Phase-change memory cells characterization in an analog in-memory computing perspective

Phase-change memory cells characterization in an analog in-memory computing perspective

Alessio Antolini[1,2*], Andrea Lico[1,2], Eleonora Franchi Scarselli[1,2],
Marcella Carissimi[3], Marco Pasotti[3]
[1]DEI – University of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy.
[2]ARCES – University of Bologna, Via Vincenzo Toffano 2/2, 40125 Bologna, Italy.
[3]STMicroelectronics, Via Camillo Olivetti 2, 20864 Agrate Brianza, Italy.
*Correspondence: alessio.antolini2@unibo.it

***Abstract*— Power consumption related to data transfers between processing and memory units has become a critical issue in the recent data-centric outlook of integrated circuits. In the context of In-memory Computing (IMC), where data conveyance is narrowed performing computations directly within the memory unit, Phase-change Memory (PCM) technology has become an attractive candidate due to its intrinsic multilevel storage capability. The test vehicle of this work is an embedded PCM (ePCM) provided by STMicroelectronics and designed in 90-nm smart power BCD technology with a Ge-rich Ge-Sb-Te (GST) alloy for automotive applications. In this framework, a preliminary characterization of PCM cells has been carried out, aimed at evaluating their performance as enabling devices for analog in-memory computing (AIMC) applications.**

Keywords—In-memory computing (IMC), Phase-change memory (PCM), programming algorithm, analog computing.

I. Introduction

The rising spread of data-centric applications, such as deep learning and artificial intelligence, has dictated new constraints in the field of computing architectures and algorithms. Among innovative solutions recently being researched, In-memory Computing (IMC) has been proposed as a valid strategy to overcome traditional Von Neumann architectures, where physically separated processing and memory units implicate a massive amount of both energy consumption and computing latency due to internal data transfers. To address this issue, IMC strategy intends the execution of computing tasks directly within the memory array.

In this context, phase-change memory (PCM) technology, a valid candidate among non-volatile memories for both stand-alone and embedded applications [1], has been applied to analog computing [2], artificial neural networks [3], [4] and spiking neural networks [5], [6]. In particular, the Analog In-memory Computing (AIMC) scenario aims all the computations to be carried out in an analog way, avoiding digital data to be conveyed between memory and processing units [3]. In all these applications, multilevel storage is an attractive peculiarity, as it allows the implementation of analog sums and multiplications simply exploiting Ohm's and Kirchhoff's laws [1], where data are stored in PCM cells in terms of conductance. PCM relies on the reversible transition of a chalcogenide material between its crystalline (or SET) and amorphous (or RESET) state. The amorphous phase tends to have high electrical resistivity, while the crystalline phase exhibits a low resistivity, several orders of magnitude lower. The programming process is performed through appropriate current pulses, named SET and RESET, whose application turns into a conductance increase or decrease, respectively.

However, from a practical viewpoint, several problems afflict this picture, due to the very nature of PCM cells [1], [5], [7], [8]. First, low-frequency noise affects cells behavior, as random electron traps are located in the cell lattice. Moreover, cell conductance tends to drift in time due to relaxation phenomena of the crystal lattice. Time drift implies a conductance decrease which differs from cell to cell, thus increasing the dispersion of stored conductance levels. These challenges lead to dispersion and inaccuracy of the stored conductance levels. In this work, a characterization in a AIMC perspective of embedded PCM (ePCM) cells test in 90-nm smart power BCD technology is proposed, aimed at evaluating their performances in terms of drift, dispersion and noise, combining the effect of temperature as well [9]. The multilevel feature of cells has been achieved through the exploitation of an iterative dedicated programming algorithm. This paper is organized as follows: in Section II the experimental setup is shown; in Section III the characterization results are discussed; the conclusion presents final assessments and proposes next research developments.

