Received 12 June 2019; revised 22 July 2019; accepted 23 July 2019. Date of publication 30 July 2019; date of current version 13 August 2019. The review of this paper was arranged by Editor M. Liu.

Digital Object Identifier 10.1109/JEDS.2019.2931769

Toward Reliable Multi-Level Operation in RRAM Arrays: Improving Post-Algorithm Stability and Assessing Endurance/Data Retention

EDUARDO PÉREZ [1], CRISTIAN ZAMBELLI [2] (Member, IEEE), MAMATHAMBA KALISHETTYHALLI MAHADEVAIAH [1], PIERO OLIVO[2], AND CHRISTIAN WENGER [1,3]

1 IHP-Leibniz-Institut für Innovative Mikroelektronik, 15236 Frankfurt, Germany
2 Dipartimento di Ingegneria, Universitá degli Studi di Ferrara, 44122 Ferrara, Italy
3 Brandenburg Medical School Theodor Fontane, 16816 Neuruppin, Germany

CORRESPONDING AUTHOR: E. PÉREZ (e-mail: perez@ihp-microelectronics.com)

This work was supported in part by the German Research Foundation (DFG) in the frame of research group under Grant FOR2093, and in part by the Università degli Studi di Ferrara under the Bando per il finanziamento della ricerca scientifica "Fondo per L'Incentivazione alla Ricerca" (FIR)–2018.

ABSTRACT Achieving a reliable multi-level operation in resistive random access memory (RRAM) arrays is currently a challenging task due to several threats like the post-algorithm instability occurring after the levels placement, the limited endurance, and the poor data retention capabilities at high temperature. In this paper, we introduced a multi-level variation of the state-of-the-art incremental step pulse with verify algorithm (M-ISPVA) to improve the stability of the low resistive state levels. This algorithm introduces for the first time the proper combination of current compliance control and program/verify paradigms. The validation of the algorithm for forming and set operations has been performed on 4-kbit RRAM arrays. In addition, we assessed the endurance and the high temperature multi-level retention capabilities after the algorithm application proving a 1 k switching cycles stability and a ten years retention target with temperatures below 100 °C.

INDEX TERMS RRAM, arrays, algorithm instabilities, multi-level, data retention, accelerated test.

I. INTRODUCTION

The resistive random access memory (RRAM) technology based on metal-oxide dielectrics is one of the most promising candidates for the next generation of non-volatile memory (NVM) applications due to their potential for high-density integration, high-speed switching operations, low-power consumption, and full compatibility with the CMOS technology [1]. The functionality of this technology is based on the resistive switching (RS) effect, which is attributed to the creation and disruption of nanometer-scale conductive filaments (CFs) in the insulator layer, consisting of oxygen vacancies (V_O) [2]. The creation process moves the RRAM device in a low resistive state (LRS), whereas the disruption process brings it in a high resistive state (HRS), which are referred as set and reset operations, respectively [3]. To activate the switching behavior, a preliminary soft breakdown in the dielectric material, referred as forming operation, is required in most RRAM technologies [4], [5]. Such operation plays a crucial role in the device performance [6], [7].

Because of the stochastic nature of the CF formation and rupture, resistive states and their related switching parameters usually vary in a large scale (from device-to-device and from cycle-to-cycle), which compromises the RRAM reliability [8], [9]. Several methods have been proposed to improve the RS stability in RRAM devices, such as: active electrode modification [10], [11], metal-oxide interfaces engineering [12], doping of the switching layer [13], [14] or by using optimized programming algorithms [15], [16]. One of our prior works demonstrated a drastic reduction of the device-to-device variability by combining Al-doped HfO_2 as switching layer with an optimized pulsed programming algorithm [17].

That reduction paved the way for the introduction of the multi-bits per cell paradigm that is usually achieved

2168-6734 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

TABLE 1. Comparison with different multi-level programming algorithms in RRAM technology presented in literature and their associated reliability assessment tests.

Ref.	Algorithm foundation	N° of levels	Test vehicle	Reliability evaluations
[18]	ISPVA-like	4 - 8	Single cell	Endurance Read disturb. Low temp. retention
[19]	Compliance control	8	Single cell	Read disturb. Low temp. retention
[20]	Dispersion aware PVA	4	Single cell	-
[21]	ISPVA-like	16	Single cell	Low temp. retention
[22]	ISPVA-like	92	Single cell	-
[23]	HRS control	4	10 cells	Endurance
[24]	Compliance control with smart target	4 - 8	2× 4-kbit arrays	Endurance
This work	M-ISPVA (LRSs) ISPVA (HRS)	4	4-kbit arrays	Post-algorithm instability Endurance Low and high temp. retention

by tailoring the CF properties through multi-level programming algorithms [18]–[24]. Table 1 summarizes the most common approaches found in the literature. However, there are a number of intrinsic concerns that could limit the reliable use of a multi-bit approach. Among them, there is a post-algorithm relaxation of the LRS and HRS distributions seen in several RRAM technologies [25], a difficult control of the multi-level endurance and a relatively poor high temperature data retention. Those issues have a consequence on the realization of high density products, confining the usage scope of RRAM technology in few scenarios. Further, the use of RRAM devices in neuromorphic applications, where the mimicking of biological synapses requires an analog-like behavior and therefore a stable multi-level capability [26], would be threatened *a priori*.

In this work, we demonstrate a well-controlled and reliable multi-level operation in 1-transistor-1-resistor (1T1R) RRAM architectures based on Al-doped HfO_2 and integrated in 4-kbit arrays through the introduction of a new degree of freedom in the state-of-the-art incremental step pulse with verify algorithm (ISPVA), namely, the compliance current control, along with the target read-out current (I_{trg}). To the best of our knowledge, this is the first attempt to combine these two parameters in order to improve the reliability of multi-level RRAM technology, which will lead the way towards the multi-bit operation as well as the implementation of artificial synaptic devices [27]. The electrical characterization of the new algorithm, hereafter called multilevel-ISPVA (M-ISPVA), on RRAM array test vehicles will allow the investigation of the post-algorithm instability, memory endurance, and low/high temperature data retention capabilities. Finally, retention accelerated tests are performed to project the lifetime reliability up to 10 years as a further assessment for our developed algorithm on arrays.

