Reliability of Logic-in-Memory Circuits in Resistive Memory Arrays

Tommaso Zanotti, *Student Member, IEEE*, Cristian Zambelli, *Member, IEEE*, Francesco Maria Puglisi, *Member, IEEE*, Valerio Milo, *Member, IEEE*, Eduardo Pérez, Mamathamba K. Mahadevaiah, Oscar G. Ossorio, Christian Wenger, Paolo Pavan, *Senior Member, IEEE*, Piero Olivo, and Daniele Ielmini, *Fellow, IEEE*

***Abstract*— Logic-in-Memory (LiM) circuits based on RRAM devices and the material implication logic are promising candidates for the development of low-power computing devices, that could fulfill the growing demand of distributed computing systems. However, these circuits are affected by many reliability challenges that arise from device non-idealities (e.g., variability) and the characteristics of the employed circuit architecture. Thus, an accurate investigation of the variability at the array level is needed to evaluate the reliability and performance of such circuit architectures. In this work, we explore the reliability and performance of SIMPLY (i.e., a recently proposed LiM architecture with improved reliability and performance) on two 4 kbits RRAM arrays based on different resistive switching oxides integrated in the BEOL of the 0.25 μm BiCMOS process. We analyze the trade-off between reliability and energy consumption of SIMPLY architecture by exploiting the results of an extensive array-level variability characterization of the two technologies. Finally, we study the worst-case performance of a full adder implemented with the SIMPLY architecture and benchmark it on the analogous CMOS implementation.**

***Index Terms*— RRAM, BEOL, SIMPLY, Logic-in-Memory, Full adder.**

I. Introduction

Today, there are roughly 17 billion devices at the edge, which causes massive and ever-growing data exchange over communication networks. In this context, edge computing has been identified as a promising solution to relax data transfer and energy consumption limitations, providing advantages for Internet of Things (IoT) applications, smart cities and smart industries, Artificial Intelligence (AI), 5G/6G communications. However, todays ultra-low power hardware solutions are still affected by the von Neumann bottleneck (VNB) [1]–[3]. Specifically, VNB is the time- and energy-demanding process of data transfer between CPU and memory chips, and is the main showstopper for edge computing solutions. As recently suggested in [2], [4], [5], Logic-in-Memory (LiM) circuits that merge together data storage and computation could bypass VNB, thus minimizing the energy and time needed to execute logic functions. Among the most promising solutions, LiM circuits based on resistive random access memory (RRAM) and on the material implication logic (IMPLY) offer significant advantages by leveraging on the small footprint of RRAMs, on their BEOL integration potential, and on the fact that implication logic is complete, i.e., all possible logic functions can be defined by a sequence of few core operations [2], [4]–[6], namely IMPLY and FALSE operations. However, the reliability of such operations and of the material implication logic circuit tightly depends on the non-idealities of the devices, especially variability [5], [6], and on the characteristics of the employed circuit architecture [5], [6]. Thus, evaluating the benefit of introducing the LiM paradigm in edge computing requires an accurate investigation of the variability at the array level. Yet, a clear array-level analysis and demonstration of functionality of RRAM-based LiM solutions is still missing. In this work, we study the performance and feasibility of a recently proposed smart IMPLY (SIMPLY) [2], [5] LiM paradigm on two 4 kbits RRAM arrays with different resistive switching oxides integrated in the BEOL of the 0.25 μm BiCMOS process. Previous works [2], [5], restricted the evaluation of the performance of SIMPLY to RRAM technologies taken from the literature for which only little information regarding cycle-to-cycle (C2C) and device-to-device (D2D) variability is available. Here we exploit the extensive array-level variability characterization of the two RRAM technologies to study the performance of the SIMPLY architecture by evaluating the trade-off between reliability and energy consumption. In addition, we estimate the worst-case energy consumption of a 1-bit full adder (FA) implemented in the SIMPLY architecture, and benchmark it against CMOS implementations. Results show that the SIMPLY implementation of the 1-bit FA outperforms the CMOS one by more than two orders of magnitude, when the VNB is considered, with significant improvement margins left.

This work was supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 648635) and in part by the German Research Foundation (DFG) in the frame of research group FOR2093.