II. Experimental setup

A. Test vehicle and evaluation board

This experimental activity has been performed on an embedded PCM (ePCM) test chip designed and manufactured by STMicroelectronics [10] in 90-nm smart power BCD technology featuring a specifically optimized Ge-rich Ge-Sb-

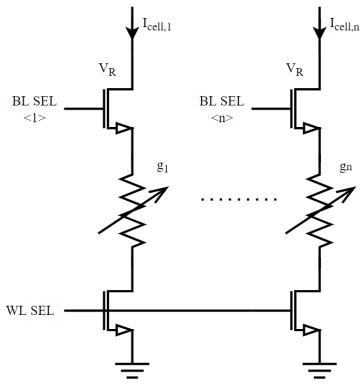

Fig. 1. Sketch of the test vehicle elementary cells. Each memory element is addressable with a word line selector (WL SEL) and a bit line selector (BL SEL). A voltage V_R can be applied to the BL node to read the cell currents.

This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 101007321. The JU receives support from the European Union's Horizon 2020 research and innovation programme and France, Belgium, Czech Republic, Germany, Italy, Sweden, Switzerland, Turkey.

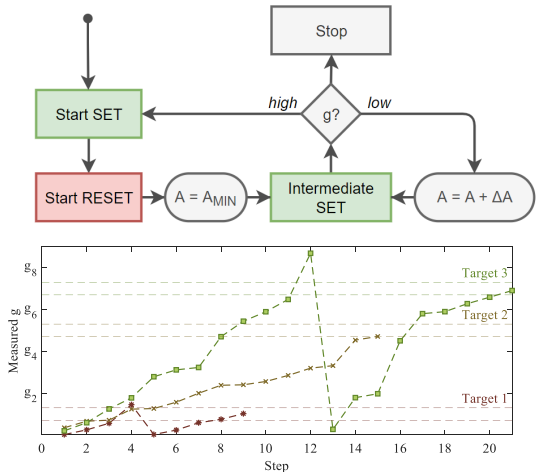

Fig. 2. Top: employed programming algorithm of PCM cells. Bottom: typical evolution of the normalized conductance g of three sample cells during their programming sequence steps; normalized target conductances are g_1, g_5 and g_7, with a normalized tolerance of 0.3.

Te (GST) alloy. The ePCM elementary cell, sketched in Fig. 1, is based on an NMOS selector and occupies 0.19 μm^2 of silicon area [11]. A PCM evaluation board has been employed and customized to allow the configuration of current pulses applied to cells. Every programming or measurement process is achieved with a graphic user interface (GUI), which is available on a personal computer. To perform the generic cell conductance measure, cell current has been measured with a precise source-meter unit (SMU) with a read voltage V_R = 100 mV.

B. *Experimental procedure*

In this work, a number of n_c = 500 memory cells have been iteratively programmed for seven different conductances to examine their behaviors in relation to their programmed level and to time. The selection of these values falls between a minimum G_{MIN} and a maximum G_{MAX} achievable conductance, which are reached by a fully RESET-state and a fully SET-state cell, respectively. G_{MIN} can be considered equal to zero; therefore, seven intermediate target conductances $\hat{G}$ have been chosen, defining accordingly seven different normalized conductances targets $\hat{g} = \hat{G}/G_{MAX}$ equally distributed between 0 and G_{MAX}. Among the state of the art, several strategies to program cells to a specific conductance are proposed [12]–[14]. In this study, the programming algorithm employed in [15] has been exploited to program the n_c cells for each one of the above-mentioned normalized conductance targets. Upon the definition of a conductance target interval by specifying its mean value $\hat{g}$ and relative tolerance Δg, each cell is first stimulated with a start SET and a start RESET pulse, which both have a high amplitude current, as they grant better temporal drift retention [12][15]. Then, for each cell, an intermediate SET sequence begins with a single minimum SET amplitude A_{MIN}, with the aim of gradually increase the cell conductance. After a predefined time T_{WAIT} (which, in this study, is in the order of ms because of the latency of the employed GUI), the cell conductance is measured: if it falls within the target interval, the sequence is terminated, otherwise, if the conductance is lower than the required limit, the cell is stimulated with a new intermediate SET pulse increasing its amplitude by a user defined step ΔA. If, instead, the conductance is above the upper limit, the whole process is restarted from the initial SET