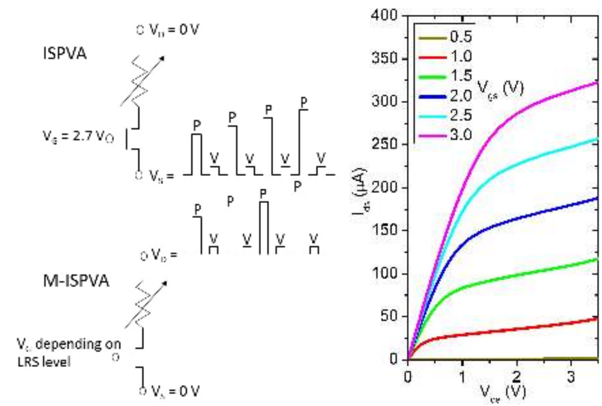

FIGURE 1. Schematic illustration of the ISPVA (HRS) and of the M-ISPVA (LRSs) approaches (left), and I-V characteristics of the transistor setting the compliance current (right).

II. EXPERIMENTAL

The test vehicles used for the implementation of the multi-level approach are 4-kbit memory arrays organized in 64 pages, each consisting of 64 1T1R cells. The architecture of the chip is described in more detail in [28]. The 1T1R devices are constituted by a NMOS transistor, manufactured in 0.25 μm CMOS technology, whose drain is connected in series to a variable resistor integrated on the metal line 2 of the CMOS process. Such resistor is a metal-insulator-metal (MIM) structure consisting of a TiN/Al:HfO_2/Ti/TiN stack with 150 nm TiN top and bottom electrode layers deposited by magnetron sputtering, a 7 nm Ti layer (under the TiN top electrode) and a 6 nm Al:HfO_2 layer based on ∼10% Al-doped HfO_2 ($Hf_{1-x}Al_xO_y$) grown by atomic layer deposition (ALD) at 300 °C. After patterning the MIM structure with an area of about 0.4 μm^2, an additional thin Si_3N_4 layer was deposited to protect the RRAM cell.

The key of a good multi-level approach in RRAM devices is in the accurate control of the multiple conductive states. The strategy followed in this work is to define one HRS and three LRSs, which leads to four levels cells. The ISPVA has shown to be an effective way to reduce the device-to-device variability by establishing I_{trg} values: a top value for HRS and a bottom value for LRS [17]. As schematically depicted in Fig. 1, during the ISPVA programming a sequence of increasing voltage amplitude pulses (P) of 10 μs duration are applied either on the drain side of the transistor connected to the MIM resistor during forming and set operations or on the source terminal of the transistor during reset operations. After every pulse (P) a read-verify operation (V) is performed by applying a read-out pulsed voltage of 0.2 V amplitude and 10 μs duration on the drain side of the transistor to check whether the I_{trg} target is achieved. In positive case the programming operation is stopped and the read-out current measured is saved as the read-out current value just after the transition. If the I_{trg} target is not achieved, the programming operation is stopped when a maximum voltage amplitude value is achieved: 3.5 V for reset and set operations and 5 V for the forming operation.

However, the definition of I_{trg} as a lower limit for LRS in the ISPVA approach is found deficient in ensuring three discrete levels as desired in our multi-level paradigm. As already claimed by Kim *et al.* [21], the abrupt nature of forming and set processes could easily program the RRAM devices into a conductive state outside of the desired range. Therefore, an upper target must be defined along with I_{trg} by using the current compliance control provided by the transistor in the 1T1R architecture, thus defining the M-ISPVA algorithm for LRSs. In the test vehicles considered in this work, the compliance current is effectively modulated by the gate voltage (V_g) applied on the transistor, whose I-V characteristics are shown in Fig. 1. Three $<I_{trg}, V_g>$ pairs were defined to evaluate the multi-level capabilities during forming: <20, 1.2>, <30, 1.4>, and <40, 1.6> <μA, V>. In the following we will refer to the LRS levels generated by those pairs as LRS1, LRS2, and LRS3, respectively. It is worth to point out that in order to ensure a good HRS, an I_{trg} value equal to 5 μA was considered along with a $V_g = 2.7$ V exploited in the ISPVA algorithm for the reset operation to minimize the series resistance of the transistor.

III. RESULTS AND DISCUSSION

A. FORMING OPERATION

In order to activate the RS behavior, the forming operation was performed by using each LRS $<I_{trg}, V_g>$ pair on different batches of pages of the array. The amplitude of voltage pulses applied on the transistor's drain during the M-ISPVA was swept in the range of 2-5 V with a step of 0.01 V. The three post-algorithm cumulative distribution functions (CDFs) of the read-out currents just after the forming transition and at the end of the M-ISPVA are shown in Fig. 2. An architecture constraint of the test setup imposes read-verify operations (V) to be continuously performed on the array until the M-ISPVA finishes (even on cells already formed). The read-out current measured during the last read-verify operation of the whole sequence is saved as the read-out current value at the end of the operation. The time interval between the switching transition of the RRAM cells and the end of the algorithm is averagely 3 minutes. This leads to the observation of post-algorithm instabilities yielding to perturbations of the conductive state of the RRAM devices under test [25]. The most likely reason for the occurrence of such instabilities is the further electrical stress applied by the additional read-verify operations. This is appreciable in the figure showing a lower tail of the CDFs with about 13% of the cells redistributing their read-out current below the I_{trg} and about 3% of the cells crossing the adjacent LRS level. This result is in accordance with [29]. Fig. 3 illustrates the device-to-device correlation between the current values measured just after the forming transition and at the end of the M-ISPVA. The lack of correlation between the current values shows that the tailing and the redistribution are stochastic processes whose detrimental effects on the multi-level operation should be prevented.

