RRAM Device Performances under Capacitive-Enhanced Current Programming Scheme

Thomas Bauvent, Paola Trotti, Olivier Billoint, Jean-François Nodin, Yasser Moursy, Gabriel Molas, Gaël Pillonnet

RRAM Device Performances under Capacitive-Enhanced Current Programming Scheme

Thomas Bauvent, Paola Trotti, Olivier Billoint, Jean-François Nodin, Yasser Moursy, Gabriel Molas, Gaël Pillonnet, Univ. Grenoble Alpes, CEA, Leti, France

***Abstract*—Embedded resistive random access memories (RRAM) are commonly written using voltage programming scheme. In this work, we study the device performance under an alternative programming approach. Utilizing the parasitic line capacitance combined with a current source, this scheme lowers the required programming current by a factor of 10 for a given conductance level. This effectively reduces writing energy and alleviates constraints on integration density due to electromigration and IR drop. The proposed scheme is demonstrated on 130 nm CMOS technology. The measurements show a low raw bit failure rate of 5.10^{-5} through 200k cycles. The read margin can be widened up to 25 µA using a write-verify strategy. These metrics highlight the efficiency of the proposed scheme with respect to the conventional voltage techniques.**

***Index Terms*—resistive random access memory, RRAM, OxRAM, HfO_2, operation scheme, current driver, current overshoot, switching variability**

INTRODUCTION

RESISTIVE random access memories (RRAM) are emerging as a promising candidate for non-volatile memory (NVM) applications. Various memory suppliers are now bringing it up to the embedded market [1], [2]. Oxide based RRAM are made of a stack of metal / oxide / metal and are generally called OxRAM. The cells can be switched from a high-resistive state (HRS) to a low-resistive state (LRS) by electrical means; we call it the set operation. The reset operation makes the cells go back from LRS to HRS. Most of the set schemes are based on voltage drivers. During set operation, the abrupt reduction of the resistance induces a current surge. To prevent damaging the cell, voltage drivers incorporates current limiting elements. To enhance diverse metrics, refinements of this general scheme have been studied such as write termination [3]–[5], delayed write termination [6], [7], and voltage ramps [8], [9].

The conductance in LRS is an important parameter as it is linked to metrics such as the read margin or the retention. While RRAM writing is field activated [10], the conductance in LRS has been reported to be strongly linked to the maximum current passing through the cell [11]. In light of this observation, it seems attractive to control the current rather than the voltage.

T. Bauvent, P. Trotti, O. Billoint, J-F. Nodin, Y. Moursy, G. Pillonnet are with Univ. Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France (e-mail: thomas.bauvent@cea.fr).

G. Molas is now with Weebitnano LTD.

Published version available at https://doi.org/10.1109/LED.2023.3274219
Digital Object Identifier 10.1109/LED.2023.3274219

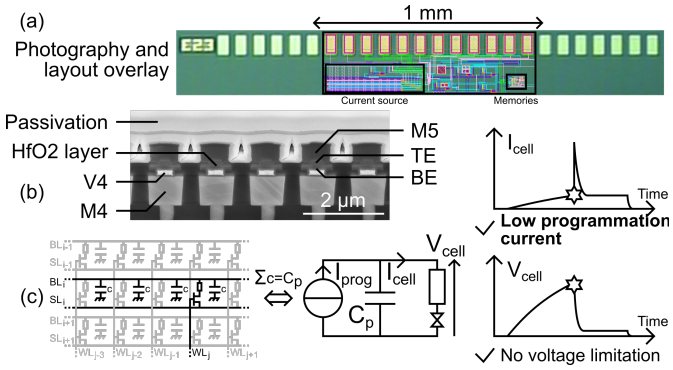

Fig. 1. (a) Photography of our test circuit. (b) Cross-section of the RRAM co-integrated in the BEoL (c) The writing scheme studied in this letter. The capacitance C_p models the parasitic line capacitance on BL in memory matrices.

Current control writing has mainly been reported on in DC / quasi-static studies [12]–[14]. To the best of the authors' knowledge, only a single publication [4] wrote about using a current driver for writing RRAM at time scales of the order of the microsecond. However, its reliance on a voltage limiter subdues its potential benefits. The advantages and tradeoffs of current based programming against the voltage based approach were not clarified so far.