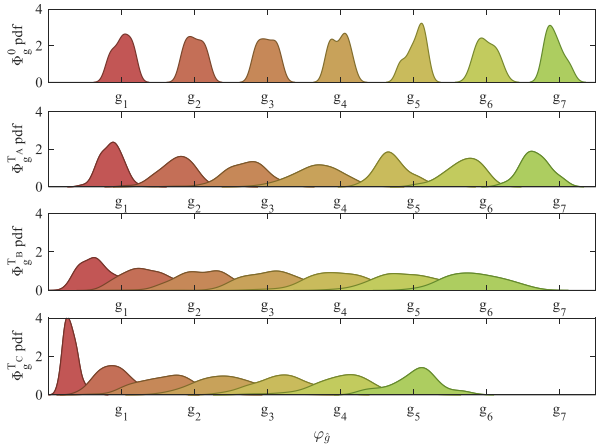

Fig. 3. Probability density function of measured normalized cells conductances. The four plots are related to the distribution after programming (t_0), after 24 hours at T_A = 25°C (t_1), after 24 hours at T_B = 90°C (t_2) and after 24 hours at T_C = 150°C (t_3), respectively.

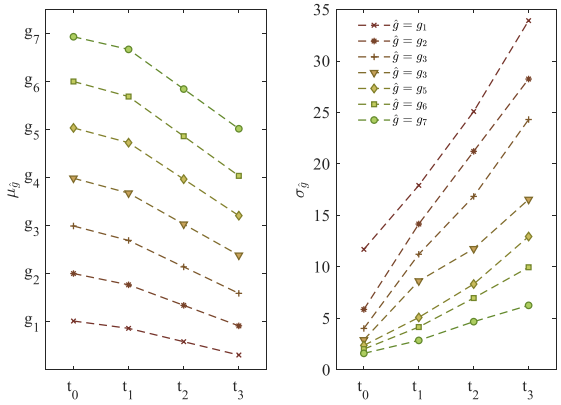

Fig. 4. Left: measured mean normalized conductance $\mu_{\hat{g}}$ after programming (t_0), and in the three conditions corresponding to t_1, t_2, t_3; different curves refer to the seven target conductances $\hat{g}$. Right: measured conductance relative dispersion $\sigma_{\hat{g}}$ in the same conditions.

and RESET pulses. This employed sequence is outlined in Fig. 2 (top). The values of A_{MIN}, ΔA and of intermediate SET pulses amplitudes depend on each conductance target. Fig. 2 (bottom) shows the programming sequences of three sample cells.

According to the purposes of this work, n_c cells have been programmed to the normalized target conductances $\hat{g}$ with a normalized tolerance Δg = 0.3. All cells have reached the target conductance with averagely 16 total programming steps, and a maximum of 54. For each normalized target conductance $\hat{g}$, the set of n_c cell conductances $\varphi^0_{\hat{g},i=1\dots n_c}$ measured after the programming procedure is defined as $\Phi^0_{\hat{g}} := \{\varphi^0_{\hat{g},1}, \varphi^0_{\hat{g},2}, \dots, \varphi^0_{\hat{g},n_c}\}$. Thus, the initial mean value $\mu^0_{\hat{g}}$ and relative dispersion $\sigma^0_{\hat{g}}$ of cell conductances can be defined as:

$$\mu^0_{\hat{g}} := \frac{1}{n_c}\sum_{i=1}^{n_c} \varphi^0_{\hat{g},i} \tag{1}$$

$$\sigma^0_{\hat{g}} := \frac{100}{\mu^0_{\hat{g}}}\sqrt{\frac{1}{n_c-1}\sum_{i=1}^{n_c}\left(\varphi^0_{\hat{g},i} - \mu^0_{\hat{g}}\right)^2} \quad [\%] \tag{2}$$

Then, the cells have been monitored in time to evaluate the effects of conductance drift, dispersion, and noise on $\mu^0_{\hat{g}}$ and $\sigma^0_{\hat{g}}$, which can be arbitrary defined through the programming