T. Zanotti, F.M. Puglisi and P. Pavan are with Dipartimento di Ingegneria "Enzo Ferrari", Universitá di Modena e Reggio Emilia, 41125 Modena, Italy. (e-mail: tommaso.zanotti@unimore.it)

C. Zambelli and P. Olivo are with Dipartimento di Ingegneria, Universitá degli Studi di Ferrara, 44122 Ferrara, Italy.

V. Milo and D. Ielmini are with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano and IU.NET, 20133 Milano, Italy.

E. Perez, M. K. Mahadevaiah, and Ch. Wenger are with IHP-Leibniz-Institut für innovative Mikroelektronik, 15236 Frankfurt (Oder), Germany.

Ch. Wenger is also with BTU Cottbus-Senftenberg, 01968 Cottbus, Germany.

O. G. Ossorio is with Dpto. Electricidad y Electronica, Universidad de Valladolid, 47011 Valladolid, Spain.

II. SIMPLY Logic-in-Memory architecture

The revived interest in RRAM technology arises from the possibility of storing and manipulating the information in the same place, by realizing LiM paradigms in which information is not stored as voltage at circuit nodes (like in CMOS logic) but as the resistance value of RRAM devices (HRS = logic 0 and LRS = logic 1). Specifically, the paradigm based on the material implication logic is among the most effective since it is "complete", thus all the possible logic operations can be implemented as a sequence of two operations, namely the FALSE (i.e., the reset of a single device) and the IMPLY. In the typical arrangement, the IMPLY operation is executed by simultaneously pulsing two devices (P and Q, holding the input bits) with two different voltages (labeled V_{SET} and $V_{COND)}$ in such a way that P holds its state and Q changes state according to the truth table in Fig.1a. However, as thoroughly analyzed in [4], [6], this arrangement suffers from several issues, such as high energy consumption, the degradation of the logic states stored in the devices, and a strong sensitivity to driving voltage variations [6]. Recently, the SIMPLY architecture was introduced to overcome the aforementioned issues [2], [5]. The SIMPLY architecture is sketched in Fig.1b, and can be easily implemented in a 1T-1R array by shunting together the bottom electrodes of a group of RRAM devices. The

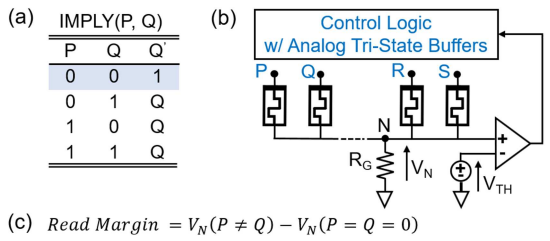

Fig. 1. (a) Truth table of the P IMPLY Q operation. The blue rectangle highlights that the state of Q changes (Q') only when the input combination is P=Q=0. (b) Schematic of SIMPLY architecture on a RRAM array. (c) Read margin (RM) definition considering ideal devices.

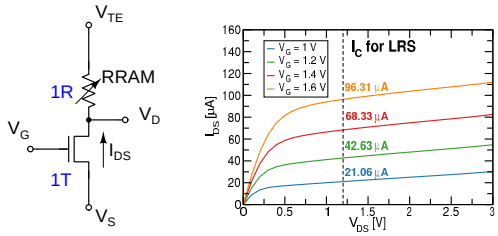

Fig. 2. 1T-1R cell's architecture (left) and I_{DS}-V_{DS} characteristics of the transistor exploited for I_C extraction at different V_G (right).

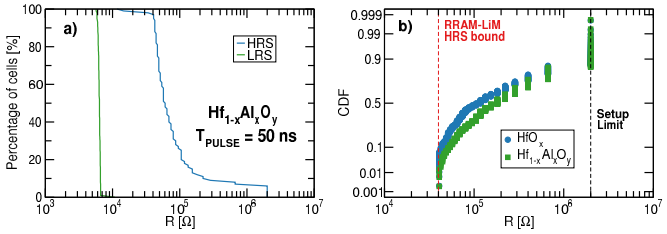

Fig. 3. (a) Demonstration of the functionality of an RRAM array based on $Hf_{1-x}Al_xO_y$ when t_P = 50 ns. Similar results can be obtained for HfO_x arrays. Adapted from [10]. (b) CDF of the HRS extracted from the 4 kb arrays evidencing the bound for LiM application at 40 kΩ.