Another issue RRAM faces is high programming current. Since the main conducting mechanism in LRS is filamentary, the LRS conductance does not scale with the device [15]. This problem cannot be solved by programming the cells to a lower conductance since it is detrimental to its retention and to the read margin [16], [17].

In this letter, we explore a novel current driven writing scheme. To tackle the current scaling issue, we exploit the current spike caused by the inherent parasitic line capacitance in RRAM matrices. This enables this scheme to write cells to a high conductance with limited current amplitude. This helps relax issues with IR drop and pushes back the electromigration barrier, which are both concerns for scaling. It is also beneficial toward lowering the programming energy consumption.

EXPERIMENTAL

The measurements presented hereafter were obtained on the stack TiN / Ti / HfO_2-5nm / TiN with a cell surface of 300 x 450 nm. We used a 130 nm CMOS technology with RRAM co-integrated in between M4 and M5 [15]. The RRAM are individually selected using pass-gates as we do not need current limitation in our proposed scheme.

Fig. 1c presents the operating principle of our writing scheme. As explained in the next section, with a current driver, the parasitic line capacitance C_p plays an important part in defining the dynamics of the system. To have a realistic capacitance value C_p with respect to Mb matrices, the current source must be co-integrated with the RRAM bank. With post-layout extraction, we estimated that C_p was approximately 600 fF in our integrated circuit. With scaling, and at constant bank size, this capacitance will tend to lower. Its value has no effect on the current spike amplitude, but is directly proportional to its energy [18]. This leads us to believe that there might be some need to adapt the scheme for more advanced technology nodes.

Capacitive-Enhanced Current Driver

When a voltage driver is used, the parasitic line capacitance C_p does not play a major role. Other parasitic capacitances play a role in current spiking, such as the one on the node linking the cell to its current limiter [18]. However, when using a current driver, C_p plays an integral part in the writing process. That is why we call the driver in this scheme *Capacitive-Enhanced Current Driver* (CECD).

Fig. 2 presents a comparison between these two driving methods on a set operation. With voltage driving, when the set event occurs, the output current of the driver quickly rises to the compliance. As a result, the current limiter makes the voltage on the RRAM drop, but the current through the device stays at the compliance level until the voltage pulse ends.

In this work, with CECD, both the current passing through and the voltage on the RRAM rise at a limited rate which is proportional to I_{prog}/C_p where I_{prog} is the output current of the driver. When the set occurs, the voltage drops and the current passing through the RRAM spikes as C_p discharges through the RRAM. As long as the switching dynamics of the RRAM is faster than the RC time constant of the parallel combination of C_p and the impedance of the cell (≥ 5 kΩ x 600 fF = 3 ns in this work), the amplitude of the spike does not depend on the value of the capacitance [18]. The capacitance value only affects the total extra energy received. Since the conductance of the LRS is mainly defined by the maximum peak current passing though the RRAM device [11], we can expect the conductance in LRS to be well controlled, independently of the precise value of C_p. In the case of memory matrices, this should make this scheme robust to cell placement.

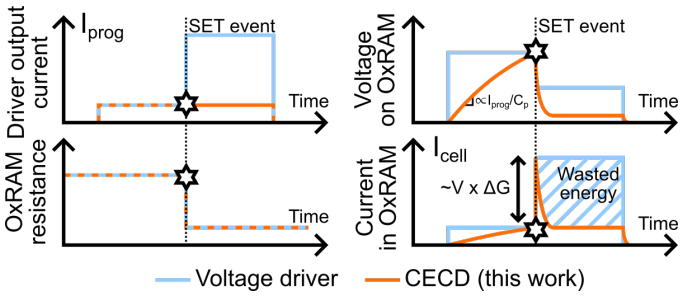

Fig. 2. A comparaison of the time evolution of the voltage and current in RRAM cells with a voltage driver on one side, and CECD on the other.

Thanks to the capacitive current spike, the maximum current passing through the RRAM is higher than I_{prog}. As such, with CECD, the output current of the driver can be significantly lower than the one in a voltage driver, for an equivalent LRS conductance. This reduces the constraints due to IR drop in memory matrices, pushes back the electromigration barrier, and reduces the total energy required per set.