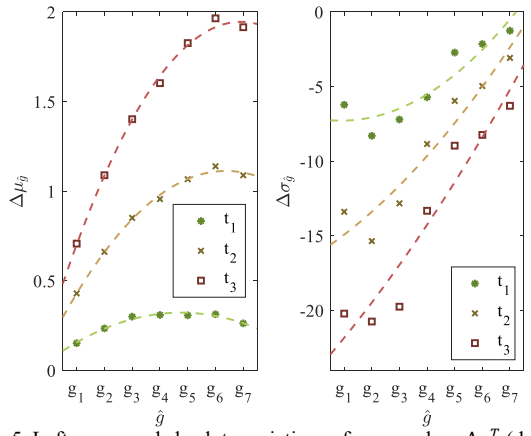

Fig. 5. Left: measured absolute variations of mean values $\Delta\mu_{\hat{g}}^{T}$ (dots) as a function of targets $\hat{g}$, and their fitting functions (dashed lines), in the three conditions corresponding to t_1, t_2, t_3. Right: measured absolute variations of dispersions $\Delta\sigma_{\hat{g}}^{T}$ (dots) as a function of targets $\hat{g}$, and their fitting functions (dashed lines) in the same three conditions.

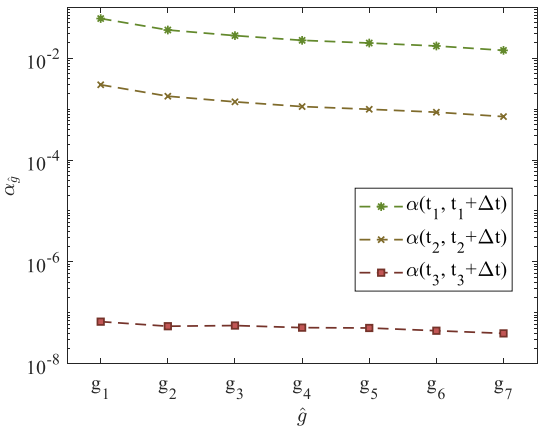

Fig. 6. Mean drift coefficients $\alpha_{\hat{g}}$ of cells sets as a function of the targets $\hat{g}$ in the three conditions corresponding to t_1, t_2, t_3 with Δt = 12 hours.

algorithm. Since these features are strictly related to the intrinsic physical structure of PCM cells [16], different temperatures T have been included in this study. To this purpose, the considered cell conductances sets $\Phi_{\hat{g}}^{0}$ have been measured 24 hours after programming at room temperature $T = T_A$ (approximately 25°C), defining thus a new cell set $\Phi_{\hat{g}}^{T_A} := \left\{\varphi_{\hat{g},1}^{T_A}, \varphi_{\hat{g},2}^{T_A}, \dots, \varphi_{\hat{g},n_c}^{T_A}\right\}$, with its mean value $\mu_{\hat{g}}^{T_A}$ and relative dispersion $\sigma_{\hat{g}}^{T_A}$, both defined accordingly to (1) and (2). Afterward, the test chip has been baked at $T = T_B =$ 90°C. To monitor the dynamics of each cell, the conductances were then measured at room temperature (to avoid leakage current increase due to high temperature) after 24 hours of bake; the same process has been repeated for $T = T_C =$ 150°C, defining thus $\Phi_{\hat{g}}^{T_B}$, $\mu_{\hat{g}}^{T_B}$, $\sigma_{\hat{g}}^{T_B}$, $\Phi_{\hat{g}}^{T_C}$, $\mu_{\hat{g}}^{T_C}$ and $\sigma_{\hat{g}}^{T_C}$.