series transistor is appropriately biased to act as the resistor R_G. The IMPLY operation is performed by: *i)* applying a small V_{read} voltage pulse (200 mV in this work) to both P and Q [2], [5]; *ii)* comparing the voltage at node N (V_N in Fig.1b) against a threshold (V_{TH}) to determine if P=Q=0; *iii)* pulse V_{SET} on Q keeping the driver of P at high impedance only if P=Q=0. In principle, the condition P=Q=0 is easy to detect since V_N is lower in this case than in all other cases, ensuring a sufficient read margin (RM), defined as in Fig.1c. When considering ideal devices, RM is a deterministic quantity dependent on the memory window (i.e., the ratio of HRS to LRS resistance), V_{read}, and R_G. However, the combined effect of D2D and C2C variability, Random Telegraph Noise (RTN), driving voltage variations, and process tolerances results in a relatively wide distribution of RM, potentially impairing the circuit functionality. Therefore, the circuit reliability is tightly coupled to the intrinsic variability of the RRAM technology exploited in its integration, and its statistical characterization allows verifying the reliability level that can be achieved by the proposed LiM circuit when implemented in the RRAM technologies under study.

III. Variability Characterization on RRAM Arrays

To statistically assess both the D2D and the C2C variability, we performed electrical characterization measurements on the 4 kbits 1T-1R arrays whose architecture is described in [7]. We remind that the array is based on the 0.25 μm BiCMOS process from IHP and that the select transistor in the 1T-1R cells is an n-MOS with W = 1.14 μm and L = 0.24 μm. The transistor allows modulating the compliance current I_C by tuning V_G during operations, thus enabling a tight control of the cell conductance and enhanced power-control features in LiM circuits. Fig. 2 shows the I_{DS}-V_{DS} characteristics of a transistor in the array exposing the different I_C. The memristive element is integrated during BEOL on top of the second metal level (M2 - in series with drain) featuring a 600 × 600 nm^2 area. To provide an exploration of the RRAM technology impact on the LiM circuit, we considered two different memristive stacks integrated in separated arrays, namely a TiN/Ti/HfO_x/TiN and a TiN/Ti/$Hf_{1-x}Al_xO_y$/TiN structure. Their process characteristics can be retrieved in [8].

Resistive switching of the memristive element is enabled for all the cells in the array through a Forming operation which consists of the application of the ISPVA algorithm for best yield [9] with duration t_p = 1 μs and a top electrode voltage V_{TE} from 2 V to 5 V in steps of 10 mV. Reset operation is then performed to reach HRS by using a single pulse gate voltage V_G = 2.8 V to minimize transistor resistance and a source voltage V_S = 1.8 V. Same pulse duration is considered. We would like to point that such t_p is chosen to simplify the requirements of the measurement setup, although the functionality of the two RRAM technologies was also proven with t_p = 50 ns [10] (see Fig. 3a). The result of the Reset operation on the 4 kbits array allows extracting the HRS bound dictated by the chosen RRAM technology for the LiM target application. As shown in Fig. 3b, we consider 40 kΩ for both memristive stacks. It is worth to mention that a higher HRS resistance results in a lower power consumption of LiM circuit, so we consider the former value as a worst-case condition that allows speculating on the performance and reliability limits. Concerning the Set operation, we used a V_{TE} = 1.2 V and a single pulse duration t_p = 1 μs associated with four different V_G values from 1 V to 1.6 V in 200 mV steps. This allows a tuning of the LRS on four levels (L_1 to L_4) devising the I_C set by the transistor, while providing a strategy for power consumption reduction policies to be applied on LiM circuits. The LRS read currents (I_{read}) measured with a V_{TE} = 200 mV for L_1 - L_4 correspond approximately to 10 μA, 20 μA, 30 μA, and 40 μA.