Results and Discussion

Fig. 3a shows the distributions of read currents obtained on a device for different I_{prog}. It shows that the distribution is mostly independent from the current pulse amplitude as long as it is above a given threshold, which in our experiment was found to be below 23.5 µA. In voltage driven schemes, a minimum voltage is required for writing the cells. Here, this limitation translates to a minimum current through the relation linking the voltage, current and impedance of the cell. To obtain the same conductance level with a typical voltage driver, a compliance current of at least 200 µA would be required. The irregularities near the high current tails seem to be linked to LRS instabilities that occur stochastically during cycling. This does not affect our conclusion on the independence of the LRS on I_{prog}, which is very contrasting compared to the importance of the compliance current in voltage schemes.

Fig. 3b shows the distributions of read currents obtained on a device for different pulse duration from 1 µs to 10 µs. Again, the read current distributions are found to be independent of the pulse duration, which is to be expected given that this scheme is naturally self-terminated. Indeed, by comparing the measured

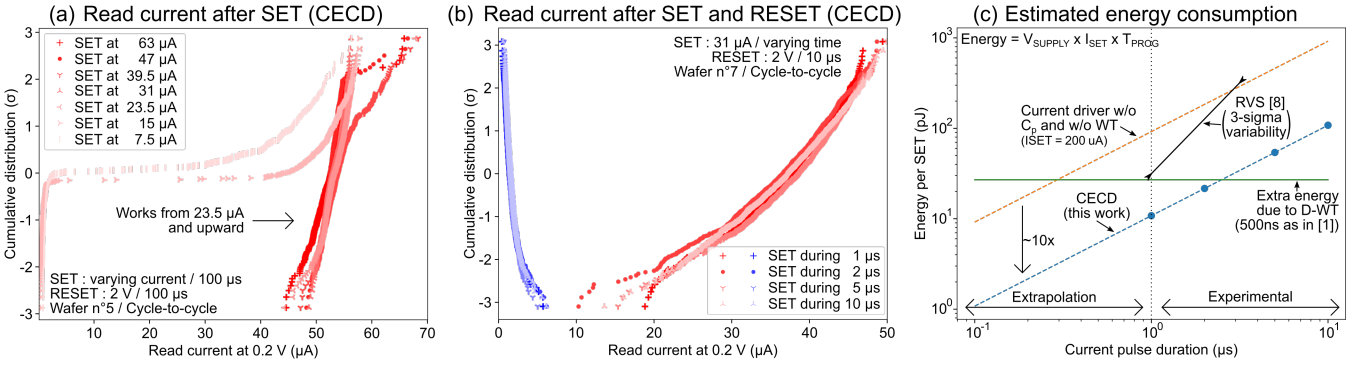

Fig. 3. (a) Cycle-to-cycle distribution of read current for different write pulse current amplitude obtained with CECD. (b) Cycle-to-cycle distribution of read current for different write pulse duration. (c) Estimation of energy consumption of the CECD scheme compared to a current driver scheme ignoring the effect of the parasitic capacitance [4], the extra energy invested in D-WT [7], and Ramp Voltage Stress (RVS) [8].
The difference of spread in the LRS distribution between (a) and (b) is due to the use of two different dies.