III. Results and Discussion

A. Evolution of cells distributions

In the first subplot of Fig. 3, the probability density functions (pdfs) of $\Phi_{\hat{g}}^{0} = \Phi_1^0, \Phi_2^0, \dots, \Phi_7^0$ are shown. In this case, the distributions of the seven cell sets are separated, their mean values are near the conductance targets $\hat{g}$, and their boundaries lays under the normalized target tolerance $\Delta g = 0.3$. These conditions are implicitly granted by the adoption of the aforementioned single-cell iterative programming algorithm. In the further subplots, reporting the pdfs of $\Phi_{\hat{g}}^{T_A}$, $\Phi_{\hat{g}}^{T_B}$, and $\Phi_{\hat{g}}^{T_C}$, respectively, it can be easily observed that cells distributions tend to decrease their mean conductance $\mu_{\hat{g}}^{T}$, while their relative dispersion $\sigma_{\hat{g}}^{T}$ increases. As a result, the considered conductances distributions tend to overlap, as memory cells have lost their initial conductance under the combined effect of time and temperature due to of random alterations to their internal structure. The values of $\mu_{\hat{g}}^{0}$ and $\mu_{\hat{g}}^{T}$ are shown in Fig. 4 (left), while the right plot reports the values of $\sigma_{\hat{g}}^{0}$ and $\sigma_{\hat{g}}^{T}$. The mean values of cells sets tend to decrease uniformly with a slight dependence on the mean initial conductance of $\Phi_{\hat{g}}^{0}$, whereas, the cells sets dispersion increase is more evident for cells set with the lower value of target $\hat{g}$. Moreover, as the programming tolerance Δg has been chosen equal for all $\hat{g}$, $\sigma_{\hat{g}}^{0}$ results to be inversely proportional to the target conductance.

In order to describe the behaviors of cells sets $\Phi_{\hat{g}}^{0}$, the absolute variations of mean values $\Delta\mu_{\hat{g}}^{T} := \mu_{\hat{g}}^{T} - \mu_{\hat{g}}^{0}$ and dispersions $\Delta\sigma_{\hat{g}}^{T} := \sigma_{\hat{g}}^{T} - \sigma_{\hat{g}}^{0}$ of cells sets are plotted as dots in Fig. 5 left and right, respectively. As previously shown, $\Delta\mu_{\hat{g}}^{T}$ increases with time and bake temperature. Moreover, $\Delta\mu_{\hat{g}}^{T_A}$ is slightly dependent on the target conductance $\hat{g}$ and varies between 0.1 and 0.3; $\mu_{\hat{g}}^{T_B}$ instead is greater for the higher values of $\hat{g}$ and ranges from 0.3 to 1.1, while $\mu_{\hat{g}}^{T_C}$ varies from 0.5 to 1.9, and shows a strong dependence on $\hat{g}$. For what concerns the cells set dispersion, $\Delta\sigma_{\hat{g}}^{T}$ increases when cells conductance target $\hat{g}$ is greater. In particular, $\Delta\sigma_{\hat{g}}^{T_A}$ varies from -5% to -0.5%, $\Delta\sigma_{\hat{g}}^{T_B}$ varies from -15% to -1%, and $\Delta\sigma_{\hat{g}}^{T_C}$ varies from -25% to -5%. All the measured values of $\Delta\mu_{\hat{g}}^{T}$ and $\Delta\sigma_{\hat{g}}^{T}$ can be fitted with 2nd-order polynomial functions of conductance target $\hat{g}$, which are plotted in Fig. 5 as dashed lines. The fitting functions of $\Delta\mu_{\hat{g}}^{T}$ show a more incisive dependence on the 2nd-order term $\hat{g}^2$, whereas $\Delta\sigma_{\hat{g}}^{T}$ has a stronger dependence on the 1st-order term $\hat{g}$.

B. Effects on drift coefficient

In this paragraph, the effect of time and temperature on the drift dynamic of cells is discussed. The conductance $\varphi(t)$ of a generic PCM cell can described over time t by the empirical law [17]:

$$\varphi(t) = \varphi_0 \left(\frac{t}{t_0}\right)^{-\alpha} \tag{3}$$

where t_0 is a normalizing time factor, φ_0 is the cell conductance at time t_0, and α is a parameter referred to as drift exponent, which is expected to be between 0.01 and 0.11 at room temperature [9]. In this context, three additional measurements have been performed after Δt = 12 hours from t_1, t_2 and t_3, respectively. Comparing these measurements with their corresponding of Fig. 4 left, and inverting Equation (3), drift coefficient has been then estimated. The influence of targets $\hat{g}$ and temperature on α are depicted in Fig. 6, where the mean drift coefficients $\alpha_{\hat{g}}$, which are the mean drift coefficient of $\Phi_{\hat{g}}^{T}$, are plotted in logarithmic scale as a function of the targets $\hat{g}$, and for the three different test conditions. Results show that the drift coefficient slightly depends on the target conductance value, and it is more significant for low $\hat{g}$.