Different LRS levels come with different variability characteristics. Fig. 4 shows a characterization study of the variability for levels L_1 - L_4 of both RRAM technologies. In the figure, the standard deviation σ_R for the C2C and D2D distributions is plotted as a function of the median device resistance indicated as R. Variability data were collected for a subset of 1024 1T-1R devices (i.e., a block of 16 wordlines set with the same LRS level) integrated in the 4 kbits test vehicles and for 1000 consecutive Set/Reset cycles. The cycling routine is performed considering the proper V_G for the Set operation in each subset. As it can be seen, the C2C variability dominates over the D2D for all LRS levels following the universal trend for σ_R which is proportional to R^2, as in [11], [12]. However, the D2D variability is the one with the highest σ_R absolute values, thus being critical for the performance of the LiM circuits like those in this work. Overall, $Hf_{1-x}Al_xO_y$ material displays a better control of the D2D and C2C variability as demonstrated by the lower scatter of the $< \sigma_R; R >$ points in the plots. LRS level L_1 has however a higher C2C σ_R compared to that of HfO_x.

IV. Reliability of SIMPLY in RRAM Arrays

The results of the variability characterization in Section III are exploited to verify the reliability and the energy efficiency of the

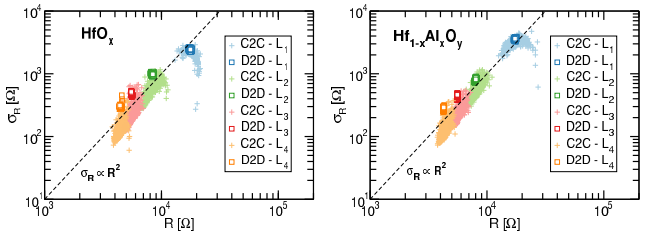

Fig. 4. Standard deviation σ_R of the resistance as a function of the median resistance R for HfO$_x$ (left) and Hf$_{1-x}$Al$_x$O$_y$ RRAM technologies (right), respectively.

SIMPLY architecture. Note that: *i)* the only requirement for a reliable circuit operation of the SIMPLY architecture is that a sufficient RM is available at the input of the comparator in all cases, even in the presence of variability; *ii)* the variability characterization in Section III is performed at the array level, natively including contributions from the device (C2C, D2D and RTN) and from the non-idealities of the peripheral circuits (e.g., possible variations of V_{read}). This makes RM a comprehensive metric of the circuit reliability. Indeed, circuit implementations that result in higher RMs are more reliable and allow using simpler comparator or sense amplifier designs, though more complex designs can be used for smaller RMs. To estimate the RM and compare the performance of the two RRAM technologies, we used the C2C and D2D joint variability data to compute the distributions of V_N when P=Q=0 and when $P \neq Q$ for both the technologies and each LRS level (see Fig.5). The joint probability distribution of LRS for each technology and I_{read} was estimated by combining together 100 random samples for each $< \sigma_R; R >$ pair of Fig.4. To identify the worst-case RM that allows evaluating the performance and reliability limits, we assume HRS fixed at the worst-case of the HRS distribution (R_{HRS} = 40 kΩ for both technologies, Fig.3) and thus $V_N(P$=Q=0) is constant for each technology and LRS level and determined by the value of R_G (see Fig.1a). The optimal R_G value maximizing RM for each technology and LRS level was chosen as:

$$R_G = \sqrt{\left(R_{LRS,MAX}^{-1} + R_{HRS,MAX}^{-1}\right)^{-1} \cdot \frac{R_{HRS,MIN}}{2}} \quad (1)$$

where $R_{LRS,MAX}$ is the μ+3σ value of the joint LRS distribution, and $R_{HRS,MIN}$ and $R_{HRS,MAX}$ are the $\mu \pm 3\sigma$ values of the HRS distribution, respectively ($R_{HRS,MIN}$ = $R_{HRS,MAX}$ = 40 kΩ in this case). As shown in Figs. 5 and 6a, RM grows with I_{read} and the two RRAM technologies show comparable reliability at the same LRS level, except at I_{read} = 10 μA where HfO$_x$ devices guarantee a quite larger RM than Hf$_{1-x}$Al$_x$O$_y$ devices. This stems from the lower C2C variability of HfO$_x$ devices at low I_{read}, which leads to a smaller tail of L_1 as shown in Fig.4. To compare the reliability metric in Fig.6a to the energy efficiency performance, we report in Fig.6b the worst-case energy consumption of the read step of SIMPLY operation (averaged over the four possible input configurations) at different I_{read}. As expected, the energy per operation increases with I_{read}, establishing a trade-off between energy efficiency and reliability. No relevant differences between the two technologies are observed.