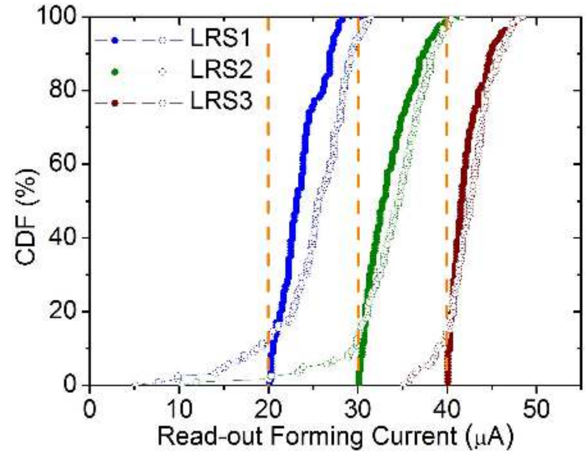

FIGURE 2. CDFs of the read-out currents just after the forming transition (full symbols) and at the end of the forming operation (empty symbols) for the three LRS levels : <20, 1.2> (in blue), <30, 1.4> (in green), and <40, 1.6> (in red) <μA, V>.

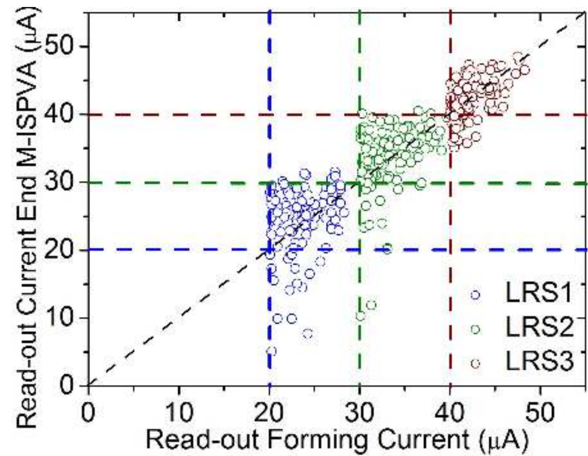

FIGURE 3. Device-to-device correlation of the read-out currents measured just after the forming transition and at the end of the M-ISPVA for the three LRS levels: <20, 1.2> (in blue), <30, 1.4> (in green), and <40, 1.6> (in red) <μA, V>.

The definition of suitable values for $<I_{trg}, V_g>$ pairs in the M-ISPVA is fundamental in providing multi-level capability while preventing post-algorithm instabilities, therefore outperforming ISPVA [17] and compliance-based algorithms [19] in this context. Multiple combinations of $<I_{trg}, V_g>$ were tested during the forming operation besides those exploited for former tests. Three I_{trg} values were considered, namely 20, 30, and 40 μA in combination with four V_g values: 1.0, 1.2, 1.4, and 1.6 V. The results of these measurements are summarized in Fig. 4. On the one hand, if the V_g value is too small, the RRAM cells can not achieve the target value I_{trg} performing the forming transition. In this specific situation the current value saved as the read-out current just after the transition is the value measured at the end of the forming operation. This is the reason why both CDFs appear partially or even totally overlapped. On the other hand, the dispersion of the CDF increases beyond control

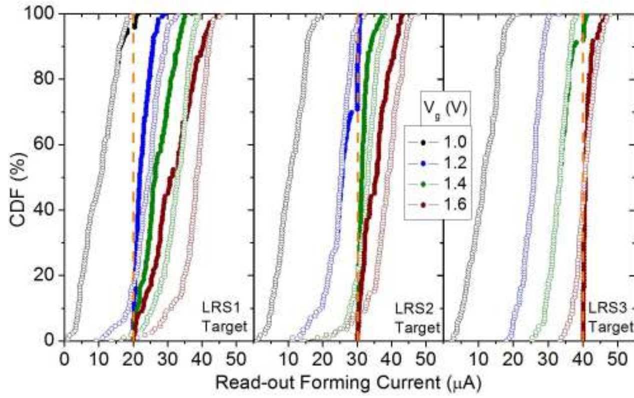

FIGURE 4. CDFs of the read-out currents just after the forming transition (solid symbols) and at the end of the forming operation (open symbols) for the twelve <I_{trg}, V_g> pairs.

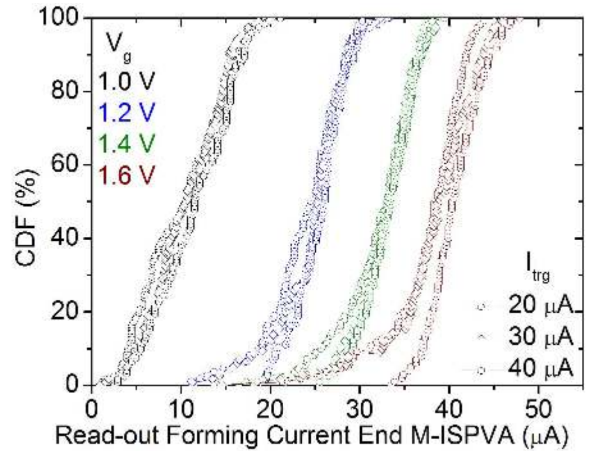

FIGURE 6. CDFs of the read-out currents at the end of the M-ISPVA during the forming operation for all <I_{trg}, V_g> pairs.

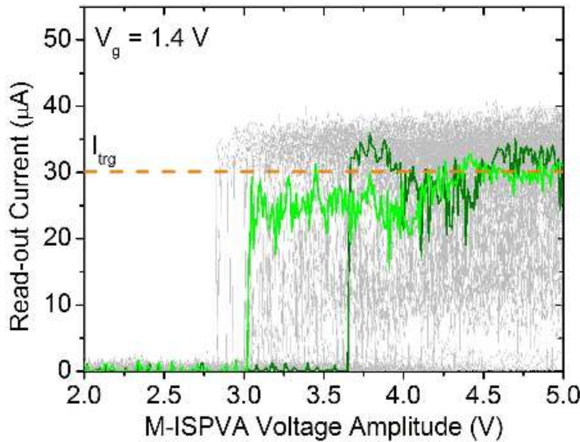

FIGURE 5. Evolution of the read-out forming current values during the M-ISPVA by using the <30, 1.4> <μA, V> pair on 128 RRAM devices (in grey). The figure highlights the evolution on a cell reaching the forming I_{trg} with one step (in dark green), and a cell requiring more than one step to reach the forming I_{trg} (in bright green).

with large V_g values, which makes difficult to fit several conductive levels within the available read-out current range (from 0 to 50 μA).