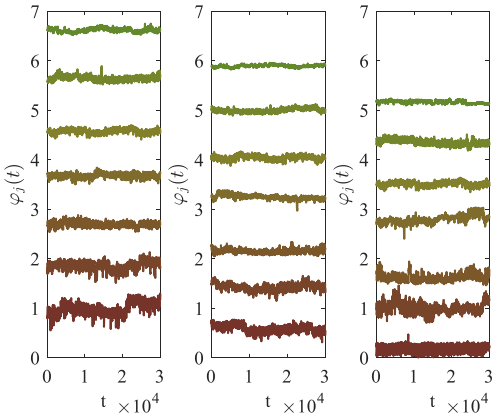

Fig. 7. Example of noise measurement of seven sample cells at $[t_1, t_1 + \Delta t]$ (left), $[t_2, t_2 + \Delta t]$ (center), $[t_3, t_3 + \Delta t]$ (right), with $\Delta t = 5$ min.

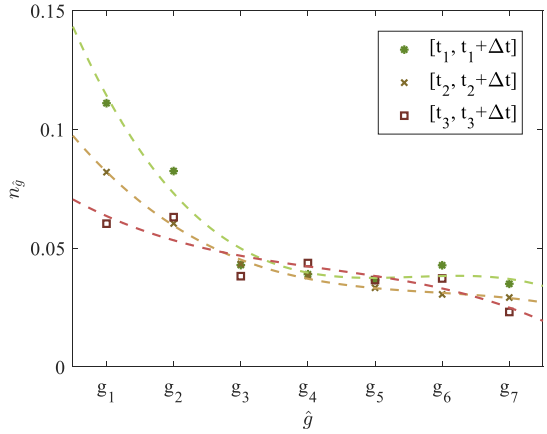

Fig. 8. Mean noise $n_{\hat{g}}^T$ of cells sets as a function of target conductances $\hat{g}$ in the three conditions corresponding to t_1, t_2, t_3.

As α is greater than 0.01 when $T = 25$°C, cells tend concordantly to weakly drift after $\Delta t = 12$ hours. When $T = 90$°C or 150°C, α is near to 0 in the time interval of Δt, concluding that, in these two latter conditions, time drift can be considered negligible. In fact, as suggested in [9], temperature accelerates drift dynamics in GST-rich PCM cells, thus, its effect becomes trifling in few hours when high temperature is applied.

C. Noise

An analysis of cells noise concludes this work. To characterize this aspect, cells conductances have been measured over a time interval Δt = 5 min, collecting n_s = 30000 samples for each cell. Then, the mean time standard deviation $n_{\hat{g}}^T$ for each cells set $\Phi_{\hat{g}}^T$ is considered:

$$n_{\hat{g}}^T = \frac{1}{n_c} \sum_{i=1}^{n_c} \sqrt{\sum_{j=1}^{n_s} \left(\varphi_{i,\hat{g}}^T(t_j) - \bar{\varphi}_{i,\hat{g}}^T\right)^2} \qquad (4)$$

where $\bar{\varphi}_{i,\hat{g}}^T = \langle \varphi_i(t) \rangle$ is the time average conductance observed in Δt for each set $\Phi_{\hat{g}}^T$. An example of noise measurement is reported in Fig. 7, where seven sample-cells conductances $\varphi(t)$ are showed. Measures have been performed at $t_1 + \Delta t$, $t_2 + \Delta t$, $t_3 + \Delta t$. The mean $n_{\hat{g}}^T$ as a function of target conductances $\hat{g}$ is reported in Fig. 8. Results show that noise is more relevant for lower values of $\hat{g}$, where it is about three times greater. Moreover, $n_{\hat{g}}^T$ does not significantly differ for different temperatures, especially for $\hat{g}_3 \dots \hat{g}_6$. Accordingly, noise in PCM elements is related to amorphous phase of cells, which is more significant in low-conductance ones.

IV. Conclusion

In this study, a preliminary characterization of phase-change memory (PCM) cells in an analog in-memory computing (AIMC) perspective has been carried out. Drift, dispersion and noise have been analyzed in relation to memory elements programmed with a dedicated programming algorithm, showing their dependences on conductance targets and temperature. Further directions of this research context will concern the development of an accurate model to describe and simulate cells behavior, as well the optimization of the employed programming algorithms.