V. LIM FULL ADDER IMPLEMENTATION

To benchmark the performance of the SIMPLY architecture on more complex operations, we designed a 1-bit ripple carry full adder (FA) with the two RRAM technologies and estimated the worst-case performance. The FA was implemented with 8 RRAM devices performing the sequence of operations reported in [2], which includes 28 steps of core operations (18 IMPLY and 10 FALSE). During a FA cycle, the energy consumption will also depend on the configuration of input bits, since it dictates the number of reset (during FALSE), set (during IMPLY when P=Q=0), and read (during IMPLY in all cases) operations that are executed. Here we consider the worst-case energy consumption during set and reset operations $E = V_P \cdot I_C \cdot t_P$, where V_P and t_P are the applied voltage magnitude and width. We consider the worst-case energy for each input combination also for the read operation. The minimum and maximum worst-case energies per FA cycle are proportional to I_C (see Table I). Values on the order of nJ are obtained for a t_P of 1 μs as the one used in Section III ($t_{IMPLY} = 4 \cdot t_P$, $t_{FALSE} = 2 \cdot t_P$, $t_{FA} = 10 \cdot t_{FALSE} + 18 \cdot t_{IMPLY}$). These estimates include the comparator energy overhead. For the comparator we used the sense amplifier design from [2]. With a power supply voltage of 1.6 V the comparator dissipates around 70 fJ/comparison when t_P is 1 μs, which is far less than the worst-case energy per read operation,

TABLE I
1-BIT FULL ADDER ENERGY PER OPERATION (INDICATING MIN - MAX RANGE) VS I_{read} WITH t_p = 1 μs

I_{read}	Read	Write	FA
10 μA	(2.4 - 2.6) · 10^{-11} J	(5.1 - 5.6) · 10^{-10} J	(5.3 - 5.8) · 10^{-10} J
20 μA	(3.2 - 3.5) · 10^{-11} J	(1.0 - 1.1) · 10^{-9} J	(1.1 - 1.2) · 10^{-9} J
30 μA	(3.9 - 4.2) · 10^{-11} J	(1.7 - 1.8) · 10^{-9} J	(1.7 - 1.9) · 10^{-9} J
40 μA	(4.4 - 4.8) · 10^{-11} J	(2.3 - 2.5) · 10^{-9} J	(2.4 - 2.6) · 10^{-9} J

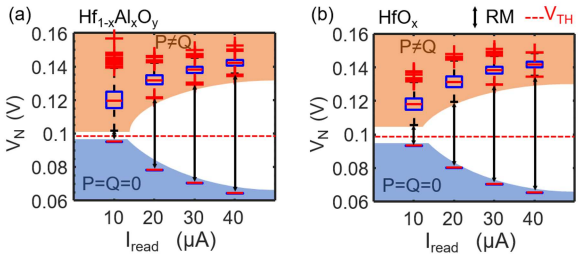

Fig. 5. (a)-(b) Distribution of V_N due to C2C and D2D variability, when P=Q=0 (blue area) and $P \neq Q$ (orange area) for the SIMPLY operation for different I_{read} and RRAM technology. Only the worst-case (i.e., lowest resistance value due to variability) HRS resistance is considered. The read margins (RM black arrows) and the threshold voltages (V_{TH} dashed red) for the comparator are evidenced. Black whiskers indicate the extreme points of the distributions. Red crosses indicate outliers. V_N when P=Q=1 is always much higher than in all other cases (thus is not reported in these box plots).

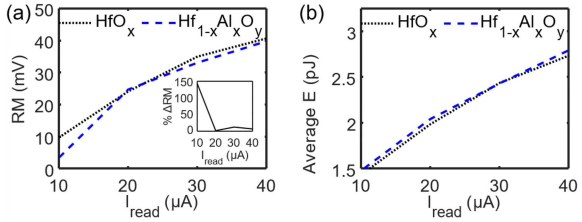

Fig. 6. (a) Worst-case RM (-3σ deviation from the mean of the RM distribution due to variability) for the two different RRAM technologies and different I_{read} values. The inset shows the relative difference between the RM obtained with HfO$_x$ and Hf$_{1-x}$Al$_x$O$_y$ devices. (b) Worst-case energy consumption of the read operation during SIMPLY operation for HfO$_x$ and Hf$_{1-x}$Al$_x$O$_y$ RRAM technologies at different I_{read} values. It is computed as the average of the worst-case energy consumption (+3σ deviation from the mean of the energy distribution due to variability) of each input configuration.