To explain this behavior, the evolution of the read-out currents measured during the M-ISPVA forming operation was carefully investigated. For instance, we will consider the batch of 128 RRAM cells formed by using the <30, 1.4> <μA, V> pair (see Fig. 5). In the pristine state (in which the read-out current is essentially equal to 0 μA) almost the whole voltage applied by the M-ISPVA on the RRAM device drops on the MIM cell. When a soft breakdown happens in the device's dielectric, a sudden increase of the read-out current takes place. According to Ielmini [30], the amplitude of such increase is limited by the voltage divider formed between the transistor resistance (controlled by V_g) and the MIM cell resistance. However, a successful forming transition crossing I_{trg} is not granted for all devices, since some of them requires additional voltage steps to converge to the target current. In addition, even if the forming transition is performed, the post-algorithm instabilities are appreciable. Such instabilities are totally independent on the I_{trg} value defined and can be limited only by the current compliance control imposed by V_g, as shown in Fig. 6. This is an added value provided by the M-ISPVA compared to other algorithms.

B. RESET AND SET OPERATIONS

After forming the three LRS levels, the feasibility of switching the RRAM devices from the LRS1, LRS2, and LRS3 to the HRS and viceversa was tested by performing reset and set operations with ISPVA and M-ISPVA, respectively. In both algorithms, the amplitude of voltage pulses applied on the source (reset) and drain (set) was swept in the range of 0.2-3.5 V with a voltage step of 0.1 V. As shown in Fig. 7, the CDFs of the read-out currents after the reset operation are completely located below the I_{trg} defined for the HRS (5 μA) in the ISPVA algorithm. After reset, the three LRSs were obtained after the set operation with M-ISPVA using the same <I_{trg}, V_g> pairs defined for the forming operation. About 5% of the devices for each LRS level cross the I_{trg} value of the adjacent level at the end of the algorithm. The CDF corresponding to the LRS3 displays a tail at high current values, however this does not interfere with the multi-level definition. Such tail can be attributed to a proximity with the upper limit of the read-out current range (50 μA) where the switching behavior of our devices starts to be unstable. Unlike in the forming operation, in set operation the redistribution of the current values during the time interval between the switching transition and the end of the algorithm (30 seconds in average) is almost suppressed, leading to a more reliable multi-level operation (see Fig. 8). An explanation of this phenomenon can be found in [31]. At the beginning of the first reset operation, the oxygen vacancies placed at the interface between the

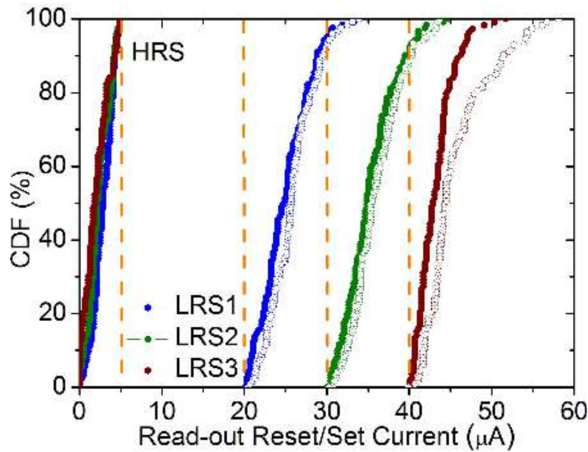

FIGURE 7. CDFs of the read-out currents after the reset operation with ISPVA (HRS) and after the set operation with M-ISPVA (LRSs). LRSs data are shown after the set transition (solid symbols) and at the end of the algorithm (open symbols) : <20, 1.2> (in blue), <30, 1.4> (in green), and <40, 1.6> (in red) <μA, V>.

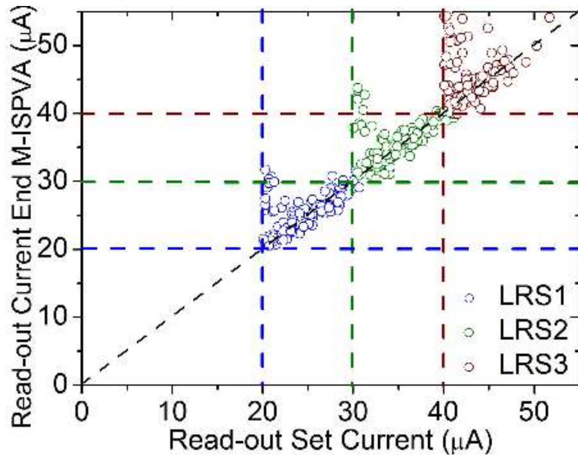

FIGURE 8. Device-to-device correlation of the read-out currents measured just after the set transition and at the end of the M-ISPVA for the three LRS levels: <20, 1.2> (in blue), <30, 1.4> (in green), and <40, 1.6> (in red) <μA, V>.

Al:HfO_2 insulator and the bottom electrode drift towards the CF, thus strengthening the CF tip in the first voltage steps of the ISPVA algorithm applied during reset operation. Such increase, independent on the starting LRS level, vanishes when some of the oxygen vacancies drift and open a gap in the CF. The subsequent set operation will fill the gap previously opened only by a few oxygen vacancies. This process stabilizes the CF strongly reducing the post-algorithm instabilities for further operations.