References

[1] G. W. Burr, B. N. Kurdi, J. C. Scott, C. H. Lam, K. Gopalakrishnan, and R. S. Shenoy, "Overview of candidate device technologies for storage-class memory," IBM J. Res. Dev., vol. 52, no. 4–5, pp. 449–464, 2008.

[2] Z. Sun, G. Pedretti, E. Ambrosi, A. Bricalli, W. Wang, and D. Ielmini, "Solving matrix equations in one step with cross-point resistive arrays," Proc. Natl. Acad. Sci. U. S. A., vol. 116, no. 10, pp. 4123–4128, 2019.

[3] W. Haensch, T. Gokmen, and R. Puri, "The Next Generation of Deep Learning Hardware: Analog Computing," Proc. IEEE, vol. 107, no. 1, pp. 108–122, 2019.

[4] V. Joshi et al., "Accurate deep neural network inference using computational phase-change memory," Nat. Commun., vol. 11, no. 1, 2020.

[5] D. Ielmini and S. Ambrogio, "Emerging neuromorphic devices," Nanotechnology, vol. 31, no. 9, p. 092001, Feb. 2020.

[6] V. Milo, G. Malavena, C. M. Compagnoni, and D. Ielmini, "Memristive and CMOS devices for neuromorphic computing," Materials (Basel)., vol. 13, no. 1, p. 166, 2020.

[7] G. Hackmann, F. Sun, N. Castaneda, C. Lu, and S. Dyke, "A Holistic Approach to Decentralized Structural Damage Localization Using Wireless Sensor Networks," in 2008 Real-Time Systems Symposium, Nov. 2008, pp. 35–46.

[8] X. Sun et al., "PCM-Based Analog Compute-In-Memory: Impact of Device Non-Idealities on Inference Accuracy," IEEE Trans. Electron Devices, vol. 68, no. 11, pp. 5585–5591, 2021.

[9] F. G. Volpe, A. Cabrini, M. Pasotti, and G. Torelli, "Drift induced rigid current shift in Ge-Rich GST phase change memories in low resistance state," 2019 26th IEEE Int. Conf. Electron. Circuits Syst. ICECS 2019, pp. 418–421, 2019.

[10] M. Carissimi et al., "2-Mb Embedded Phase Change Memory with 16-ns Read Access Time and 5-Mb/s Write Throughput in 90-nm BCD Technology for Automotive Applications," ESSCIRC 2019 - IEEE 45th Eur. Solid State Circuits Conf., pp. 135–138, 2019.

[11] M. Pasotti et al., "A 32-KB ePCM for Real-Time Data Processing in Automotive and Smart Power Applications," IEEE J. Solid-State Circuits, vol. 53, no. 7, pp. 2114–2125, 2018.

[12] F. Bedeschi et al., "A bipolar-selected phase change memory featuring multi-level cell storage," IEEE J. Solid-State Circuits, vol. 44, no. 1, pp. 217–227, 2009.

[13] N. Papandreou et al., "Programming algorithms for multilevel phase-change memory," Proc. - IEEE Int. Symp. Circuits Syst., pp. 329–332, 2011.

[14] A. Cabrini, S. Braga, A. Manetto, and G. Torelli, "Voltage-driven multilevel programming in phase change memories," Proc. 2009 IEEE Int. Work. Mem. Technol. Des. Testing, MTDT 2009, pp. 3–6, 2009,

[15] A. Antolini et al., "Characterization and programming algorithm of phase change memory cells for analog in-memory computing," Materials (Basel)., vol. 14, no. 7, 2021, doi: 10.3390/ma14071624.

[16] S. Raoux et al., "Phase-change random access memory: A scalable technology," IBM J. Res. Dev., vol. 52, no. 4–5, pp. 465–479, 2008.

[17] D. Ielmini, A. L. Lacaita, and D. Mantegazza, "Recovery and drift dynamics of resistance and threshold voltages in phase-change memories," IEEE Trans. Electron Devices, vol. 54, no. 2, pp. 308–315, 2007.