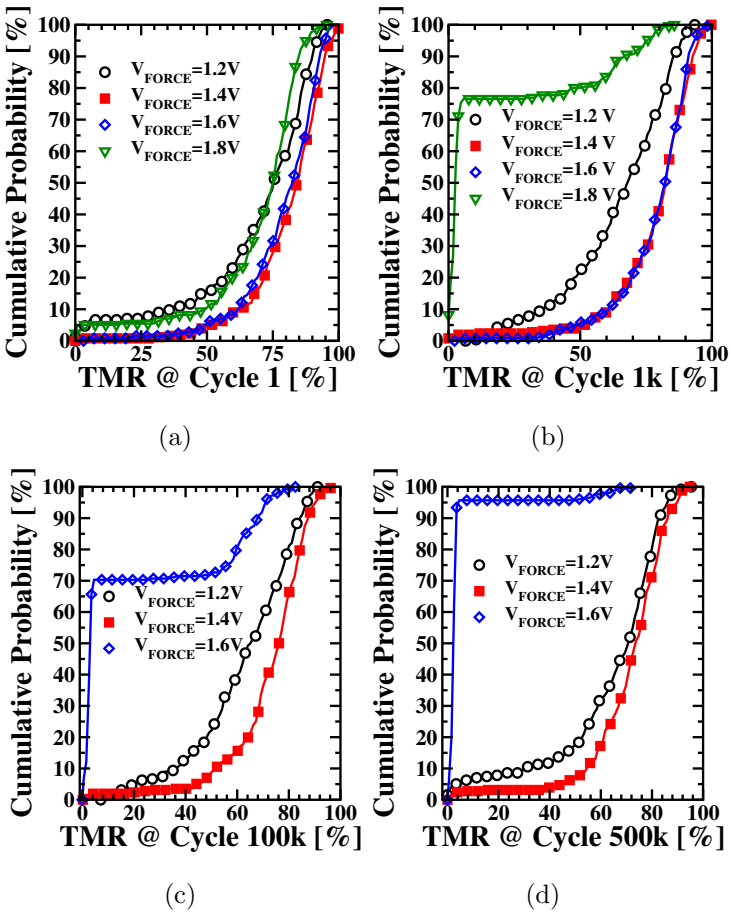


Figure 4.12: TMR measured during cycling with different V_{FORCE} conditions at cycle 1 (a), 1k (b), 100k (c) and 500k (d) .

negative) between the two measurements yields the magnetic orientation of the storage layer and thus the stored data.

In this approach, the read cycle is longer than in the standard method, but this self-reference reading approach strongly improves the tolerance to process variation compared to the standard resistance measurement reading scheme. Indeed, MTJ diameter or tunnel barrier thickness variations lead to large spreads of minimum and even more maximum MTJ resistances. The 12 σ separation criterion may not be fulfilled at small technological nodes, and thus the standard reading scheme could not be reliably used to read the stored data. This issue is solved with a self-reference reading scheme.

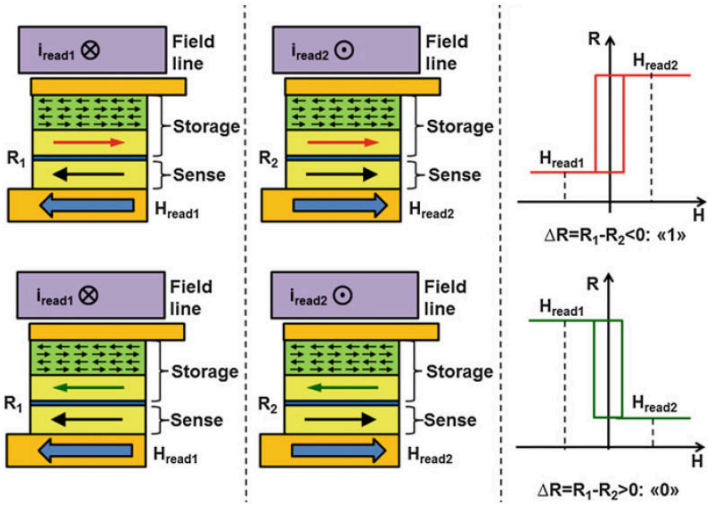


Figure 4.13: Self-reference reading scheme. The sense layer is switched in a first predetermined direction (left) and then in the opposite one (middle). The MTJ resistance is measured after both steps. The difference of resistance between the two steps, either negative (right, top) or positive (right, bottom), yields the storage layer pinning direction and thus the stored data.

Moreover, the TMR of the MTJ does not need to be as high as in standard TAS-MRAM. Indeed, since only the sign of the resistance change is used to read out the stored data, a TMR ratio of 10 % is sufficient to ensure a correct reading of the cell. The writing of the storage layer is achieved similarly to TAS-MRAM devices by the application of a combination of an external field and a heating pulse that allows the storage layer to be switched and then pinned in the opposite direction as the system cools back below its antiferromagnet blocking temperature. During writing, the pulse applied on the field line is always longer than the heating pulse to ensure that the storage layer is still pinned in the desired direction while cooling down after switching.

In this section, 1 kbit self-referenced TAS-MRAM arrays are characterized in terms of hysteresis, endurance and retention. Test setup, cells and array structures are the same of 4.2: the only difference is in the sensing layer.

4.4.1 Hysteresis analysis

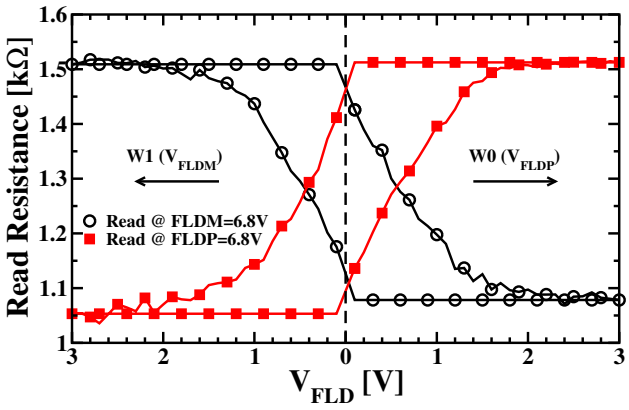
To evaluate the impact of heating (V_{IOF}) and switching (V_{FLD}) voltages on write operations, a preliminary hysteresis analysis of both parameters has been performed on fresh devices. The field line voltage hysteresis was measured by applying write pulses with $V_{PULSE,IOF}$ = 1.6V, $T_{PULSE,IOF}$ = 500ns, $V_{PULSE,FLD}$ increased from 0.1 to 3V with V_{STEP} = 0.1V, $T_{PULSE,FLD}$ = 700ns and $T_{RISE/FALL}$ = 100ns for both pulses. The read resistances measured during hysteresis are reported in Fig. 4.14 (a). In this case, two reading operations are always performed where the sense layer magnetization is set in opposite directions by the field lines FLDM and FLDP by applying 6.8V on the field line for 70 μ, with V_{IOF} fixed to 0.3 V for 50 μ.

The IOF line voltage hysteresis was measured by applying write pulses with $V_{PULSE,IOF}$ increased from 0.1 to 2V with V_{STEP} = 0.1V, $T_{PULSE,IOF}$ = 500ns, $V_{PULSE,IOF}$ increased from 0.1 to 3V with V_{STEP} = 0.1V, $T_{PULSE,FLD}$ = 700ns and $T_{RISE/FALL}$ = 100ns for both pulses. The read resistances measured during hysteresis are reported in Fig. 4.14 (b). It is possible to observe from the hysteresis results that V_{FLD} >2 V and V_{IOF} > 1.4 V are required to obtain a proper switching.

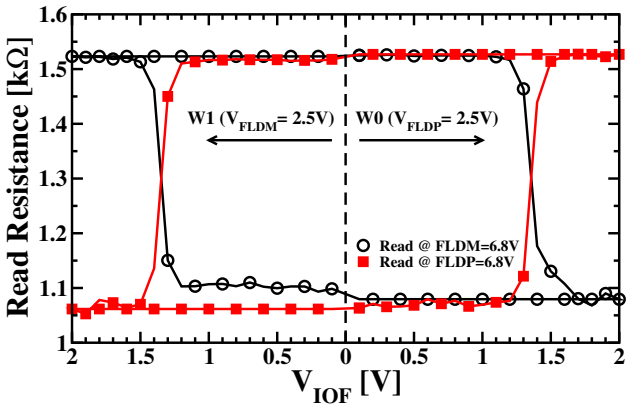
4.4.2 Endurance

To evaluate the cells performance and reliability during cycling and the effect of the cell degradation, 500k Write "0" and Write "1" operations have been performed. Read and Write conditions used during cycling are reported in Tab. 4.1. Fig. 4.15 shows the cumulative distributions of the read resistances measured after Write "0" and "Write "1", measured at cycle 1 and after 500k cycles. Left tails on the distributions appear due to the cell degradation. The differential resistance, calculated as $\Delta R = R_{read,FLDP} - R_{read,FLDM}$ is difference between the read resistances measured when the field line voltage is applied on FLDP and FLDM, respectively.

Fig. 4.16 shows ΔR average values and the standard deviation calculated



(a)



(b)

Figure 4.14: FLD (a) and IOF (b) voltage hysteresis.

Table 4.1: Read and Write operations parameters.

Operation	V_{IOF}	T_{IOF}	V_{FLDP}	T_{FLDP}	V_{FLDM}	T_{FLDM}
Read FLDP	0.3 V	50μs	6.8 V	70μs	-	-
Read FLDM	0.3 V	50μs	-	-	6.8 V	70μs
Write "0"	1.6 V	500 ns	2.5 V	700 ns	-	-
Write "1"	1.6 V	500 ns	-	-	2.5 V	700 ns

after Write "0" and Write "1" operations during 500k cycles. The cumulative distributions ΔR obtained after Write "0" and Write "1" operations, at cycle 1 and after 500k cycles are reported in Fig. 4.17. It can be observed that

even if the resistance window is smaller compared to standard TAS-MRAM, the differential read allows to successfully separate the ΔR distributions.