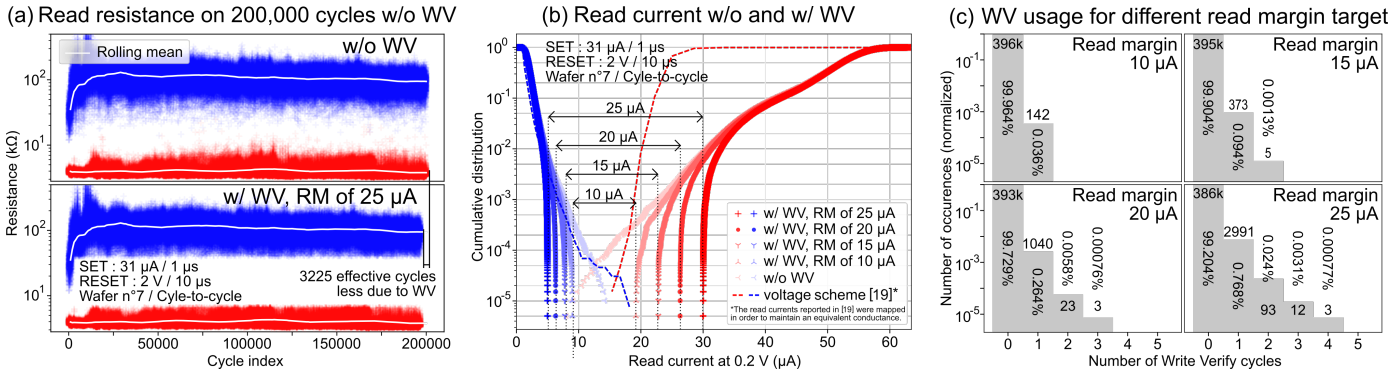

Fig. 4. (a) Scatter plot of the read resistance of 200 000 consecutive SET / READ / RESET / READ cycles with CECD. The bottom plot illustrates the effect of the WV with a RM of 25 µA. (b) Corresponding distribution (excluding the 2 000 seasoning cycles) with and without write verify (WV) and the distribution obtained in [19] with a conventional voltage scheme on the same technology. (c) Histograms of the number of extra cycles required by our WV scheme.

read current at 0.2 V with I_{prog}, we can infer that, after the set occurred, the voltage on the RRAM drops to a value close to 0.2 V. Therefore, from the point of view of the RRAM cell, the excitation roughly ends as soon as the spike from the set event wears off. That explains why the pulse duration has no measurable impact. Furthermore, the low remnant electric field could be beneficial to reduce relaxation as observed with delayed write termination [6], [7].

Such as with the effect of the current pulse amplitude, we expect to hit a threshold in pulse duration under which the set operation would no longer be consistent. We successfully demonstrate that CECD works with pulses as short as 1 µs.

Since the energy consumption is proportional to both the pulse duration and current amplitude, shorter pulses would be beneficial not only for writing speed, but also for energy efficient programming. Fig. 3c shows a comparison of energy consumption between our scheme and [4], [7], [8]. This establishes our scheme as an energy efficient writing technique that does not require any extra write termination circuitry.

Fig. 4a shows the scatter plot of the measured resistance under 0.2 V against the cycle index, in a 200 000 cycles experiment. After a period of seasoning of 2 000 cycles, both distributions remained stable throughout the remainder of the 200 000 cycles. Fig. 4b shows the distribution excluding seasoning. The observed raw bit failure rate is 5×10^{-5}. For comparison, a distribution obtained with a conventional voltage scheme on the same technology is shown [19]. We can see that its LRS distribution is tighter, but achieves about the same raw bit failure rate. The low LRS/HRS ratio of less than one order of magnitude may seem insufficient for high capacity memories. However, by using circuit design strategies, a fully functional 128 kbit array with zero failures up to 1 M cycles was demonstrated on the same technology [19].

To further enhance the reliability, a write verify (WV) algorithm is employed. After each programming operation, if the resulting conductance is not satisfactory, the memory is cycled. This process is repeated until the conductance of the memory is within predefined bounds. Fig. 4c shows the histograms of the number of extra cycles required to meet the targeted read margin (RM). As expected, the larger the RM target, the more cycles are required. However, even for a target of 25 µA, only 4 extra cycles are required, at most, to meet our criteria on the 200 000 cycles of this experiment. In more than 99 % of the cases no extra cycle is necessary, and the maximum number of cycles is only necessary in rare events, less than 10 ppm. By comparison, the same write algorithm with the same rejection rate would only achieve a ~15 µA read margin with the conventional voltage scheme. For instance, a 25 µA read margin would hardly be possible with the voltage scheme. The lower slope of the distribution of LRS in the case of CECD, which could be seen as an inconvenience at first, actually makes WV more efficient. The WV algorithm requires extra cycles and as such consumes energy and ages the cell prematurely. However, since extra cycles are rare, its effect on global energy consumption and aging is marginal.

The low write energy exhibited by this scheme raises a question about retention. This work being the first on this new current scheme concept, we focused on its key and unique features. Therefore, we did not specifically address retention. We can report that we did not experience any issue with short-term retention during this study. We plan to fully address retention in future work.

Given our previous observation that the LRS is independent of the value of I_{prog} above a threshold, the driver in this scheme need not be any more complicated than a fixed current source. As such, the surface required per driver is much lower than any voltage control implementation.