TABLE II
COMPARISON BETWEEN THE PROPOSED FA AND A CMOS FA WHEN EXECUTING 32 PARALLEL 32-BIT FA OPERATIONS (ON A 4 KB ARRAY)

	# Devices	Energy	Delay	EDP	Norm. EDP[$]	EDP Improvement[$]
CMOS w/ VNB*	8192 - 28672 FET	$\approx$ 9.4 μJ	$\approx$ 284 μs	$\approx 2.7 \cdot 10^{-9}$ J·s	1	1
CMOS w/o VNB**	8192 - 28672 FET	$\approx 9.7 \cdot 10^{-4}$ - 7.4 pJ	$\approx 5.6 \cdot 10^{-2}$ - 4.8 μs	$\approx 1.7 \cdot 10^{-24}$ - $3.6 \cdot 10^{-17}$ J·s	$6.3 \cdot 10^{-16}$ - $1.3 \cdot 10^{-8}$	$7.5 \cdot 10^{7}$ - $1.6 \cdot 10^{15}$
This work* t_P = 1 μs**	**3232 RRAM**	**$\approx$ 594 nJ**	**$\approx$ 2.9 ms**	**$\approx 1.7 \cdot 10^{-9}$ J·s**	**0.6**	**1.57**
This work* t_P = 50 ns**	**3232 RRAM**	**$\approx$ 30 nJ**	**$\approx$ 147 μs**	**$\approx 4.4 \cdot 10^{-12}$ J·s**	**$1.6 \cdot 10^{-3}$**	**607**

*,** estimates with (w/) and without (w/o) VNB are performed considering energy and delay overhead for reading 2 kbits of data from a NAND flash memory with a 4kB memory page size [13]. CMOS FA performances were estimated projecting the time and energies for different 1-bit FA schemes taken from [3], [14], [15] where 0.18 μm, 45 nm, and 10 nm CMOS technology are used. *** I_{read} = 10 μA [$]w.r.t. to CMOS w/ VNB.

shown in Fig.6b. To show the advantages offered by the proposed LiM scheme in terms of energy efficiency, we compare the performance of the proposed architecture against CMOS FA implementations from [3], [14], [15] with and without considering the VNB [13]). The VNB overhead was computed considering the latency and energy required to read data from a flash memory considering a typical page size of 4kB [13]. We consider flash memory technology as it is currently the state of the art for non-volatile memories. We estimate the delay and energy required to compute 32 parallel 32-bits FA operations (simple ripple carry) which require slightly less (3132) than the available 4096 devices in the array. To show the potential energy efficiency improvement over CMOS we consider the case I_{read} = 10 μA. For both SIMPLY and CMOS implementations, the peripheral circuitry needed to decode instructions is comparable, and thus its overhead is neglected in both cases. As shown in Table II, the largest share of energy consumption and delay for CMOS logic comes from the VNB data exchange overhead [3], [13]–[15]. With the technologies explored in this work, when t_P = 1 μs (i.e., the pulse duration used in the characterization in Section III) the energy delay product (EDP) of SIMPLY is only slightly better than its CMOS counterpart when the VNB effect is included. However, the functionality of the RRAM devices used in this work was proven also with t_P = 50 ns (see Fig.3a), and energy projections with such a t_P show that the proposed LiM scheme outperforms CMOS (when including the VNB overhead) by more than two orders of magnitude (worst-case projection) in energy efficiency with similar computing time, as shown in Table II. In addition, the number of required devices is reduced as compared to CMOS, thus achieving higher integration density.

VI. CONCLUSIONS

In this work, we studied the reliability and performance of SIMPLY architecture integrated on two different 4 kbits RRAM arrays. We highlighted and evaluated the trade-off between circuit reliability and energy consumption by exploiting the extensive array-level variability characterization of the two technologies. Furthermore, we analyzed the performance of a 1-bit FA implemented on SIMPLY. Even when considering the worst-case, the proposed architecture is $\approx$ 600 times more efficient than the CMOS counterpart including the VNB, and achieves higher integration density. These results suggest that the proposed approach is a solution for the development of ultra-low power computing.