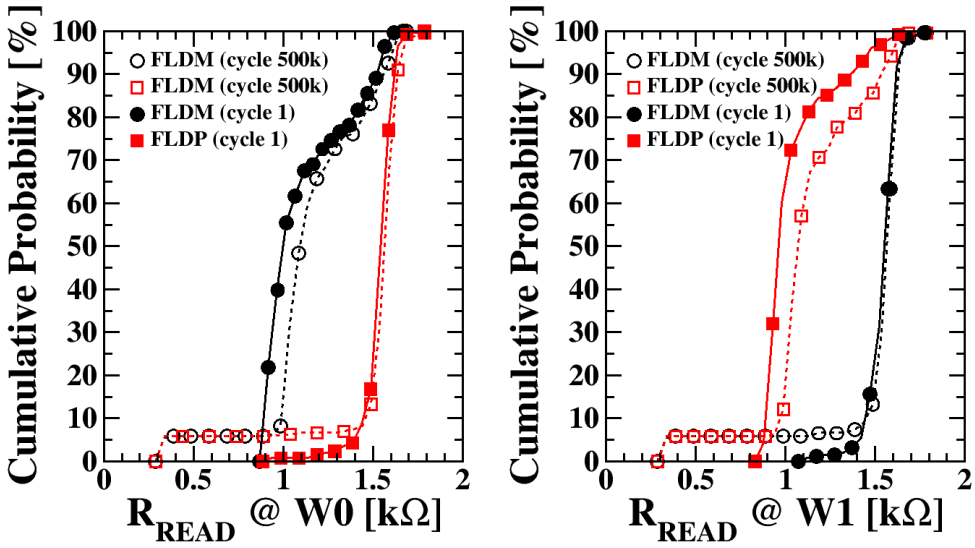


Figure 4.15: Cumulative distributions of Read Resistances measured after Write "0" (left) and Write "1" (right).

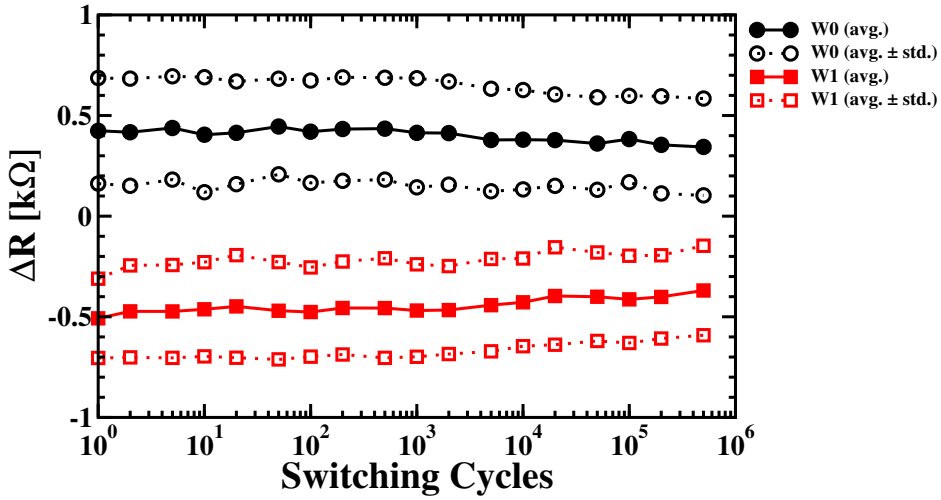


Figure 4.16: Average value and standard deviations of the differential read resistances measured after Write "0" and Write "1" during 500k endurance cycles.

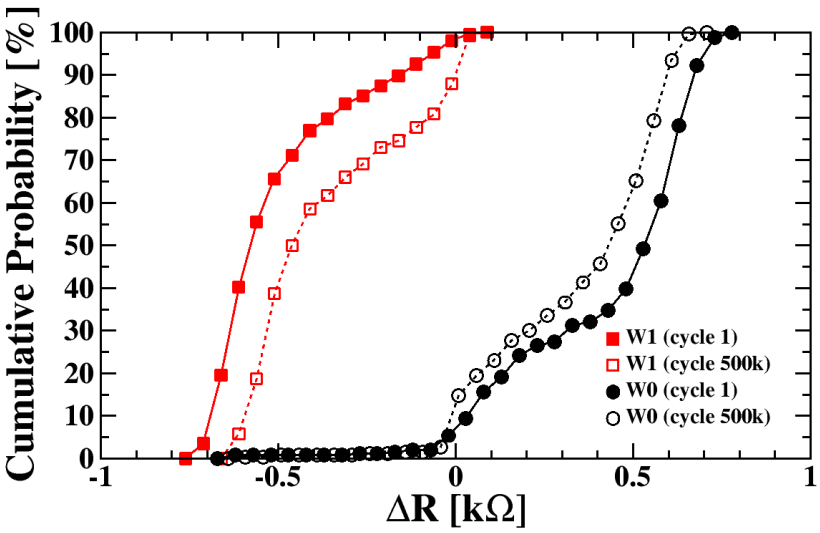
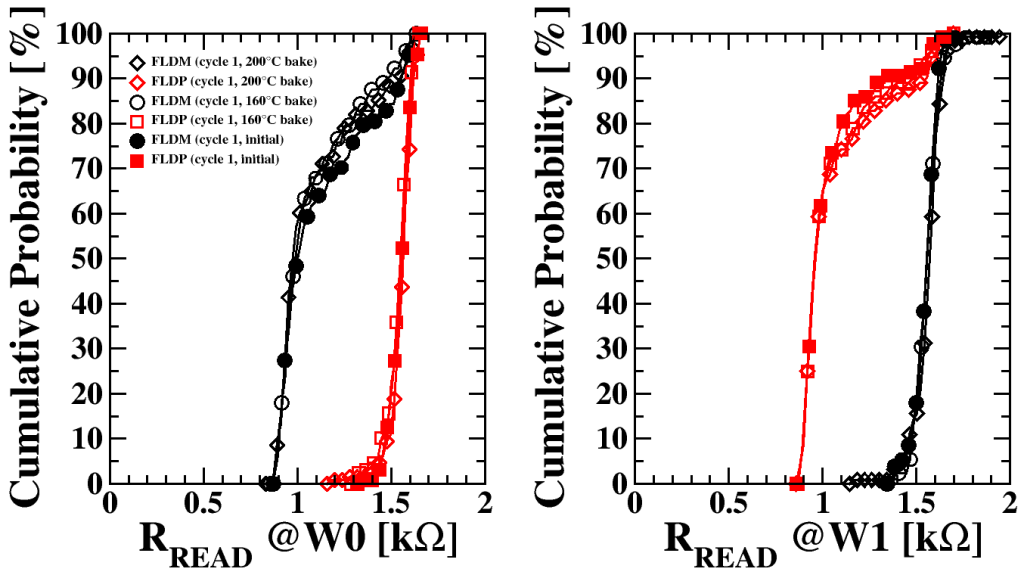


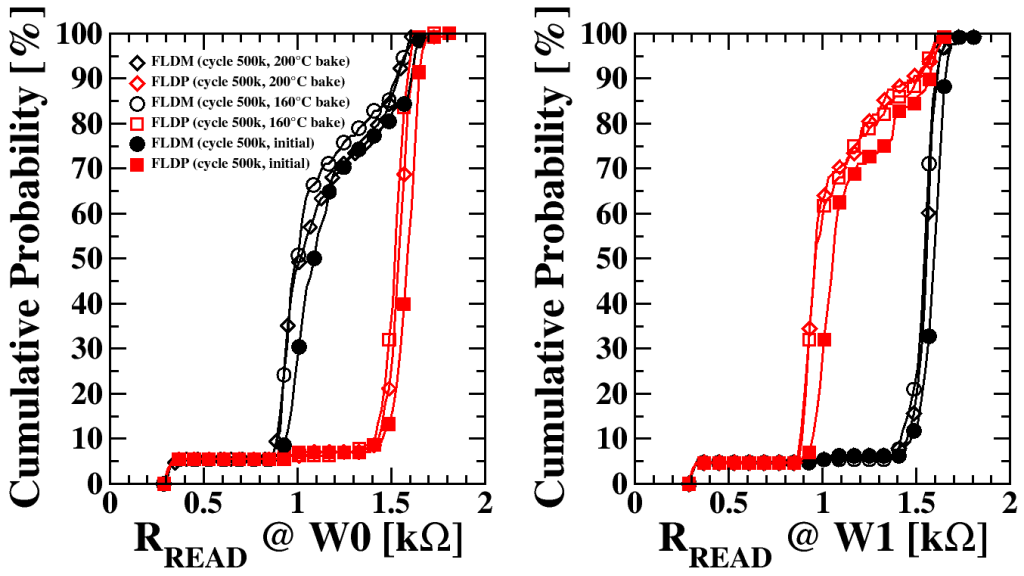
Figure 4.17: Cumulative distributions of the differential read resistances measured after Write "0" and Write "1" at cycle 1 and after 500k endurance cycles.

4.4.3 Retention

Measurements have been performed to evaluate the data retention degradation by baking ceramic packaged test chips with both fresh and cycled devices at 160°C for 150 hours and then at 200°C for 100 hours. Cumulative distributions of Read Resistances measured after Write "0" and Write "1", before and after the retention tests on fresh and cycled devices are reported in Fig. 4.18. The temperature tests show no relevant impact on the measured distributions. The average values and standard deviations of the differential resistances calculated after Write "0" and Write "1" during the retention tests for fresh and cycled devices are reported in Fig. 4.19 (a,b). A larger read window (i.e. the distance between the measured differential resistances after write operations) and lower standard deviation is observed on fresh devices, however the retention tests do not cause any relevant variation on the read windows. The average differential resistances ΔR calculated during the retentions tests on fresh and cycled devices are shown in Fig. 4.19 (c). Again, no relevant impact of the retention tests is observed.