Conclusion

We presented a current driver based scheme for set operation in RRAM. The capacitive-enhanced current driver (CECD) utilizes the inherent parasitic line capacitance to set RRAM cells to a high conductance with limited current coming from the driver. We showed that the distribution of LRS is mostly independent of the pulse amplitude, duration and line capacitance. The scheme is used to program a cell for 200 000 cycles with a raw bit failure rate of 5×10^{-5}. With write verification, it manages a read margin of 25 µA with only 4 extra cycles at worst. Since in more than 99 % of the cases the read margin is attained without requiring any extra cycle, the mean cost in time and energy of the write verification is marginal. The proposed scheme lowers the power consumption since it allows 10 x reduction in the current of the driver at a given LRS conductance compared to conventional drivers. This is a step forward to relax the constraints due to IR drop and electromigration and to promote the RRAM integration.

REFERENCES

[1] C. Peters, F. Adler, K. Hofmann and J. Otterstedt, "Reliability of 28nm embedded RRAM for consumer and industrial products," in *IMW Tech. Dig.*, 2022. pp. 1-3. DOI. 10.1109/IMW52921.2022.9779300.

[2] S. Ito, Y. Hayakawa, Z. Wei, S. Muraoka, K. Kawashima, H. Kotani, K. Kouno, M. Nakamura, G. A. Du, J. F. Chen, S. P. Yeoh, M. H. Chen, T. Mikawa and S. Yoneda, "ReRAM Technologies for Embedded Memory and Further Applications," in *IMW Tech. Dig.*, 2018. pp. 1-4. DOI. 10.1109/IMW.2018.8388846.

[3] Y. Liu, Z. Wang, A. Lee, F. Su, C.-P. Lo, Z. Yuan, C.-C. Lin, Q. Wei, Y. Wang, Y.-C. King, C.-J. Lin, P. Khalili, K.-L. Wang, M.-F. Chang and H. Yang, "4.7 A 65nm ReRAM-enabled nonvolatile processor with 6× reduction in restore time and 4× higher clock frequency using adaptive data retention and self-write-termination nonvolatile logic," in *ISSCC Dig. Tech. Papers*, 2016. pp. 84-86. DOI. 10.1109/ISSCC.2016.7417918.

[4] P. Jain, U. Arslan, M. Sekhar, B. C. Lin, L. Wei, T. Sahu, J. Alzate-vinasco, A. Vangapaty, M. Meterelliyoz, N. Strutt, A. B. Chen, P. Hentges, P. A. Quintero, C. Connor, O. Golonzka and Fis, "13.2 A 3.6Mb 10.1Mb/mm2 Embedded Non-Volatile ReRAM Macro in 22nm FinFET Technology with Adaptive Forming/Set/Reset Schemes Yielding Down to 0.5V with Sensing Time of 5ns at 0.7V," in *ISSCC*, 2019. pp. 212-214. DOI. 10.1109/ISSCC.2019.8662393.

[5] W.-H. Chen, W.-J. Lin, L.-Y. Lai, S. Li, C.-H. Hsu, H.-T. Lin, H.-Y. Lee, J.-W. Su, Y. Xie, S.-S. Sheu and M.-F. Chang, "A 16Mb dual-mode ReRAM macro with sub-14ns computing-in-memory and memory functions enabled by self-write termination scheme," in *IEDM*, 2017. pp. 28.2.1-28.2.4. DOI. 10.1109/IEDM.2017.8268468.

[6] J. Yang, X. Xue, X. Xu, Q. Wang, H. Jiang, J. Yu, D. Dong, F. Zhang, H. Lv and M. Liu, "24.2 A 14nm-FinFET 1Mb Embedded 1T1R RRAM with a 0.022μm2 Cell Size Using Self-Adaptive Delayed Termination and Multi-Cell Reference," in *ISSCC*, 2021. pp. 336-338. DOI. 10.1109/ISSCC42613.2021.9365945.

[7] C.-C. Chou, Z.-J. Lin, P.-L. Tseng, C.-F. Li, C.-Y. Chang, W.-C. Chen, Y.-D. Chih and T.-Y. J. Chang, "An N40 256K×44 embedded RRAM macro with SL-precharge SA and low-voltage current limiter to improve read and write performance," in *ISSCC*, 2018. pp. 478-480. DOI. 10.1109/ISSCC.2018.8310392.

[8] G. Sassine, C. Cagli, J.-F. Nodin, G. Molas and E. Nowak, "Novel Computing Method for Short Programming Time and Low Energy Consumption in HfO2 Based RRAM Arrays," *IEEE Journal of the Electron Devices Society,* vol. 6, pp. 696-702, 2018. DOI. 10.1109/JEDS.2018.2830999.