REFERENCES

[1] J. Backus, "Can Programming Be Liberated from the Von Neumann Style?: A Functional Style and Its Algebra of Programs," *Commun. ACM*, vol. 21, no. 8, pp. 613–641, 1978, doi: 10.1145/359576.359579.

[2] T. Zanotti, F. M. Puglisi, and P. Pavan, "A Smart Logic-in-Memory Architecture for Low-Power non-von Neumann Computing," *IEEE Journal of the Electron Devices Society*, 2020, (in press), doi: 10.1109/JEDS.2020.2987402.

[3] M. Aguirre-Hernandez and M. Linares-Aranda, "CMOS Full-Adders for Energy-Efficient Arithmetic Applications," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 4, pp. 718–721, 2011, doi: 10.1109/TVLSI.2009.2038166.

[4] S. Kvatinsky, G. Satat, N. Wald, E. G. Friedman, A. Kolodny, and U. C. Weiser, "Memristor-Based Material Implication (IMPLY) Logic: Design Principles and Methodologies," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 10, pp. 2054–2066, 2014, doi: 10.1109/TVLSI.2013.2282132.

[5] F. M. Puglisi, T. Zanotti, and P. Pavan, "SIMPLY: Design of a RRAM-Based Smart Logic-in-Memory Architecture using RRAM Compact Model," in *ESSDERC*, 2019, pp. 130–133, doi: 10.1109/ESSDERC.2019.8901731.

[6] T. Zanotti, F. M. Puglisi, and P. Pavan, "Reliability-aware design strategies for stateful logic-in-memory architectures," *IEEE Trans. on Device and Materials Reliability*, vol. 20, no. 2, pp. 278–285, 2020, doi: 10.1109/TDMR.2020.2981205.

[7] A. Grossi, D. Walczyk, C. Zambelli, E. Miranda, P. Olivo, V. Stikanov, A. Feriani, J. Su, 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 Trans. on Electron Devices*, vol. 62, no. 8, pp. 2502–2509, 2015, doi: 10.1109/TED.2015.2442412.

[8] V. Milo, C. Zambelli, P. Olivo, E. Perez, M. K. Mahadevaiah, O. G. Ossorio, C. Wenger, and D. Ielmini, "Multilevel HfO_2-based RRAM devices for low-power neuromorphic networks," *APL Materials*, vol. 7, no. 8, p. 081120, 2019, doi: 10.1063/1.5108650.

[9] 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, pp. 17 – 25, 2016, doi: 10.1016/j.sse.2015.10.003.

[10] E. Perez, O. Gonzlez Ossorio, S. Duenas, H. Castan, H. Garcia, and C. Wenger, "Programming Pulse Width Assessment for Reliable and Low-Energy Endurance Performance in Al:HfO_2-Based RRAM Arrays," *Electronics*, vol. 9, p. 864, 2020, doi: 10.3390/electronics9050864.

[11] A. Fantini, L. Goux, R. Degraeve, D. J. Wouters, N. Raghavan, G. Kar, A. Belmonte, Y. . Chen, B. Govoreanu, and M. Jurczak, "Intrinsic switching variability in HfO_2 RRAM," in *IEEE IMW*, 2013, pp. 30–33, doi: 10.1109/IMW.2013.6582090.

[12] 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, doi: 10.1109/TED.2014.2330200.

[13] S.-Y. Park, D. Jung, J.-U. Kang, J.-S. Kim, and J. Lee, "CFLRU: A Replacement Algorithm for Flash Memory," in *Proc. of the 2006 Int. Conf. on Compilers, Architecture and Synthesis for Embedded Systems*. ACM, 2006, pp. 234–241, doi: 10.1145/1176760.1176789.

[14] A. K. Yadav, B. P. Shrivatava, and A. K. Dadoriya, "Low power high speed 1-bit full adder circuit design at 45nm CMOS technology," in *2017 Int. Conf. on Recent Innovations in Signal processing and Embedded Systems (RISE)*, 2017, pp. 427–432, doi. 10.1109/RISE.2017.8378203.

[15] S. Sharma and G. Soni, "Comparision analysis of FinFET based 1-bit full adder cell implemented using different logic styles at 10, 22 and 32NM," in *2016 Int. Conf. on Energy Efficient Technologies for Sustainability (ICEETS)*, 2016, pp. 660–667, doi. 10.1109/ICEETS.2016.7583835.