As for forming operation, it is important to define appropriate <I_{trg}, V_g> pairs to establish multiple LRSs in RRAM devices without endangering their reliability with unoptimized M-ISPVA conditions. Fig. 9 shows the CDFs of the read-out currents just after the set transition and at the end of the M-ISPVA for the same twelve <I_{trg}, V_g> pairs considered in the forming analysis. Similarly, if the V_g value

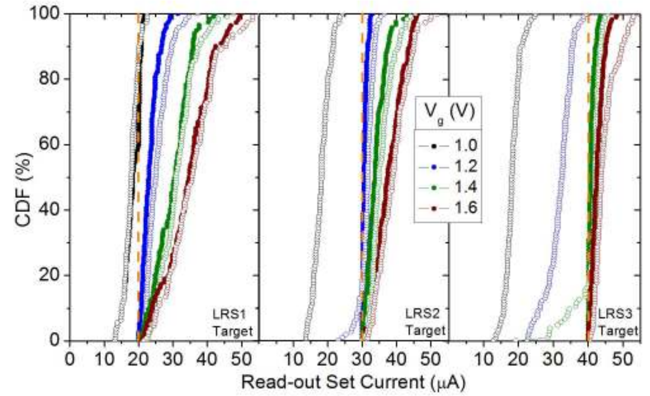

FIGURE 9. CDFs of the read-out currents just after the set transition (solid symbols) and at the end of the set operation (open symbols) for the twelve <I_{trg}, V_g> pairs.

is too small the RRAM cells can not perform the set transition. However, during the set operation this limitation was found in a smaller group of pairs. The overlapping of both CDFs in this situation was already clarified in the forming section. Once again, if the V_g value is too large the dispersion of the CDF becomes uncontrollable to fit different LRS levels within the read-out current range (for instance <20, 1.6> <μA, V>). It is worth to point out that if the CDF shape is narrow and too close to the I_{trg} value after the set transition, as depicted in Fig. 9 for <30, 1.2> and <40, 1.4> <μA, V>, some cells could suffer from post-algorithm instability. Additionally, as illustrated in Fig. 10, the voltages required to perform the set operation by using these pairs are strongly increased. Therefore, a trade-off should be exercised. Extremely narrow current CDFs lead to more effective definition of discrete LRS levels within the read-out current range. However, the energy consumption of the memory increases significantly because of the larger number of M-ISPVA pulses and voltages required to perform set operations, as well as, because of the error detection and correction systems required to handle the minor post-algorithm instabilities.

The evolution of the read-out currents measured during the whole M-ISPVA on a group of 128 RRAM cells are shown in Fig. 11. The results confirm the enhanced suppression of the post-algorithm instabilities compared to those observed during the forming operation. The broader range of current values covered by the RRAM devices during the M-ISPVA after the set transition by using <30, 1.6> <μA, V> instead of <30, 1.4> <μA, V> is a further evidence of the detrimental increase of the CDF dispersion observed in Fig. 9 when the V_g is too large. The current evolution for the pair <30, 1.2> <μA, V> shown in Fig. 11 illustrates that the I_{trg} value is achieved progressively after the initial current jump. This behavior explains the extremely narrow CDF shown in Fig. 9 and why several devices cross down the I_{trg} limit, which provoke the post-algorithm instabilities.

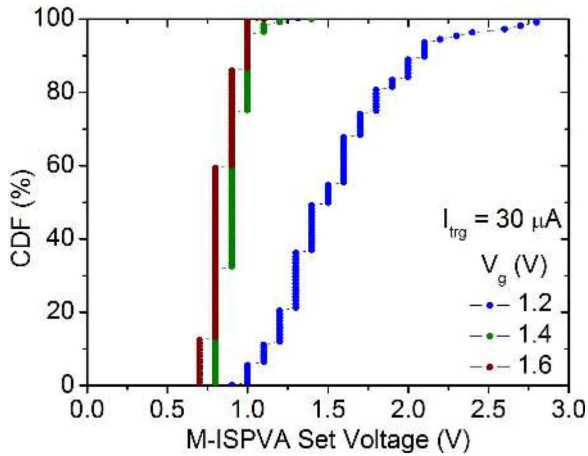

FIGURE 10. CDFs of the set voltages corresponding to three <I_{trg}, V_g> pairs with I_{trg} equal to 30 μA.

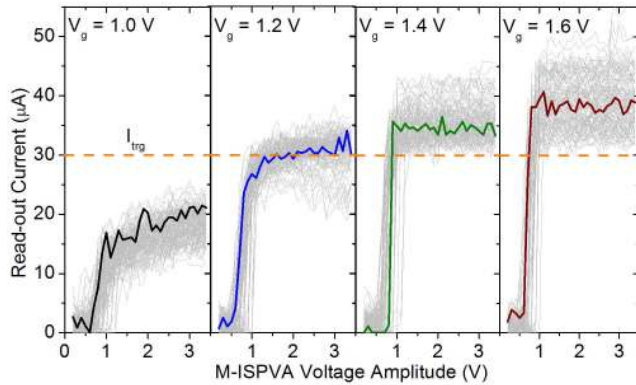

FIGURE 11. Evolution of the read-out current values during the M-ISPVA for the set operation on 128 RRAM devices (in grey) as a function of V_g by considering I_{trg} equal to 30 μA. The figure highlights the evolution on a specific cell.

C. ENDURANCE

After proving the effectiveness of the M-ISPVA algorithm in countering the post-algorithm instabilities we have evaluated the multi-level endurance capabilities of the RRAM array. To this extent, we have performed 1k switching cycles (i.e., consecutive set/reset operations) by monitoring the read-out current of the four different resistance states at every cycle. This procedure is comparable to what has been used in [24]. As shown in Fig. 12, the M-ISPVA approach is not threatening the reliability of the memory since the LRSs and HRS are not degraded. Further, the post-algorithm instability is not a concern since the read window margin between different states is kept intact at the end of the programming algorithm, thus strongly reducing the number of cells moving to a wrong resistance level.

D. LOW AND HIGH TEMPERATURE DATA RETENTION

The data retention still remains a key issue in RRAM technology [17], [21], [32]. Thermal stability in multi-level switching is even more crucial to prevent the overlapping

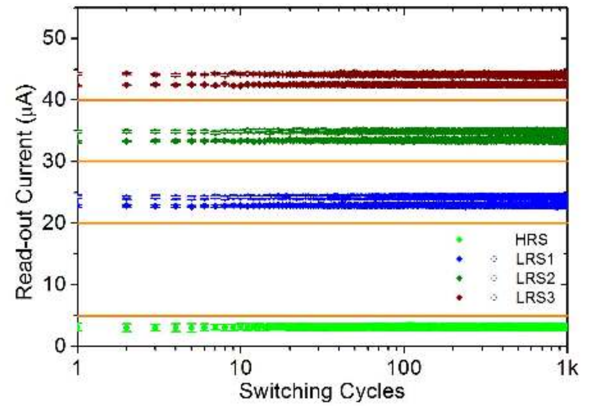

FIGURE 12. Endurance characterization of the four resistance levels achieved with M-ISPVA (LRSs) and ISPVA (HRS) along 1k switching cycles. For LRSs data are shown after the set transition (solid symbols) and at the end of the algorithm (open symbols). The error bars refer to the device-to-device dispersion for 128 RRAM devices.