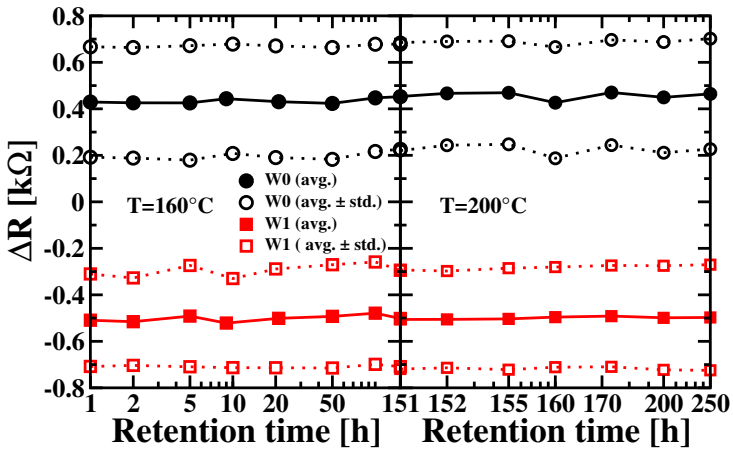


(a)

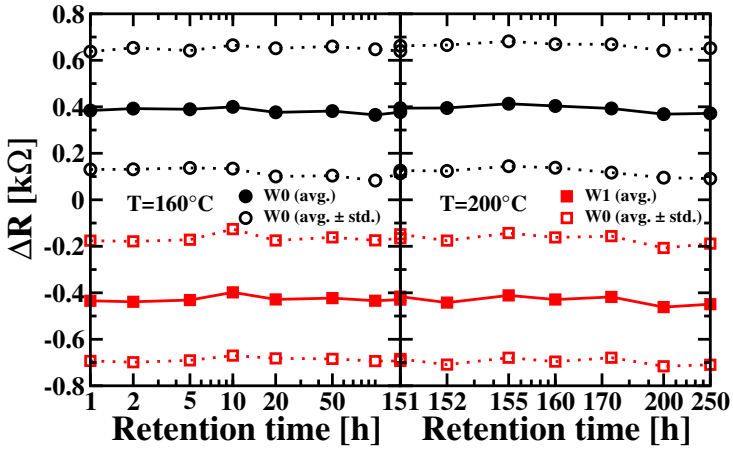


(b)

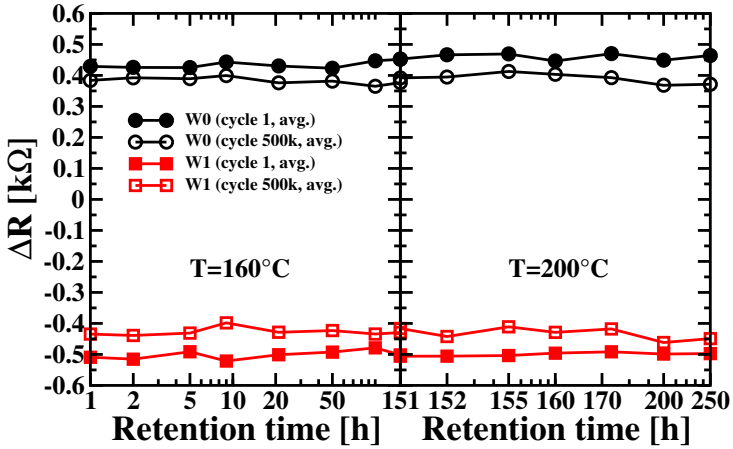
Figure 4.18: Cumulative distributions of Read Resistances measured after Write "0" and Write "1" before and after the retention tests on fresh (a) and cycled (b) devices.



(a)



(b)



(c)

Figure 4.19: Average value and standard deviations of the differential read resistances measured during 150 hours of retention at 160°C and 100 hours at 200°C, on fresh (a) and cycled (b) devices and comparison of average differential read resistances during the retention test (c).

Chapter 5

Conclusions

In this thesis the reliability of three promising candidates for the Flash memory technology replacement was investigated. Looking into more details at the aforementioned nonvolatile memory technologies, due to the very different nature of the physics behind them different advantages, disadvantages and reliability issues were observed compared to Flash. Even if all of them were able to overcome Flash for several performance and reliability features, not even one of them could today replace Flash in storage applications or fit all the requirements of the memory hierarchy to became the universal memory technology of the future. However, their characteristics make them suitable for satisfying different market requirements. A comparison of the main features in terms of performance and reliability between standard Flash technology and the emerging technologies considered in this thesis is reported in Tab. 5.1.

The experimental results evidenced that the main reliability issue for Charge Trap NAND technology is the retention. The use of enhanced program and read algorithms allows to reduce the charge loss, increasing the overall memory reliability. The SSD simulations evidenced that the reliability increase obtained with such algorithms allows to satisfy the Quality of Service requirements of enterprise SSD for a longer number of endurance cycles compared to the standard programming algorithm. The endurance

Table 5.1: Performance and reliability comparison.

Metric	Flash NAND	CT-NAND	RRAM	TAS-MRAM
Write/Read Speed	Slow (≈ms)	Slow (≈ms)	Fast (≈100 ns)	Fast (≈100 ns)
Endurance	Low ($\approx 10^4$)	Low ($\approx 10^5$)	Medium ($\approx 10^6$)	High ($>10^6$)
Retention	Medium	Low	Medium	High
Integration Density	High (≈Tbit)	High (>Tbit)	Medium (≈Mbit)	Low (<Mbit)
Reliability Limiting Factor	Scalability	Retention	Intrinsic variability	Process
Expected Applications	SSD, Storage	SSD, Storage	Embedded, Space, Neuromorphic, Wearable	Embedded, Automotive

gain in different SSD architectures for enterprise environments is quantified by a factor four. As a consequence, the advantages in terms of SSD's Quality of Service obtained by using enhanced algorithms are demonstrated to be outstanding for cold storage scenarios in which the read operations are much more than the write operations, hence the lower programming speed of the enhanced program algorithms will not impact the performances. In conclusion, the high scalability of Charge Trap technology still makes it the most attractive solution for the 3D integration in hyper-scaled arrays since its reliability problems can be mitigated through programming algorithms and error correction techniques. As future work, could be interesting to evaluate the impact of the proposed algorithms on 3D CT arrays as well as the expected SSD performance and reliability.

RRAM technology results showed that its main reliability issue is the

variability from cell-to-cell and cycle-to-cycle. It has been shown that in HRS the variability limiting element is the MIM stack, while in LRS is the selector giving a guideline for variability reduction in both conduction regimes. As future work could be interesting to study how these intrinsic variability limits behave when enhanced program algorithms are used. Several solutions were proposed to reduce and control the variability such as process optimization and program and verify algorithms usage: both solutions showed promising results. However, such variability issues are expected to increase on arrays larger than the 4kbits considered in this thesis. Hence, RRAMs seems to be promising only for applications were small array sizes are requested. In embedded and wearable applications RRAM speed and low power features could be interesting advantages. Space applications is another appealing market because of RRAM radiation hard intrinsic features. Finally, RRAMs are gathering a lot of interest for neuromorphic applications, in which a high variability is requested to emulate the physics of a neural network in an effective way.

TAS-MRAM characterization results evidenced a low cell-to-cell and cycle-to-cycle variability. In this technology the reliability is process-dependent only, hence as soon as the process will be fully optimized this is not expected to be a limit anymore. Self Referenced TAS-MRAM showed increased reliability and process variability immunity even at high temperatures up to 500k endurance cycles, at the cost of a longer read. However, the high power requirements (around mA per bit in program) and the process complexity still prevent this technology from its commercialization. To this extent, process and stack optimizations will be necessary to reduce the power requirements without impacting the reliability. TAS-MRAM could be an interesting technology for automotive/embedded applications, requiring small array sizes and very high reliability features.

Bibliography

[1] A. Grossi, C. Zambelli, and P. Olivo, *Reliability of 3D NAND Flash Memories*, R. Micheloni, Ed. Dordrecht: Springer Netherlands, 2016.

[2] L. Perniola, G. Molas, G. Navarro, E. Nowak, V. Sousa, E. Vianello, and B. D. Salvo, "Universal Signatures from Non-Universal Memories: Clues for the Future..." in *IEEE 8th International Memory Workshop (IMW)*, May 2016, pp. 1–3.

[3] L. Zuolo, C. Zambelli, R. Micheloni, S. Galfano, M. Indaco, S. Di Carlo, P. Prinetto, P. Olivo, and D. Bertozzi, "SSDExplorer: A virtual platform for fine-grained design space exploration of solid state drives," in *Design, Automation Test in Europe Conference Exhibition (DATE)*, Mar 2014, pp. 1–6.

[4] L. Zuolo, C. Zambelli, R. Micheloni, D. Bertozzi, and P. Olivo, "Analysis of reliability/performance trade-off in solid state drives," in *IEEE International Reliability Physics Symposium (IRPS)*, Jun 2014, pp. 4B.3.1–4B.3.5.

[5] L. Zuolo, C. Zambelli, R. Micheloni, S. Galfano, M. Indaco, S. Di Carlo, P. Prinetto, P. Olivo, and D. Bertozzi, "SSDExplorer: a Virtual Platform for Performance/Reliability-oriented Fine-Grained Design Space Exploration of Solid State Drives," *to appear on IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, 2015.

[6] L. Zuolo, C. Zambelli, A. Grossi, R. Micheloni, S. Bates, and P. Olivo, "Memory System Architecture Optimization for Enterprise All-RRAM Solid State Drives," in *IEEE 8th International Memory Workshop (IMW)*, May 2016, pp. 1–4.

[7] A. Grossi, L. Zuolo, F. Restuccia, C. Zambelli, and P. Olivo, "Quality-of-service implications of enhanced program algorithms for charge-trapping nand in future solid-state drives," *IEEE Transactions on Device and Materials Reliability*, vol. 15, no. 3, pp. 363–369, Sept 2015.

[8] A. Arreghini, F. Driussi, D. Esseni, L. Selmi, M. J. van Duuren, and R. van Schaijk, "Experimental Extraction of the Charge Centroid and of the Charge Type in the P/E Operation of Sonos Memory Cells," in *IEEE International Electron Devices Meeting*, Dec 2006, pp. 1–4.

[9] L. Vandelli, A. Padovani, L. Larcher, A. Arreghini, G. V. den bosch, M. Jurczak, J. V. Houdt, V. D. Marca, and P. Pavan, "Role of holes and electrons during erase of TANOS memories: Evidences for dipole formation and its impact on reliability," in *IEEE International Reliability Physics Symposium*, May 2010, pp. 731–737.

[10] A. Arreghini, N. Akil, F. Driussi, D. Esseni, L. Selmi, and M. J. van Duuren, "Characterization and modeling of long term retention in SONOS non volatile memories," in *European Solid State Device Research Conference*, Sept 2007, pp. 406–409.

[11] C.-P. Chen, H.-T. Lue, C.-C. Hsieh, K.-P. Chang, C.-C. Hsieh, and C.-Y. Lu, "Study of fast initial charge loss and it's impact on the programmed states V_T distribution of charge-trapping NAND Flash," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2010, pp. 5.6.1–5.6.4.