[9] Y. Feng, P. Huang, Y. Zhao, Y. Shan, Y. Zhang, Z. Zhou, L. Liu, X. Liu and J. Kang, "Improvement of State Stability in Multi-Level Resistive Random-Access Memory (RRAM) Array for Neuromorphic Computing," *IEEE Electron Device Letters,* vol. 42, no. 8, pp. 1168-1171, 2021. DOI. 10.1109/LED.2021.3091995.

[10] D. Ielmini, F. Nardi and S. Balatti, "Evidence for Voltage-Driven Set/Reset Processes in Bipolar Switching RRAM," *IEEE Transactions on Electron Devices,* vol. 59, no. 8, pp. 2049-2056, 2012. DOI. 10.1109/TED.2012.2199497.

[11] D. Ielmini, "Resistive switching memories based on metal oxides: mechanisms, reliability and scaling," *Semiconductor Science and Technology,* vol. 31, no. 6, p. 063002, 2016. DOI. 10.1088/0268-1242/31/6/063002.

[12] M. H. Li, Y. Jiang, V. Y. -Q. Zhuo, E. -G. Yeo, L. -T. Law and K. -G. Lim, "Self-compliance SET switching and multilevel TaOx resistive memory by current-sweep operation," in *14th Annual Non-Volatile Memory Tech. Symp. (NVMTS)*, 2014. pp. 1-3. DOI. 10.1109/NVMTS.2014.7060851.

[13] B. Chen, B. Gao, S. W. Sheng, L. F. Liu, X. Y. Liu, Y. S. Chen, Y. Wang, R. Q. Han, B. Yu and J. F. Kang, "A Novel Operation Scheme for Oxide-Based Resistive-Switching Memory Devices to Achieve Controlled Switching Behaviors," *IEEE Electron Device Letters,* vol. 32, no. 3, pp. 282-284, 2011. DOI. 10.1109/LED.2010.2101577.

[14] W. Lian, H. Lv, Q. Liu, S. Long, W. Wang, Y. Wang, Y. Li, S. Zhang, Y. Dai, J. Chen and M. Liu, "Improved Resistive Switching Uniformity in Cu/HfO2/Pt Devices by Using Current Sweeping Mode," *IEEE Electron Device Letters,* vol. 32, no. 8, pp. 1053-1055, 2011. DOI. 10.1109/LED.2011.2157990.

[15] J. Sandrini, L. Grenouillet, V. Meli, N. Castellani, I. Hammad, S. Bernasconi, F. Aussenac, S. Van Duijn, G. Audoit, M. Barlas, J. F. Nodin, O. Billoint, G. Molas, R. Fournel, E. Nowak, F. Gaillard and C. Cagli, "OxRAM for embedded solutions on advanced node," in *IEDM*, 2019. pp. 30.5.1-30.5.4. DOI. 10.1109/IEDM19573.2019.8993484.

[16] 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 *IEDM*, 2013. pp. 10.1.1-10.1.4. DOI. 10.1109/IEDM.2013.6724598.

[17] C. Y. Chen, A. Fantini, L. Goux, R. Degraeve, S. Clima, A. Redolfi, G. Groeseneken and M. Jurczak, "Programming-conditions solutions towards suppression of retention tails of scaled oxide-based RRAM," in *IEDM*, 2015. pp. 10.6.1-10.6.4. DOI. 10.1109/IEDM.2015.7409671.

[18] P. R. Shrestha, D. M. Nminibapiel, J. P. Campbell, J. T. Ryan, D. Veksler, H. Baumgart and K. P. Cheung, "Analysis and Control of RRAM Overshoot Current," *IEEE Transactions on Electron Devices,* vol. 65, no. 1, pp. 108-114, 2018. DOI. 10.1109/TED.2017.2776860.

[19] B. Giraud, S. Ricavy, Y. Moursy, C. Laffond, I. Sever, V. Gherman, M. Pezzin, F. Lepin, M. Diallo, K. Zenati, S. Dumas, M. Vershkov, A. Bricalli, G. Piccolboni, J.-P. Noel, A. Samir, G. Pillonnet, Y. Thonnart and G. Molas, "Benefits of Design Assist Techniques on Performances and Reliability of a RRAM Macro," *IEEE International Memory Workshop (IMW),* in press, 2023.