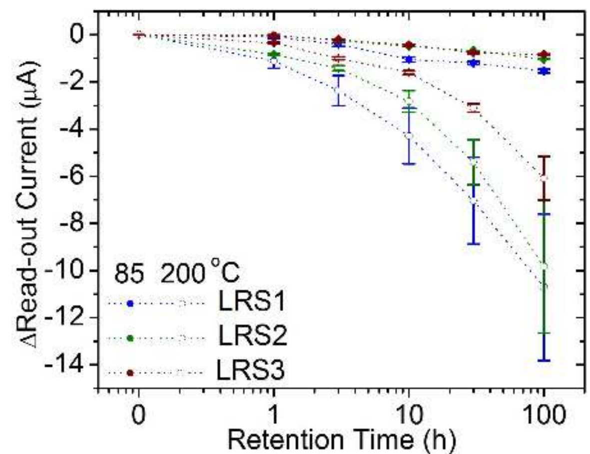

FIGURE 13. Variation of the average and dispersion of the current values sampled during the retention test compared to the beginning of the bake at 85 °C (solid symbols) and 200 °C (open symbols) for the three LRSs conductive levels.

between adjacent LRS levels. In order to assess the low and high temperature data retention of the four conductive levels, the RRAM devices in the array were baked at 85 °C and 200 °C for 100 hours. According to one of our prior publications [17], HRS remains stable during the retention tests. Therefore, we monitored the three LRSs at different sampling times. In Fig. 13 the average and the dispersion of the read-out currents measured at each sampling time are depicted as the variation from their initial values at 0 hours (ΔRead-out). At 85 °C the behavior is stable and only a slight decrease of the mean values occurs. In contrast, at 200 °C the thermal stress has a strong impact. The mean values feature a monotone decrease with a faster rate strongly depending on the LRS conductivity [33]. In addition, the dispersion values grow with the decrease of the LRS conductivity, which is caused by the growth of the tails of the current CDFs, as already observed in [17].

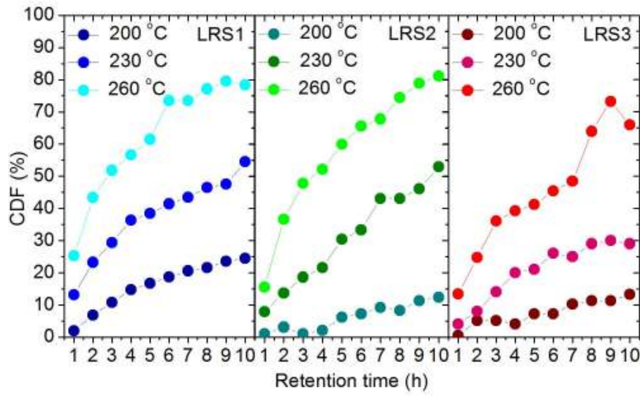

FIGURE 14. Distributions of the retention failure times for the three LRSs: LRS1 (blue), LRS2 (green), and LRS3 (red); at the three annealing temperatures: 200, 230, and 260 °C.

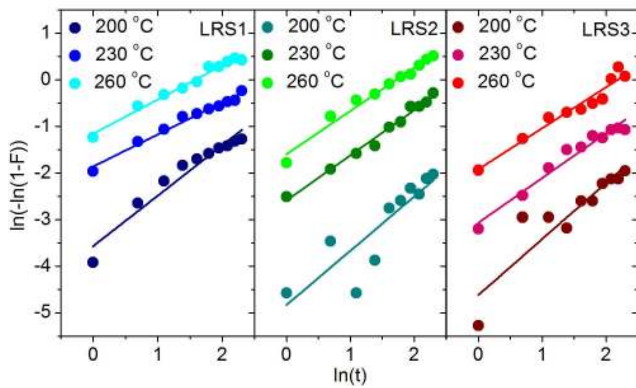

FIGURE 15. Weibull plots of the retention failure times distributions for the three LRSs: LRS1 (blue), LRS2 (green), and LRS3 (red); at the three annealing temperatures: 200, 230, and 260 °C.

The stability of the multi-level information storage must be guaranteed for at least 10 years [34], [35]. To assess the maximum working temperature that allows such a data retention lifetime, we exploited a high temperature accelerated test [35], [36]. Sufficient statistical information from the LRSs retention failures can be obtained in a practical period of time, namely 10 hours, by using bake temperatures of 200 °C, 230 °C, and 260 °C [36], as shown in Fig. 14. As depicted in Fig. 15, the distributions of the retention failure times can be modeled by using a Weibull distribution. The retention mean time to failure (MTTF) of each LRS level programmed on the RRAM devices can be calculated for the three temperatures according to the following equation [36]:

$$MTTF = \alpha\Gamma(1 + 1/\beta) \quad (1)$$

where Γ(x) is the Euler's gamma function and α and β the scale and shape parameters of the Weibull distribution, respectively. Assuming an Arrhenius dependency of the MTTF with the temperature we can state [36]:

$$MTTF = A\exp\left(\frac{E_a}{kT}\right) \quad (2)$$

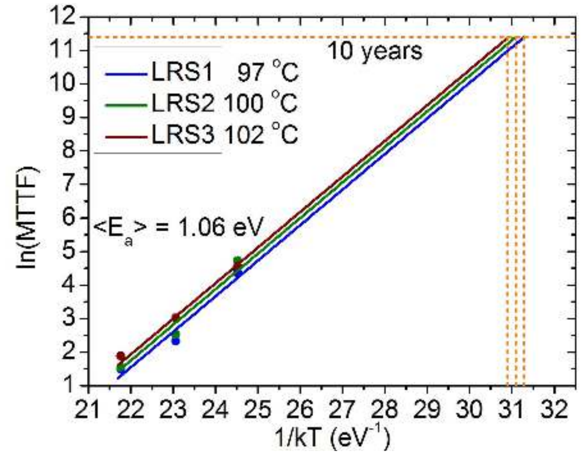

FIGURE 16. Extrapolation of MTTF to a value of 10 years lifetime for the three LRSs: LRS1 (blue), LRS2 (green), and LRS3 (red). The maximum temperature value that guarantees such a retention lifetime for each LRS is depicted.