[12] J. K. Park, D. I. Moon, Y. K. Choi, S. H. Lee, K. H. Lee, S. H. Pyi, and B. J. Cho, "Origin of transient Vth shift after erase and its impact on

2D/3D structure charge trap flash memory cell operations," in *IEEE International Electron Devices Meeting*, Dec 2012, pp. 2.4.1–2.4.4.

[13] H. Park, G. Bersuker, D. Gilmer, K. Lim, M. Jo, H. Hwang, A. Padovani, L. Larcher, P. Pavan, W. Taylor, and P. Kirsch, "Charge loss in tanos devices caused by vt sensing measurements during retention," in *IEEE International Memory Workshop (IMW)*, May 2010, pp. 1–2.

[14] W. Tsai, N. Zous, C. Liu, C. Liu, C. Chen, T. Wang, S. Pan, C.-Y. Lu, and S. Gu, "Data retention behavior of a sonos type two-bit storage flash memory cell," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2001, pp. 32.6.1–32.6.4.

[15] N. Mielke, T. Marquart, W. Ning, J. Kessenich, H. Belgal, E. Schares, F. Trivedi, E. Goodness, and L. Nevill, "Bit error rate in NAND flash memories," in *IEEE International Reliability Physics Symposium (IRPS)*, Apr 2008, pp. 9–19.

[16] G. Ghidini, C. Scozzari, N. Galbiati, A. Modelli, E. Camerlenghi, M. Alessandri, A. D. Vitto, G. Albini, A. Grossi, T. Ghilardi, and P. Tessariol, "Cycling degradation in TANOS stack," *Microelectronics Engineering*, vol. 86, pp. 1822–1825, 2009.

[17] H. Park, G. Bersuker, M. Jo, D. Veksler, K. Lim, D. Gilmer, N. Goel, C. Kang, C. Young, M. Chang, H. Hwang, H. Tseng, P. Kirsch, and R. Jammy, "Tunnel oxide degradation in TANOS devices and its origin," in *International Symposium on VLSI Technology Systems and Applications (VLSI-TSA)*, Apr 2010, pp. 50–51.

[18] H. Park, G. Bersuker, D. Gilmer, K. Lim, M. Jo, H. Hwang, A. Padovani, L. Larcher, P. Pavan, W. Taylor, and P. Kirsch, "Charge loss in TANOS devices caused by V_T sensing measurements during retention," in *IEEE International Memory Workshop (IMW)*, May 2010, pp. 1–2.

[19] Intel, "Intel SSD Data Center S3700 Series Quality of Service," Jun 2013. [Online]. Available: http://www.intel.com/content/www/us/en/solid-state-drives/ssd-dc-s3700-quality-service-tech-brief.html

[20] C. C. Yeh, W. J. Tsai, T. C. Lu, H. Y. Chen, H. C. Lai, N. Zous, Y. Y. Liao, G. D. You, S. Cho, C. Liu, F. S. Hsu, L. T. Huang, W. S. Chiang, C. J. Liu, C. F. Cheng, M. H. Chou, C. H. Chen, T. Wang, W. Ting, S. Pan, J. Ku, and C.-Y. Lu, "Novel operation schemes to improve device reliability in a localized trapping storage SONOS-type flash memory," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2003, pp. 7.5.1–7.5.4.

[21] K.-D. Suh, B.-H. Suh, Y.-H. Um, J.-K. Kim, Y.-J. Choi, Y.-N. Koh, S.-S. Lee, S.-C. Kwon, B.-S. Choi, J.-S. Yum, J.-H. Choi, J.-R. Kim, and H.-K. Lim, "A 3.3 V 32 Mb NAND flash memory with incremental step pulse programming scheme," in *IEEE International Solid-State Circuits Conference (ISSCC)*, Feb 1995, pp. 128–129.

[22] A. Grossi, C. Zambelli, and P. Olivo, "Bit error rate analysis in charge trapping memories for ssd applications," in *IEEE International Reliability Physics Symposium*, June 2014, pp. MY.7.1–MY.7.5.

[23] S. Amoroso, A. Mauri, N. Galbiati, C. Scozzari, E. Mascellino, E. Camozzi, A. Rangoni, T. Ghilardi, A. Grossi, P. Tessariol, C. Monzio Compagnoni, A. Maconi, A. Lacaita, A. Spinelli, and G. Ghidini, "Reliability constraints for tanos memories due to alumina trapping and leakage," in *IEEE International Reliability Physics Symposium (IRPS)*, May 2010, pp. 966–969.

[24] L. Crippa and R. Micheloni, "MLC Storage," in *Inside NAND Flash memories*. Springer-Verlag, 2010, pp. 261–298.

[25] A. Padovani, L. Larcher, D. Heh, and G. Bersuker, "Modeling TANOS Memory Program Transients to Investigate Charge-Trapping Dynamics," *IEEE Electron Device Letters*, vol. 30, no. 8, pp. 882–884, 2009.

[26] J.-W. Im, W.-P. Jeong, D.-H. Kim, S.-W. Nam, D.-K. Shim, M.-H. Choi, H.-J. Yoon, D.-H. Kim, Y.-S. Kim, H.-W. Park, D.-H. Kwak, S.-W. Park, S.-M. Yoon, W.-G. Hahn, J.-H. Ryu, S.-W. Shim, K.-T. Kang, S.-H. Choi, J.-D. Ihm, Y.-S. Min, I.-M. Kim, D.-S. Lee, J.-H. Cho, O.-S. Kwon, J.-S. Lee, M.-S. Kim, S.-H. Joo, J.-H. Jang, S.-W. Hwang, D.-S. Byeon, H.-J. Yang, K.-T. Park, K. hyun Kyung, and J.-H. Choi, "A 128Gb 3b/cell V-NAND flash memory with 1Gb/s I/O rate," in *IEEE International Solid-State Circuits Conference (ISSCC)*, Feb 2015, pp. 1–3.

[27] A. Nitayama and H. Aochi, "Bit Cost Scalable (BiCS) flash technology for future ultra high density storage devices," in *International Symposium on VLSI Technology Systems and Applications (VLSI-TSA)*, Apr 2010, pp. 130–131.

[28] C. Zambelli and P. Olivo, "Statistical investigation of anomalous fast erase dynamics in charge trapping nand flash," *IEEE Electron Device Letters*, vol. 34, no. 4, pp. 514–516, 2013.

[29] C. Zambelli, A. Chimenton, and P. Olivo, "Analysis of Edge Wordline Disturb in multimegabit charge trapping flash NAND arrays," in *IEEE International Reliability Physics Symposium (IRPS)*, Apr 2011, pp. MY.4.1–MY.4.5.

[30] Y. Lee, H. Yoo, I. Yoo, and I.-C. Park, "6.4Gb/s multi-threaded BCH encoder and decoder for multi-channel SSD controllers," in *IEEE International Solid-State Circuits Conference (ISSCC)*, Feb 2012, pp. 426–428.

[31] R. Frickey, Intel-Corporation, "Data Integrity on 20nm SSDs," 2012.

[32] J. Yang, Silicon Motion, "High-Efficiency SSD for Reliable Data Storage Systems," 2012.

[33] PCI-SIG, "PCIe Base 2.1 Specification," 2005. [Online]. Available: https://www.pcisig.com/specifications/pciexpress/base2

[34] D. Cobb and A. Huffman, "NVM Express and the PCI Express SSD Revolution," 2012. [Online]. Available: http://www.nvmexpress.org/wp-content/uploads/2013/04/IDF-2012-NVM-Express-and-the-PCI-Express-SSD-Revolution.pdf

[35] "Flexible I/O tester." [Online]. Available: http://freecode.com/projects/fio

[36] C. Sun, A. Soga, T. Onagi, K. Johguchi, and K. Takeuchi, "A workload-aware-design of 3D-NAND flash memory for enterprise SSDs," in *International Symposium on Quality Electronic Design (ISQED)*, Mar 2014, pp. 554–561.

[37] "Standard JESD219, Solid-State Drive (SSD) Endurance Workloads," 2010.

[38] H. S. P. Wong, H. Y. Lee, S. Yu, Y. S. Chen, Y. Wu, P. S. Chen, B. Lee, F. T. Chen, and M. J. Tsai, "[metal-oxide rram," *Proceedings of the IEEE*.

[39] D. Ielmini, "Resistive switching memories based on metal oxides: mechanisms, reliability and scaling," *Semiconductor Science and Technology*, vol. 31, no. 6, p. 063002, 2016. [Online]. Available: http://stacks.iop.org/0268-1242/31/i=6/a=063002

[40] R. Waser, *Nanoelectronics and Information Technology: Advanced Electronic Materials and Novel Devices*. New York, NY, USA: John Wiley & Sons, Inc., 2003.

[41] S. F. Karg *et al.*, "Transition-metal-oxide-based resistance-change memories," *IBM Journal of Research and Development*, vol. 52, no. 4.5, pp. 481–492, 2008.

[42] Y. Y. Chen, B. Govoreanu, L. Goux, R. Degraeve, A. Fantini, G. S. Kar, D. J. Wouters, G. Groeseneken, J. A. Kittl, M. Jurczak, and L. Altimime, "Balancing SET/RESET pulse for $>10^{10}$ endurance in HfO_2/Hf 1T1R bipolar RRAM," *IEEE Trans. on Electron Devices*, vol. 59, no. 12, pp. 3243–3249, Dec 2012.

[43] C. Y. Chen, L. Goux, A. Fantini, A. Redolfi, S. Clima, R. Degraeve, Y. Y. Chen, G. Groeseneken, and M. Jurczak, "Understanding the impact of programming pulses and electrode materials on the endurance properties of scaled ta2o5 rram cells," in *2014 IEEE International Electron Devices Meeting*, Dec 2014, pp. 14.2.1–14.2.4.

[44] D. Carta, I. Salaoru, A. Khiat, A. Regoutz, C. Mitterbauer, N. M. Harrison, and T. Prodromakis, "Investigation of the Switching Mechanism in TiO2-Based RRAM: A Two-Dimensional EDX Approach," *ACS Applied Materials and Interfaces*, vol. 8, no. 30, pp. 19 605–19 611, 2016.

[45] M. Fujimoto, H. Koyama, M. Konagai, Y. Hosoi, K. Ishihara, S. Ohnishi, and N. Awaya, "TiO2 anatase nanolayer on TiN thin film exhibiting high-speed bipolar resistive switching," *Applied Physics Letters*, vol. 89, no. 22, p. 223509, 2006.

[46] S. Yu, X. Guan, and H.-S. P. Wong, "Conduction mechanism of TiN/HfOx/Pt resistive switching memory: A trap-assisted-tunneling model," *Applied Physics Letters*, vol. 99, no. 6, p. 063507, 2011.

[47] G. Bersuker, D. C. Gilmer, D. Veksler, P. Kirsch, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, A. Shluger, V. Iglesias, M. Porti, and M. NafrÃa, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *Journal of Applied Physics*, vol. 110, no. 12, p. 124518, 2011.

[48] A. Grossi, C. Zambelli, P. Olivo, P. Pellati, M. Ramponi, C. Wenger, J. Alvarez-Herault, and K. Mackay, "An Automated Test Equipment for Characterization of emerging MRAM and RRAM arrays," *IEEE Transactions on Emerging Topics in Computing*, vol. PP, no. 99, pp. 1–1, 2016.

[49] C. Zambelli, A. Grossi, P. Olivo, D. Walczyk, T. Bertaud, B. Tillack, T. Schroeder, V. Stikanov, and C. Walczyk, "Statistical analysis of resistive switching characteristics in ReRAM test arrays," in *IEEE International Conference on Microelectronic Test Structures (ICMTS)*, March 2014, pp. 27–31.

[50] C. Zambelli, A. Grossi, P. Olivo, D. Walczyk, J. Dabrowski, B. Tillack, T. Schroeder, R. Kraemer, V. Stikanov, and C. Walczyk, "Electrical characterization of read window in reram arrays under different SET/RESET cycling conditions," in *IEEE International Memory Workshop (IMW)*, May 2014, pp. 1–4.

[51] D. Walczyk, T. Bertaud, M. Sowinska, M. Lukosius, M. A. Schubert, A. Fox, D. Wolansky, A. Scheit, M. Fraschke, G. Schoof, C. Wolf, R. Kraemer, B. Tillack, R. Korolevych, V. Stikanov, C. Wenger, T. Schroeder, and C. Walczyk, "Resistive switching behavior in TiN/HfO2/Ti/TiN devices," in *International Semiconductor Conference Dresden-Grenoble (ISCDG)*, Sept 2012, pp. 143–146.

[52] E. A. Miranda, C. Walczyk, C. Wenger, and T. Schroeder, "Model for the Resistive Switching Effect in HfO_2 MIM Structures Based on the Transmission Properties of Narrow Constrictions," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 609–611, 2010.

[53] C. Walczyk, D. Walczyk, T. Schroeder, T. Bertaud, M. Sowinska, M. Lukosius, M. Fraschke, D. Wolansky, B. Tillack, E. Miranda, and C. Wenger, "Impact of Temperature on the Resistive Switching Behav-

ior of Embedded HfO_2-Based RRAM Devices," *IEEE Transactions on Electron Devices*, vol. 58, no. 9, pp. 3124–3131, 2011.

[54] C. Zambelli, A. Grossi, P. Olivo, C. Walczyk, and C. Wenger, "RRAM Reliability/Performance Characterization through Array Architectures Investigations," in *IEEE Computer Society Annual Symposium on VLSI*, July 2015, pp. 327–332.

[55] N. Raghavan, A. Fantini, R. Degraeve, P. Roussel, L. Goux, B. Govoreanu, D. Wouters, G. Groeseneken, and M. Jurczak, "Statistical insight into controlled forming and forming free stacks for hfox {RRAM}," *Microelectronic Engineering*, vol. 109, pp. 177 – 181, 2013.

[56] G. Jegert, A. Kersch, W. Weinreich, and P. Lugli, "Ultimate scaling of TiN/ZrO2/TiN capacitors: Leakage currents and limitations due to electrode roughness," *Journal of Applied Physics*, vol. 109, no. 1, p. 014504, 2011.

[57] P.-S. Chen, Y.-S. Chen, H.-Y. Lee, T.-Y. Wu, P.-Y. Gu, F. Chen, and M.-J. Tsai, "Impact of Flattened TiN Electrode on the Memory Performance of HfO2 Based Resistive Memory," *Electrochemical and Solid-State Letters*, vol. 15, no. 4, pp. H136–H139, 2012.

[58] S. Mueller, J. Mueller, A. Singh, S. Riedel, J. Sundqvist, U. Schroeder, and T. Mikolajick, "Incipient Ferroelectricity in Al-Doped HfO2 Thin Films," *Advanced Functional Materials*, vol. 22, no. 11, pp. 2412–2417, 2012.

[59] N. Raghavan, R. Degraeve, A. Fantini, L. Goux, D. J. Wouters, G. Groeseneken, and M. Jurczak, "Stochastic variability of vacancy filament configuration in ultra-thin dielectric RRAM and its impact on OFF-state reliability," in *IEEE International Electron Devices Meeting*, Dec 2013, pp. 21.1.1–21.1.4.

[60] L. Larcher, O. Pirrotta, F. M. Puglisi, A. Padovani, P. Pavan, and L. Vandelli, "Progresses in Modeling HfOx RRAM Operations and Variability," *ECS Transactions*, vol. 64, no. 14, pp. 49–60, 2014.

[61] T. J. Chu, T. C. Chang, T. M. Tsai, H. H. Wu, J. H. Chen, K. C. Chang, T. F. Young, K. H. Chen, Y. E. Syu, G. W. Chang, Y. F. Chang, M. C. Chen, J. H. Lou, J. H. Pan, J. Y. Chen, Y. H. Tai, C. Ye, H. Wang, and S. M. Sze, "Charge quantity influence on resistance switching characteristic during forming process," *IEEE Electron Device Letters*, vol. 34, no. 4, pp. 502–504, 2013.

[62] C. F. T., L. HengYuan, C. YuSheng, H. YenYa, Z. LiJie, C. Pang-Shiu, C. WeiSu, G. PeiYi, L. WenHsing, W. SuMin, T. ChenHan, S. S. Shyuan, T. M. Jinn, and H. Ru, "Resistance switching for rram applications," *SCIENCE CHINA Information Sciences*, vol. 54, no. 5, pp. 1073–1086, 2011.

[63] M. Sowinska, T. Bertaud, D. Walczyk, S. Thiess, P. Calka, L. Alff, C. Walczyk, and T. Schroeder, "In-operando hard X-ray photoelectron spectroscopy study on the impact of current compliance and switching cycles on oxygen and carbon defects in resistive switching Ti/HfO2/TiN cells," *Journal of Applied Physics*, vol. 115, no. 20, p. 204509, 2014.

[64] H.-C. Shih, C.-Y. Chen, C.-W. Wu, C. H. Lin, and S. S. Sheu, "Training-based forming process for RRAM yield improvement," in *VLSI Test Symposium*, May 2011, pp. 146–151.

[65] P. Huang, B. Chen, Y. J. Wang, F. F. Zhang, L. Shen, R. Liu, L. Zeng, G. Du, X. Zhang, B. Gao, J. F. Kang, X. Y. Liu, X. P. Wang, B. B. Weng, Y. Z. Tang, G. Q. Lo, and D. L. Kwong, "Analytic model of endurance degradation and its practical applications for operation scheme optimization in metal oxide based rram," in *2013 IEEE International Electron Devices Meeting*, Dec 2013, pp. 22.5.1–22.5.4.

[66] K. Higuchi, T. Iwasaki, and K. Takeuchi, "Investigation of Verify-Programming Methods to Achieve 10 Million Cycles for 50nm HfO2 ReRAM," in *IEEE International Memory Workshop*, May 2012, pp. 1–4.

[67] E. Miranda, D. Jimenez, and J. Sune, "The quantum point-contact memristor," *IEEE Electron Device Letters*, vol. 33, no. 10, pp. 1474–1476, 2012.

[68] A. Chen and M.-R. Lin, "Variability of resistive switching memories and its impact on crossbar array performance," in *IEEE International Reliability Physics Symposium (IRPS)*, April 2011, pp. MY.7.1–MY.7.4.

[69] T. Ninomiya, Z. Wei, S. Muraoka, R. Yasuhara, K. Katayama, and T. Takagi, "Conductive filament scaling of TaO_x bipolar ReRAM for improving data retention under low operation current," *IEEE Transactions on Electron Devices*, vol. 60, no. 4, pp. 1384–1389, 2013.

[70] A. Grossi, D. Walczyk, C. Zambelli, E. Miranda, P. Olivo, V. Stikanov, A. Feriani, J. Sune, G. Schoof, R. Kraemer, B. Tillack, A. Fox, T. Schroeder, C. Wenger, and C. Walczyk, "Impact of Intercell and Intracell Variability on Forming and Switching Parameters in RRAM Arrays," *IEEE Transactions on Electron Devices*, vol. 62, no. 8, pp. 2502–2509, Aug 2015.

[71] F. Crupi, F. Filice, A. Grossi, C. Zambelli, P. Olivo, E. Perez, and C. Wenger, "Implications of the Incremental Pulse and Verify Algorithm on the Forming and Switching Distributions in RERAM Arrays," *IEEE Transactions on Device and Materials Reliability*, vol. 16, no. 3, pp. 413–418, 2016.

[72] E. Perez, A. Grossi, C. Zambelli, P. Olivo, and C. Wenger, "Impact of the Incremental Programming Algorithm on the Filament Conduction

in HfO2-Based RRAM Arrays," *IEEE Journal of the Electron Devices Society*, vol. 5, no. 1, pp. 64–68, 2017.

[73] P. Lorenzi, R. Rao, and F. Irrera, "Forming kinetics in HfO_2-based RRAM cells," *IEEE Transactions on Electron Devices*, vol. 60, no. 1, pp. 438–443, 2013.

[74] H.-T. Liu, B.-H. Yang, H.-B. Lv, X.-X. Xu, Q. Luo, G.-M. Wang, M.-Y. Zhang, S.-B. Long, Q. Liu, and M. Liu, "Effect of pulse and dc formation on the performance of one-transistor and one-resistor resistance random access memory devices," *Chinese Physics Letters*, vol. 32, no. 02, pp. 1–3, 2015.