where *A* is a pre-exponential constant, *k* is the Boltzmann's constant, *T* is the operating temperature of the device and E_a is the extracted activation energy. The three E_a values obtained for LRS1, LRS2, and LRS3 according to the data in Fig. 15 are: 1.05 eV, 1.16 eV, and 0.97 eV, respectively. These results are in line with the values already reported in literature [36] and evidence a universal mechanism associated to the physical processes involved in the degradation of the CF regardless the morphology of the CF in the conductivity range of 20-50 μA [33], [34]. An average E_a value of 1.06 eV can be assumed to extrapolate the maximum temperature value that guarantees a MTTF of 10 years (see Fig. 16). According to Ielmini *et al.* [33], a stronger CF leads to a more stable LRS. However, the improvement achieved on the data retention by increasing the CF conductivity from 20 to 40 μA is only about 5 °C, which is extremely small. In terms of multi-level operation, the temperature limit must be established as the most restrictive value, namely 97 °C.

IV. CONCLUSION

In this study, we evaluated the feasibility of reliable multi-level operation in Al-doped HfO_2 RRAM arrays by introducing the M-ISPVA algorithm. This new approach considers two parameters to shape the resistance levels: I_{trg} and V_g. Four conductive levels were successfully established in each 1T1R device of the memory. The importance of defining suitable <I_{trg}, V_g> pairs in order to reduce post-algorithm instabilities while accomplishing endurable multi-bit operation was evidenced. The endurance up to 1k switching cycles has been demonstrated. Finally, the low and high temperature data retention were assessed reporting an E_a of 1.06 eV in accelerated tests. This value materializes in a 10 years lifetime with working temperatures below 100 °C, ensuring reliable multi-bit data storage for many NVM applications.

REFERENCES

[1] H.-S. P. Wong *et al.*, "Metal–Oxide RRAM," *Proc. IEEE*, vol. 100, no. 6, pp. 1951–1970, Jun. 2012. doi: 10.1109/JPROC.2012.2190369.

[2] Y. S. Lin *et al.*, "Resistive switching mechanisms relating to oxygen vacancies migration in both interfaces in Ti/HfO_x/Pt memory devices," *J. Appl. Phys.*, vol. 113, no. 6, 2013, Art. no. 064510. doi: 10.1063/1.4791695.

[3] B. J. Choi *et al.*, "Resistive switching mechanism of TiO_2 thin films grown by atomic-layer deposition," *J. Appl. Phys.*, vol. 98, no. 3, 2005, Art. no. 033715. doi: 10.1063/1.2001146.

[4] G. Bersuker *et al.*, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *J. Appl. Phys.*, vol. 110, no. 12, 2011, Art. no. 124518. doi: 10.1063/1.3671565.

[5] B. Traoré, P. Blaise, E. Vianello, L. Perniola, B. De Salvo, and Y. Nishi, "HfO_2-based RRAM: Electrode effects, Ti/HfO_2 interface, charge injection, and oxygen (O) defects diffusion through experiment and Ab initio calculations," *IEEE Trans. Electron Devices*, vol. 63, no. 1, pp. 360–368, Jan. 2016. doi: 10.1109/TED.2015.2503145.

[6] A. Kalantarian *et al.*, "Controlling uniformity of RRAM characteristics through the forming process," in *Proc. IEEE Int. Rel. Phys. Symp.*, Apr. 2012, pp. 1–5. doi: 10.1109/IRPS.2012.6241874.

[7] A. Grossi *et al.*, "Relationship among current fluctuations during forming, cell-to-cell variability and reliability in RRAM arrays," in *Proc. IEEE Int. Memory Workshop*, May 2015, pp. 1–4. doi: 10.1109/IMW.2015.7150303.

[8] S. Yu, X. Guan, and H.-S. P. Wong, "On the switching parameter variation of metal oxide RRAM—Part II: Model corroboration and device design strategy," *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 1183–1188, Apr. 2012. doi: 10.1109/TED.2012.2184544.

[9] A. Grossi *et al.*, "Fundamental variability limits of filament-based RRAM," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2016, pp. 1–4. doi: 10.1109/IEDM.2016.7838348.

[10] C.-Y. Lin *et al.*, "Effect of top electrode material on resistive switching properties of ZrO_2 film memory devices," *IEEE Electron Device Lett.*, vol. 28, no. 5, pp. 366–368, May 2007. doi: 10.1109/LED.2007.894652.

[11] T. Cabout *et al.*, "Role of Ti and Pt electrodes on resistance switching variability of HfO_2-based resistive random access memory," *Thin Solid Films*, vol. 533, pp. 19–23, Apr. 2013. doi: 10.1016/j.tsf.2012.11.050.

[12] Y. Y. Chen *et al.*, "Endurance/retention trade-off on HfO_2/metalcap 1T1R bipolar RRAM," *IEEE Trans. Electron Devices*, vol. 60, no. 3, pp. 1114–1121, Mar. 2013. doi: 10.1109/TED.2013.2241064.

[13] Z. Wang *et al.*, "Highly uniform, self-compliance, and forming-free ALD HfO_2-based RRAM with Ge doping," *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 1203–1208, Apr. 2012. doi: 10.1109/TED.2012.2182770.

[14] Y. Y. Chen *et al.*, "Tailoring switching and endurance/retention reliability characteristics of HfO_2/Hf RRAM with Ti, Al, Si dopants," in *Proc. Symp. VLSI Technol.*, 2014, pp. 1–2. doi: 10.1109/VLSIT.2014.6894403.

[15] J. Park *et al.*, "New set/reset scheme for excellent uniformity in bipolar resistive memory," *IEEE Electron Device Lett.*, vol. 32, no. 3, pp. 228–230, Mar. 2011. doi: 10.1109/LED.2010.2094599.

[16] A. Grossi *et al.*, "Electrical characterization and modeling of pulse-based forming techniques in RRAM arrays," *Solid-State Electron.*, vol. 115, pp. 17–25, Jan. 2016. doi: 10.1016/j.sse.2015.10.003.

[17] E. Pérez, A. Grossi, C. Zambelli, P. Olivo, R. Roelofs, and C. Wenger, "Reduction of the cell-to-cell variability in $Hf_{1-x}Al_xO_y$ based RRAM arrays by using program algorithms," *IEEE Electron Device Lett.*, vol. 38, no. 2, pp. 175–178, Feb. 2017. doi: 10.1109/LED.2016.2646758.

[18] W.-C. Chien *et al.*, "A multi-level 40nm WO_X resistive memory with excellent reliability," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2011, pp. 1–4. doi: 10.1109/IEDM.2011.6131651.

[19] A. Prakash, J. Park, J. Song, J. Woo, E.-J. Cha, and H. Hwang, "Demonstration of low power 3-bit multilevel cell characteristics in a TaO_x-based rram by stack engineering," *IEEE Electron Device Lett.*, vol. 36, no. 1, pp. 32–34, Jan. 2015. doi: 10.1109/LED.2014.2375200.

[20] F. M. Puglisi, C. Wenger, and P. Pavan, "A novel program-verify algorithm for multi-bit operation in HfO_2 RRAM," *IEEE Electron Device Lett.*, vol. 36, no. 10, pp. 1030–1032, Oct. 2015. doi: 10.1109/LED.2015.2464256.

[21] G. H. Kim *et al.*, "Four-bits-per-cell operation in an HfO_2-based resistive switching device," *Small*, vol. 13, no. 40, 2017, Art. no. 1701781. doi: 10.1002/smll.201701781.

[22] S. Stathopoulos *et al.*, "Multibit memory operation of metal-oxide bi-layer memristors," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 17532. doi: 10.1038/s41598-017-17785-1.

[23] J. Liu *et al.*, "Characteristics of multilevel storage and switching dynamics in resistive switching cell of Al_2O_3/HfO_2/Al_2O_3 sandwich structure," *J. Phys. D Appl. Phys.*, vol. 51, no. 2, 2017, Art. no. 025102. doi: 10.1088/1361–6463/aa9c15.

[24] B. Q. Le *et al.*, "Resistive RAM with multiple bits per cell: Array-level demonstration of 3 bits per cell," *IEEE Trans. Electron Devices*, vol. 66, no. 1, pp. 641–646, Jan. 2019. doi: 10.1109/TED.2018.2879788.

[25] A. Fantini *et al.*, "Intrinsic program instability in HfO_2 RRAM and consequences on program algorithms," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2015, pp. 1–4. doi: 10.1109/IEDM.2015.7409648.

[26] S. Yu, Y. Wu, R. Jeyasingh, D. Kuzum, and H.-S. Philip Wong, "An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2729–2737, Aug. 2011. doi: 10.1109/TED.2011.2147791.

[27] V. Milo *et al.*, "Multilevel HfO_2-based RRAM devices for low-power neuromorphic networks," *APL Mater.*, 2019.

[28] C. Zambelli *et al.*, "Statistical analysis of resistive switching characteristics in ReRAM test arrays," in *Proc. Int. Conf. Microelectron. Test Struct.*, 2014, pp. 27–31. doi: 10.1109/ICMTS.2014.6841463.

[29] F. Crupi *et al.*, "Implications of the incremental pulse and verify algorithm on the forming and switching distributions in ReRAM arrays," *IEEE Trans. Device Mater. Rel.*, vol. 16, no. 3, pp. 413–418, Sep. 2016. doi: 10.1109/TDMR.2016.2594119.

[30] D. Ielmini, "Modeling the universal set/reset characteristics of bipolar RRAM by field- and temperature-driven filament growth," *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4309–4317, Dec. 2011. doi: 10.1109/TED.2011.2167513.

[31] E. Perez, M. K. Mahadevaiah, C. Zambelli, P. Olivo, and C. Wenger, "Characterization of the interface-driven 1st reset operation in HfO_2-based 1T1R RRAM devices," *Solid State Electron.*, vol. 159, pp. 51–56, Sep. 2019. doi: 10.1016/j.sse.2019.03.054.

[32] M. Zhao *et al.*, "Investigation of statistical retention of filamentary analog RRAM for neuromophic computing," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2017, pp. 1–4. doi: 10.1109/IEDM.2017.8268522.

[33] D. Ielmini, F. Nardi, C. Cagli, and A. L. Lacaita, "Size-dependent retention time in NiO-based resistive-switching memories," *IEEE Electron Device Lett.*, vol. 31, no. 4, pp. 353–355, Apr. 2010. doi: 10.1109/LED.2010.2040799.

[34] B. Traoré *et al.*, "On the origin of low-resistance state retention failure in HfO_2-based RRAM and impact of doping/alloying," *IEEE Trans. Electron Devices*, vol. 62, no. 12, pp. 4029–4036, Dec. 2015. doi: 10.1109/TED.2015.2490545.

[35] J. Park *et al.*, "Investigation of state stability of low-resistance state in resistive memory," *IEEE Electron Device Lett.*, vol. 31, no. 5, pp. 485–487, May 2010. doi: 10.1109/LED.2010.2042677.

[36] E. Perez, M. K. Mahadevaiah, C. Zambelli, P. Olivo, and C. Wenger, "Data retention investigation in Al:HfO_2-based RRAM arrays by using high-temperature accelerated tests," *J. Vac. Sci. Technol. B*, vol. 37, no. 1, 2019, Art. no. 012202. doi: 10.1116/1.5054983.