[75] H.-C. Shih, C.-Y. Chen, C.-W. Wu, C.-H. Lin, and S.-S. Sheu, "Training-based forming process for RRAM yield improvement," in *IEEE VLSI Test Symposium (VTS)*, May 2011, pp. 146–151.

[76] D. C. Gilmer, G. Bersuker, H.-Y. Park, C. Park, B. Butcher, W. Wang, P. D. Kirsch, and R. Jammy, "Effects of RRAM stack configuration on forming voltage and current overshoot," in *IEEE International Memory Workshop (IMW)*, May 2011, pp. 1–4.

[77] A. Grossi, C. Zambelli, P. Olivo, E. Miranda, V. Stikanov, T. Schroeder, C. Walczyk, and C. Wenger, "Relationship among Current Fluctuations during Forming, Cell-To-Cell Variability and Reliability in RRAM Arrays," in *IEEE International Memory Workshop (IMW)*, May 2015, pp. 1–4.

[78] F. T. Chen *et al.*, "Operation of oxygen vacancy-based RRAMs," *Science China Information Sciences*, vol. 54, no. 5, pp. 1073–1086, 2011.

[79] E. Perez, C. Wenger, A. Grossi, C. Zambelli, P. Olivo, and R. Roelofs, "Impact of temperature on conduction mechanisms and switching parameters in HfO2-based 1T-1R resistive random access memories devices," *Journal of Vacuum Science and Technology B, Nanotechnology*

and Microelectronics: Materials, Processing, Measurement, and Phenomena, vol. 35, no. 1, p. 01A103, 2017.

[80] Y. Y. Chen, R. Degraeve, S. Clima, B. Govoreanu, L. Goux, A. Fantini, G. Kar, G. Pourtois, G. Groeseneken, D. Wouters, and M. Jurczak, "Understanding of the endurance failure in scaled HfO_2-based 1T1R RRAM through vacancy mobility degradation," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2012, pp. 20.3.1–20.3.4.

[81] G. Bersuker, D. C. Gilmer, D. Veksler, P. Kirsch, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, A. Shluger, V. Iglesias, M. Porti, and M. NafrÃa, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *Journal of Applied Physics*, vol. 110, no. 12, pp. 1–12, 2011.

[82] X. Lian, X. Cartoixa, E. Miranda, L. Perniola, R. Rurali, S. Long, M. Liu, and J. Sune, "Multi-scale quantum point contact model for filamentary conduction in resistive random access memories devices," *Journal of Applied Physics*, vol. 115, no. 24, pp. 1–8, 2014.

[83] S. Yu, X. Guan, and H.-S. Wong, "On the switching parameter variation of metal oxide RRAM; part ii: Model corroboration and device design strategy," *IEEE Transactions on Electron Devices*, vol. 59, no. 4, pp. 1183–1188, 2012.

[84] P. Huang, B. Chen, Y. Wang, F. Zhang, L. Shen, R. Liu, L. Zeng, G. Du, X. Zhang, B. Gao, J. Kang, X. Liu, X. Wang, B. Weng, Y. Tang, G.-Q. Lo, and D.-L. Kwong, "Analytic model of endurance degradation and its practical applications for operation scheme optimization in metal oxide based RRAM," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2013, pp. 22.5.1–22.5.4.

[85] S.-S. Sheu, M.-F. Chang, K.-F. Lin, C.-W. Wu, Y.-S. Chen, P.-F. Chiu, C.-C. Kuo, Y.-S. Yang, P.-C. Chiang, W.-P. Lin, C.-H. Lin, H.-Y. Lee,

P.-Y. Gu, S.-M. Wang, F.-T. Chen, K.-L. Su, C.-H. Lien, K.-H. Cheng, H.-T. Wu, T.-K. Ku, M.-J. Kao, and M.-J. Tsai, "A 4Mb embedded SLC resistive-RAM macro with 7.2ns read-write random-access time and 160ns MLC-access capability," in *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, Feb 2011, pp. 200–202.

[86] T.-Y. Liu, T. H. Yan, R. Scheuerlein, Y. Chen, J. K. Lee, G. Balakrishnan, G. Yee, H. Zhang, A. Yap, J. Ouyang, T. Sasaki, S. Addepalli, A. Al-Shamma, C.-Y. Chen, M. Gupta, G. Hilton, S. Joshi, A. Kathuria, V. Lai, D. Masiwal, M. Matsumoto, A. Nigam, A. Pai, J. Pakhale, C. H. Siau, X. Wu, R. Yin, L. Peng, J. Y. Kang, S. Huynh, H. Wang, N. Nagel, Y. Tanaka, M. Higashitani, T. Minvielle, C. Gorla, T. Tsukamoto, T. Yamaguchi, M. Okajima, T. Okamura, S. Takase, T. Hara, H. Inoue, L. Fasoli, M. Mofidi, R. Shrivastava, and K. Quader, "A 130.7 mm^2 2-layer 32 Gb ReRAM Memory Device in 24 nm technology," in *Proc. ISSCC*, Feb. 2013, pp. 210–212.

[87] S. Balatti, S. Ambrogio, A. Cubeta, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Voltage-dependent random telegraph noise (RTN) in HfOx resistive RAM," in *IEEE Int. Reliability Physics Symposium (IRPS)*, June 2014, pp. MY.4.1–MY.4.6.

[88] K. Morgan, R. Huang, S. Pearce, and C. de Groot, "The effect of atomic layer deposition temperature on switching properties of HfOx resistive RAM devices," in *IEEE Int. Symp. on Circuits and Systems (ISCAS)*, June 2014, pp. 432–435.

[89] H.-D. Kim, F. Crupi, M. Lukosius, A. Trusch, C. Walczyk, and C. Wenger, "Resistive switching characteristics of integrated polycrystalline hafnium oxide based one transistor and one resistor devices fabricated by atomic vapor deposition methods," *Journal of Vacuum Science and Technology B*, vol. 33, no. 5, pp. 052 204.1–052 204.5, 2015.

[90] A. Grossi, E. Perez, C. Zambelli, P. Olivo, and C. Wengeru, "Performance and Reliability Comparison of 1T-1R RRAM arrays with Amorphous and Polycrystalline HfO_2," in *Joint Int. EUROSOI Workshop and Int. Conf. on Ultimate Integration on Silicon (EUROSOI-ULIS)*, Jan 2016, pp. 80–83.

[91] A. Grossi, C. Zambelli, P. Olivo, A. Crespo-Yepes, J. Martin-Martinez, R. Rodriguez, M. Nafria, E. Perez, and C. Wenger, "Electrical characterization and modeling of 1T-1R RRAM arrays with amorphous and poly-crystalline HfO2," *Solid-State Electronics*, vol. 128, pp. 187 – 193, 2017.

[92] V. Iglesias, M. Porti, M. Nafria, X. Aymerich, P. Dudek, and G. Bersuker, "Dielectric breakdown in polycrystalline hafnium oxide gate dielectrics investigated by conductive atomic force microscopy," *Journal of Vacuum Science and Technology B*, vol. 29, no. 1, pp. 1–4, 2011.

[93] A. Chimenton, C. Zambelli, P. Olivo, and A. Pirovano, "Set of electrical characteristic parameters suitable for reliability analysis of multimegabit phase change memory arrays," in *Joint Non-Volatile Semiconductor Memory Workshop and Int. Conf. on Memory Technology and Design*, May 2008, pp. 49–51.

[94] A. Grossi, C. Zambelli, P. Olivo, E. Miranda, V. Stikanov, C. Walczyk, and C. Wenger, "Electrical characterization and modeling of pulse-based forming techniques in RRAM arrays," *Solid-State Electronics*, vol. 115, Part A, pp. 17 – 25, 2016.

[95] A. Crespo-Yepes, J. Martin-Martinez, I. Rama, M. Maestro, R. Rodriguez, M. Nafria, X. Aymerich., M. B. Gonzalez, and F. Campabadal, "Intra-device statistical parameters in variabilit-yaware modelling of resistive switching devices," in *Joint Int. EUROSOI Workshop and Int.*

Conf. on Ultimate Integration on Silicon (EUROSOI-ULIS), Jan 2016, pp. 84–87.

[96] J. Martin-Martinez, B. Kaczer, R. Degraeve, P. J. Roussel, R. Rodriguez, M. Nafria, X. Aymerich, B. Dierickx, and G. Groeseneken, "Circuit design-oriented stochastic piecewise modeling of the post-breakdown gate current in MOSFETs: Application to ring oscillators," *IEEE Trans. on Device and Materials Reliability*, vol. 12, no. 1, pp. 78–85, March 2012.

[97] Y. Y. Chen, M. Komura, R. Degraeve, B. Govoreanu, L. Goux, A. Fantini, N. Raghavan, S. Clima, L. Zhang, A. Belmonte, A. Redolfi, G. S. Kar, G. Groeseneken, D. J. Wouters, and M. Jurczak, "Improvement of data retention in HfO2/Hf 1T1R RRAM cell under low operating current," in *IEEE Int. Electron Devices Meeting (IEDM)*, Dec 2013, pp. 10.1.1–10.1.4.

[98] M. Lukosius, C. Walczyk, M. Fraschke, D. Wolansky, H. Richter, and C. Wenger, "High performance metal-insulator-metal capacitors with atomic vapor deposited HfO2 dielectrics," *Thin Solid Films*, vol. 518, no. 15, pp. 4380 – 4384, 2010.

[99] M. Cho, J. H. Kim, C. S. Hwang, H.-S. Ahn, S. Han, and J. Y. Won, "Effects of carbon residue in atomic layer deposited HfO2 films on their time-dependent dielectric breakdown reliability," *Applied Physics Letters*, vol. 90, no. 18, p. 182907, 2007.

[100] J. Robertson, "High dielectric constant gate oxides for metal oxide si transistors," *Reports on Progress in Physics*, vol. 69, no. 2, p. 327, 2006.

[101] C. Walczyk, C. Wenger, D. Walczyk, M. Lukosius, I. Costina, M. Fraschke, J. Dabrowski, A. Fox, D. Wolansky, S. Thiess, E. Miranda, B. Tillack, and T. Schroeder, "On the role of Ti adlayers for resistive switching in HfO2-based metal-insulator-metal structures: Top

versus bottom electrode integration," *Journal of Vacuum Science and Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 29, no. 1, p. 01AD02, 2011.

[102] M. Lanza, G. Bersuker, M. Porti, E. Miranda, M. Nafria, and X. Aymerich, "Resistive switching in hafnium dioxide layers: Local phenomenon at grain boundaries," *Applied Physics Letters*, vol. 101, no. 19, p. 193502, 2012.

[103] J. W. Maes, Y. Fedorenko, A. Delabie, L.-A. Ragnarsson, J. Swerts, L. Nyns, S. Van Elshocht, C. Wang, and G. Wilk, "Impact of Hf-Precursor Choice on Scaling and Performance of High-k Gate Dielectrics," *ECS Trans.*, vol. 11, no. 4, pp. 59–72, 2007.

[104] E. Miranda and J. Sune, "Electron transport through broken down ultra-thin SiO2 layers in MOS devices," *Microelectronics Reliability*, vol. 44, no. 1, pp. 1 – 23, 2004.

[105] A.Grossi, E. Nowak, C. Zambelli, C. Pellissier, S. Bernasconi, G. Cibrario, K. E. Hajjam, R. Crochemore, J. Nodin, P. Olivo, and L. Perniola, "Fundamental Variability Limits of Filament-based RRAM," in *IEEE Int. Electron Devices Meeting (IEDM)*, Dec 2016, pp. 4.7.1–4.7.4.

[106] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Statistical Fluctuations in HfOx Resistive-Switching Memory: Part I - Set/Reset Variability," *IEEE Trans. on Electron Devices*, vol. 61, no. 8, pp. 2912–2919, 2014.

[107] D. Chen, H. Kim, A. Phan, E. Wilcox, K. Label, S. Buchner, A. Khachatrian, and N. Roche, "Single-Event Effect Performance of a Commercial Embedded ReRAM," *IEEE Trans. on Nuclear Science*, vol. 61, no. 6, pp. 3088–3094, Dec 2014.

[108] S. Weeden-Wright, W. Bennett, N. Hooten, E. X. Zhang, M. McCurdy, M. King, R. Weller, M. Mendenhall, M. Alles, D. Linten, M. Jurczak, R. Degraeve, A. Fantini, R. Reed, D. Fleetwood, and R. Schrimpf, "TID and Displacement Damage Resilience of 1T1R HfO_2/Hf Resistive Memories," *IEEE Trans. on Nuclear Science*, vol. 61, no. 6, pp. 2972–2978, Dec 2014.

[109] F. Xue, L. Ping, L. Wei, Z. Bin, X. Xiaodong, W. Gang, H. Bin, and Z. Yahong, "Gate-enclosed NMOS transistors," *Journal of Semiconductors*, vol. 32, no. 8, p. 084002, 2011.

[110] A. Grossi, C. Calligaro, E. Perez, J. Schmidt, F. Teply, T. Mausolf, C. Zambelli, P. Olivo, and C. Wenger, "Radiation hard design of HfO2 based 1T1R cells and memory arrays," in *International Conference on Memristive Systems (MEMRISYS)*, Nov 2015, pp. 1–2.

[111] C. Calligaro, A. Arbat, Y. Roizin, and D. Nahmad, "A 15 Mrad(Si) 512Kbit Rad-Hard SRAM in a standard 0.18μm CMOS technology," in *13th European Conference on Radiation and its Effects on Components and Systems (RADECS)*, Sep 2012, pp. 1–1.

[112] C. Calligaro, J. Georgiou, F. Gomez, V. Liberali, D. Nahmad, A. Paccagnella, A. Prokofiev, and A. Virtanen, "A Rad-Hard 1 Mbit S-Flash OTP Memory for Space," in *Nuclear and Space Radiation Effects Conference (NSREC)*, Jul 2014, pp. 1–1.

[113] J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meservey, "Large Magnetoresistance at Room Temperature in Ferromagnetic Thin Film Tunnel Junctions," *Phys. Rev. Lett.*, vol. 74, pp. 3273–3276, 1995.

[114] T. Miyazaki and N. Tezuka, "Giant magnetic tunneling effect in Fe/Al2O3/Fe junction," *Journal of Magnetism and Magnetic Materials*, vol. 139, no. 3, pp. L231–L234, 1995.

[115] S. Bandiera and B. Dieny, *Thermally Assisted MRAM*. Dordrecht: Springer Netherlands, 2014, pp. 1–29.

[116] W. Gallagher, D. Abraham, S. Assefa, S. Brown, J. DeBrosse, M. Gaidis, E. Galligan, E. Gow, B. Hughes, J. Hummel, S. Kanakasabapathy, C. Kaiser, M. Lamorey, T. Maffit, K. Milkove, Y. Lu, J. Nowak, P. Rice, M. Samant, E. O'Sullivan, S. Parkin, R. Robertazzi, P. Trouilloud, D. Worledge, G. Wright, and S.-H. Yang, "Recent advances in MRAM technology," in *IEEE VLSI-TSA Int. Symp.*, April 2005, pp. 72–73.

[117] E. Vatajelu, H. Aziza, and C. Zambelli, "Nonvolatile memories: Present and future challenges," in *International Design Test Symposium (IDT)*, Dec 2014, pp. 61–66.

[118] R. Sousa, I. Prejbeanu, D. Stanescu, B. Rodmacq, O. Redon, B. Dieny, J. Wang, and P. Freitas, "Tunneling hot spots and heating in magnetic tunnel junctions," *J. of App. Phys.*, vol. 95, pp. 6783–6785, 2004.

[119] J. Azevedo, A. Virazel, A. Bosio, L. Dilillo, P. Girard, A. Todri-Sanial, J. Alvarez-Herault, and K. Mackay, "A complete resistive-open defect analysis for thermally assisted switching MRAMs," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 22, pp. 2326–2335, Nov 2014.

[120] I. Prejbeanu, S. Bandiera, J. Alvarez-Hérault, R. Sousa, B. Dieny, and J. Nozières, "Thermally assisted MRAMs: ultimate scalability and logic functionalities," *Journal of Physics D: Applied Physics*, vol. 46, no. 7, p. 074002, 2013.

[121] B. Dieny and O. Redon, "Magnetic tunnel junction magnetic device, memory and writing and reading methods using said device," Sep. 27 2005, uS Patent 6,950,335. [Online]. Available: https://www.google.ch/patents/US6950335

[122] I. Prejbeanu, W. Kula, K. Ounadjela, R. Sousa, O. Redon, B. Dieny, and J.-P. Nozieres, "Thermally assisted switching in exchange-biased storage layer magnetic tunnel junctions," *IEEE Trans. on Magnetics*, vol. 40, pp. 2625–2627, 2004.

[123] A. Grossi, C. Zambelli, P. Olivo, J. Alvarez-Herault, and K. Mackay, "Reliability and cell-to-cell variability of TAS-MRAM arrays under cycling conditions," in *Non-Volatile Memory Technology Symposium (NVMTS)*, Oct 2015, pp. 1–4.

[124] A. Grossi, C. Zambelli, P. Olivo, P. Pellati, M. Ramponi, J. Alvarez-Herault, and K. Mackay, "Automated characterization of TAS-MRAM test arrays," in *IEEE Int. Conference on Design Technology of Integrated Systems in Nanoscale Era (DTIS)*, April 2015, pp. 1–2.

[125] S. Yuasa, A. Fukushima, K. Yakushiji, T. Nozaki, M. Konoto, H. Maehara, H. Kubota, T. Taniguchi, H. Arai, H. Imamura, K. Ando, Y. Shiota, F. Bonell, Y. Suzuki, N. Shimomura, E. Kitagawa, J. Ito, S. Fujita, K. Abe, K. Nomura, H. Noguchi, and H. Yoda, "Future prospects of MRAM technologies," in *IEEE Int. Electron Devices Meeting (IEDM)*, Dec 2013, pp. 3.1.1–3.1.4.

[126] Q. Stainer, L. Lombard, K. Mackay, R. C. Sousa, I. L. Prejbeanu, and B. Dieny, "MRAM with soft reference layer: In-stack combination of memory and logic functions," in *IEEE Int. Memory Workshop*, May 2013, pp. 84–87.

Author's publications list

International Journals

1. "Impact of inter-cell and intra-cell variability on forming and switching parameters in RRAM arrays"

   A. Grossi, D. Walczyk, C. Zambelli, E. Miranda, P. Olivo, V. Stikanov, A. Feriani, J. Sune, G. Schoof, R. Kraemer, B. Tillack, A. Fox, T. Schroeder, C. Wenger, and C. Walczyk

   In: IEEE Transactions on Electron Devices (TED), vol. 62, no. 8, pp. 2502-2509, Aug 2015

2. "Quality-of-Service Implications of Enhanced Program Algorithms for Charge-Trapping NAND in Future Solid-State Drives"

   A. Grossi, L. Zuolo, F. Restuccia, C. Zambelli, and P. Olivo

   In: IEEE Transactions on Device and Materials Reliability (TDMR), vol. 15, no. 3, pp. 363-369, Sep 2015

3. "Electrical Characterization and Modeling of Pulse-based Forming Techniques in RRAM Arrays"

   A. Grossi, C. Zambelli, P. Olivo, E. Miranda, V. Stikanov, C. Walczyk, and C. Wenger

   In: Elsevier Solid-State Electronics, vol. 115, part A, pp. 17-25, Jan 2016

4. "An Automated Test Equipment for Characterization of emerging MRAM and RRAM arrays"

   A. Grossi, C. Zambelli, P. Olivo, P. Pellati, M. Ramponi, C. Wenger, J. Alvarez-Herault and K. Mackay

   In: IEEE Transactions on Emerging Topics in Computing (TETC), vol. PP, pp. 1-10, Jun 2016

5. "Implications of the Incremental Pulse and Verify Algorithm on the Forming and Switching Distributions in RERAM Arrays"

   F. Crupi, F. Filice, A. Grossi, C. Zambelli, P. Olivo, E. Perez and C. Wenger

   In: IEEE Transactions on Device and Materials Reliability (TDMR), vol. PP, pp. 1-6, Jul 2016

6. "Impact of the incremental programming algorithm on the filament conduction in HfO2 based RRAM arrays"

   E. Perez, A. Grossi, C. Zambelli, P. Olivo, and C. Wenger

   In: IEEE Journal of the Electron Devices Society (J-EDS), vol. PP, pp. 1-5, Oct 2016

7. "Impact of Temperature on Conduction Mechanisms and Switching Parameters in HfO2-based 1T-1R RRAM Devices"

   E. Perez, C. Wenger, A. Grossi, C. Zambelli, P. Olivo, and R. Roelofs

   In: AVS Journal of Vacuum Science and Technology B (JVSTB), vol.35, pp. 1-5, Jan 2017

8. "Reduction of the cell-to-cell variability in Hf1-xAlxOy based RRAM arrays by using program algorithms"

   E. Perez, A. Grossi, C. Zambelli, P. Olivo, R. Roelofs, and C. Wenger

   In: IEEE Electron Device Letters, vol.PP, pp. 1-4, Jan 2017

9. "Electrical Characterization and Modeling of 1T-1R RRAM Arrays with Amorphous and Poly-crystalline HfO2"

   A. Grossi, C. Zambelli, P. Olivo, A. Crespo-Yepes, J. Martin-Martinez, R. Rodriguez, M. Nafria, E. Perez, and C. Wenger

   In: Elsevier Solid-State Electronics, vol. 128, pp. 187-193, Feb 2017

International Conference Proceedings

10. "Statistical analysis of resistive switching characteristics in ReRAM test arrays"

    C. Zambelli, A. Grossi, D. Walckzyk, T. Bertaud, B. Tillack, T. Schroeder, V. Stikanov, and C. Walczyk

    In: IEEE International Conference on Microelectronic Test Structures (ICMTS), pp. 27-31, Mar 2014

11. "Electrical characterization of read window in ReRAM arrays under different SET/RESET cycling conditions"

    C. Zambelli, A. Grossi, P. Olivo, D. Walczyk, J. Dabrowski, B. Tillack, T. Schroeder, R. Kraemer ,V. Stikanov, and C. Walczyk

    In: IEEE International Memory Workshop (IMW), pp. 1-4, May 2014

12. "Bit Error Rate Analysis in Charge Trapping Memories for SSD applications"

    A. Grossi, C. Zambelli, and P. Olivo.

    In: IEEE International Reliability Physics Symposium (IRPS), pp. MY.7.1-MY.7.5, Jun 2014

13. "Automated characterization of TAS-MRAM test arrays"

    A. Grossi, C. Zambelli, P. Olivo, P. Pellati, M. Ramponi, J. Alvarez-Herault, and K. Mackay

In: IEEE International Conference On Design and Technology of Integrated Systems In Nanoscale Era (DTIS), pp. 1-2, Apr 2015

14. "Relationship among current fluctuations during forming, cell-to-cell variability and reliability in RRAM arrays"

    A. Grossi, C. Zambelli, P. Olivo, E. Miranda, V. Stikanov, T. Schroeder, C. Walczyk, and C. Wenger

    In: IEEE International Memory Workshop (IMW), pp. 1-4, May 2015

15. "RRAM Reliability and Performance Characterization through Array Architectures investigations"

    C. Zambelli, A. Grossi, P. Olivo, C. Walczyk, and C. Wenger

    In: IEEE Computer Society Annual Symposium on VLSI (ISVLSI), pp. 327-332, Jul 2015

16. "Reliability and Cell-to-Cell Variability of TAS-MRAM arrays under cycling conditions"

    A. Grossi, C. Zambelli, P. Olivo, J. Alvarez-Herault and K. Mackay

    In: IEEE Non-Volatile Memory Technology Symposium (NVMTS), pp. 1-4, Oct 2015

17. "Radiation hard design of HfO2 based 1T1R cells and memory arrays"

    A. Grossi, C. Calligaro, E. Perez, J. Schmidt, F. Teply, T. Mausolf, C. Zambelli, P. Olivo, and C. Wenger

    In: International Conference on Memristive Systems (MEMRYSIS), pp. 1-2, Nov 2015

18. "Performance and Reliability Comparison of 1T-1R RRAM arrays with Amorphous and Polycrystalline HfO2"

    A. Grossi, E. Perez, C. Zambelli, P. Olivo, and C. Wenger

    In: IEEE Joint Int. EUROSOI Workshop and Int. Conf. on Ultimate Integration on Silicon (EUROSOI-ULIS), pp.80-83, Jan 2016

19. "Memory System Architecture Optimization for Enterprise All-RRAM Solid State Drives"

    L. Zuolo, C. Zambelli, A. Grossi, R. Micheloni, S. Bates and P. Olivo

    In: IEEE International Memory Workshop (IMW), pp. 1-4, May 2016

20. "Fundamental Variability Limits of Filament-based RRAM"

    A.Grossi, E. Nowak, C. Zambelli, C. Pellissier, S. Bernasconi, G. Cibrario, K. El Hajjam, R. Crochemore, J.F. Nodin, P. Olivo and L. Perniola

    In: IEEE International Electron Device Meeting (IEDM), pp. 1-2, Dec 2016

Book chapters

21. "Reliability of 3D NAND Flash Memories"

    A. Grossi, C. Zambelli and P. Olivo

    In: 3D Flash Memories, Editor: R. Micheloni, Springer, pp. 29-62, May 2016

Others

22. "Impact of ALD process parameters on HfO2 based 1T-1R RRAM inter-cell variability and switching properties"

    A. Grossi, E. Perez, C. Zambelli, P. Olivo, R. Roelofs, and C. Wenger

    In: IEEE Semiconductor Interface Specialists Conference (SISC), pp. 1-2, Dec 2015

Acknowledgements

During these three years of work toward my Ph.D. graduation I had the pleasure to spend my time with many people that I would like to thank: they all contributed to my personal and professional growth and I am really grateful to all of them.

First of all I would like to thank Prof. Piero Olivo and Dr. Cristian Zambelli: they've been driving my first steps in the academic research and they were always available to discuss my results, giving new ideas and pushing me to improve. It was a great pleasure and honor to work with them, they have been great teachers and also good friends.

Special thanks also to Dr. Christian Wenger for giving me the opportunity to work at IHP during my second year of Ph.D and to Dr. Etienne Nowak for driving my research activities at CEA-Leti during my third year: I found both of them amazing guides, and I really enjoyed working with them and learning from them. I would also like to thank all my colleagues at Università degli Studi di Ferrara, IHP and CEA-Leti that shared with me joys and efforts during this period, in particular Lorenzo Zuolo who shared the office in Ferrara with me during these three years.

Last but not least the most sincere thanks are for my family and to Eva for all the aid they offered me. They have been always supportive, patient and ready to help me in every situation: if I've been able to bring these three years of Ph.D. to a successful conclusion it's mainly because of them.

Dottorati di ricerca

Il tuo indirizzo e-mail

alessandro.grossi@unife.it

Oggetto:

Dichiarazione di conformità della tesi di Dottorato

Io sottoscritto Dott. (Cognome e Nome)

Grossi Alessandro

Nato a:

Ferrara

Provincia:

Ferrara

Il giorno:

05/02/1988

Avendo frequentato il Dottorato di Ricerca in:

Scienze dell'Ingegneria

Ciclo di Dottorato

29

Titolo della tesi:

Emerging non volatile memories reliability

Titolo della tesi (traduzione):

Affidabilità di memorie non volatili emergenti

Tutore: Prof. (Cognome e Nome)

Olivo Piero

Settore Scientifico Disciplinare (S.S.D.)

ING-INF/01

Parole chiave della tesi (max 10):

Memory, characterization, modeling, simulations, reliability, memorie, caratterizzazione, simulazione, affidabilità

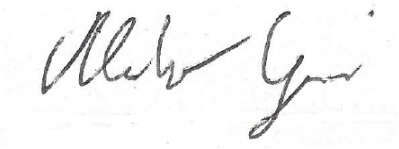
Consapevole, dichiara

CONSAPEVOLE: (1) del fatto che in caso di dichiarazioni mendaci, oltre alle sanzioni previste dal codice penale e dalle Leggi speciali per l'ipotesi di falsità in atti ed uso di atti falsi, decade fin dall'inizio e senza necessità di alcuna formalità dai benefici conseguenti al provvedimento emanato sulla base di tali dichiarazioni; (2) dell'obbligo per l'Università di provvedere al deposito di legge delle tesi di dottorato al fine di assicurarne la conservazione e la consultabilità da parte di terzi; (3) della procedura adottata dall'Università di Ferrara ove si richiede che la tesi sia consegnata dal dottorando in 2 copie di cui una in formato cartaceo e una in formato pdf non modificabile su idonei supporti (CD-ROM, DVD) secondo le istruzioni pubblicate sul sito: http://www.unife.it/studenti/dottorato alla voce ESAME FINALE

– disposizioni e modulistica; (4) del fatto che l'Università, sulla base dei dati forniti, archivierà e renderà consultabile in rete il testo completo della tesi di dottorato di cui alla presente dichiarazione attraverso l'Archivio istituzionale ad accesso aperto "EPRINTS.unife.it" oltre che attraverso i Cataloghi delle Biblioteche Nazionali Centrali di Roma e Firenze; DICHIARO SOTTO LA MIA RESPONSABILITA': (1) che la copia della tesi depositata presso l'Università di Ferrara in formato cartaceo è del tutto identica a quella presentata in formato elettronico (CD-ROM, DVD), a quelle da inviare ai Commissari di esame finale e alla copia che produrrò in seduta d'esame finale. Di conseguenza va esclusa qualsiasi responsabilità dell'Ateneo stesso per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi; (2) di prendere atto che la tesi in formato cartaceo è l'unica alla quale farà riferimento l'Università per rilasciare, a mia richiesta, la dichiarazione di conformità di eventuali copie; (3) che il contenuto e l'organizzazione della tesi è opera originale da me realizzata e non compromette in alcun modo i diritti di terzi, ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto l'Università è in ogni caso esente da responsabilità di qualsivoglia natura civile, amministrativa o penale e sarà da me tenuta indenne da qualsiasi richiesta o rivendicazione da parte di terzi; (4) che la tesi di dottorato non è il risultato di attività rientranti nella normativa sulla proprietà industriale, non è stata prodotta nell'ambito di progetti finanziati da soggetti pubblici o privati con vincoli alla divulgazione dei risultati, non è oggetto di eventuali registrazioni di tipo brevettale o di tutela. PER ACCETTAZIONE DI QUANTO SOPRA RIPORTATO

Firma del dottorando

Ferrara, li _30/01/2017_____ (data) Firma del Dottorando
_Alessandro_Grossi_____



Firma del Tutore

Visto: Il Tutore Si approva Firma del Tutore
_Piero_Olivo_____