Experimentally Validated $Pr_{0.7}Ca_{0.3}MnO_3$ RRAM Verilog-A model based Izhikevich Neuronal Dynamics

Omkar Phadke
Dept. of Electrical Engineering,
IIT Bombay, Mumbai, India
omkarphadke@ee.iitb.ac.in

Arpan De
Dept. of Electronics and Telecommunication Engineering, Jadavpur University,
Kolkata, India
arpantukan@gmail.com

Jayatika Sakhuja
Dept. of Electrical Engineering,
IIT Bombay, Mumbai, India
jayatika@ee.iitb.ac.in

Vivek Saraswat
Dept. of Electrical Engineering,
IIT Bombay, Mumbai, India
svivek@ee.iitb.ac.in

Udayan Ganguly,
Dept. of Electrical Engineering,
IIT Bombay, Mumbai, India
udayan@ee.iitb.ac.in

***Abstract*—Spiking Neural Networks (SNNs) are brain-inspired computational networks that promise an efficient solution for real-life applications such as audio processing and pattern recognition. An SNN is a complex network of neurons interconnected with synapses, whose spike times and synaptic strength is essential for information processing. SNNs consume low power and perform computations in parallel, making them attractive for hardware implementation of applications in neuromorphic engineering. For hardware implementation of SNNs, we need devices that mimic the behavior of neurons and synapses as well as their computational models to build and characterize large-scale SNNs. Earlier, a $Pr_{0.7}Ca_{0.3}MnO_3$ (PCMO) material based Resistive-RAM (RRAM) has been used to experimentally demonstrate an Integrate and Fire (IF) Neuron and Izhikevich Neuron. Here, we present an experimentally validated Verilog-A model of PCMO RRAM, which behaves like a neuron. The model captures the conductance change of the RRAM for different applied voltages and can mimic an IF Neuron and an Izhikevich Neuron. The model enables the design of large-scale SNNs and studies their behavior in a simulation domain.**

Keywords—PCMO, RRAM, SNNs, Izhikevich Neuron, Integrate and Fire Neuron

I. INTRODUCTION

Brain-inspired computing with spiking neural networks (SNNs) are thrusting efficient solutions for real-life applications such as pattern recognition and audio processing [1]. SNNs consume less power and perform computations in parallel, making them an attractive alternative for traditional computational systems based on the von Neumann architecture [2]. SNNs are comprised of neurons that are interconnected via synapses. A biological neuron 'integrates' the input signal to raise its membrane potential. Once the membrane potential exceeds a threshold value, the neuron performs a 'fire' operation. The event is referred to as neuronal spiking. A biological synapse can adjust its conductance level, or synaptic strength, based on the time difference between the spiking of neurons situated at the two ends of the synapse. Owing to spike timings, the human brain can exhibit different spiking patterns, as shown in Fig. 1(a). These spiking patterns are correlated to different brain functions [3]. To efficiently mimic human brain in neural networks, it is essential to capture the different spiking patterns of biological neurons. For a hardware implementation of SNNs, we need devices that are capable of performing an integration and fire operation (to function as a neuron) and exhibit multiple controllable conductance levels (to function as a synapse). In recent years, RRAM is proving to be a promising candidate for such applications.

For SNNs, one of the upcoming devices is the PCMO material-based RRAM. The PCMO based RRAM exhibits a gradual RESET operation (RESET: Decrease in conductance) and an abrupt SET operation (SET: Increase in conductance) with respect to the applied voltage. Recently, the PCMO based RRAM has been utilized to implement the spike time-dependent plasticity (STDP) rule of the synapse (Fig. 1b). STDP plays a vital role in the learning of neural networks [4-5]. Additionally, using the change in conductance states in PCMO RRAM, a clocked neuron has also been demonstrated, which is capable of exhibiting Integrate and Fire (IF) operation and complex Izhikevich dynamics [6-7]. Further, with PCMO RRAM-based neurons, if the set-up parameters are kept constant, including device area and input bias, we can modulate neuron spiking frequencies by changing only the initial conductance state of the device [7]. Since PCMO can be used for both a synapse and a neuron, an SNN can be fabricated out of the same material system, decreasing the complexity in hardware implementation (Fig.1b). Moreover, in conventional CMOS-based neurons, large capacitors are needed to perform the integration operation at a biological timescale; hence more area is required, as shown in Fig.1(c) [9]. In contrast, NVM-based neurons are able to generate large timescales using device physics and are still nanoscale as there is no requirement for external capacitors. For biological SNNs, RRAMs offer non-volatile memory and analog synapse capability with area efficiency. PCMO based RRAM devices are non-filamentary and area scalable. Therefore, with small area capacitor-less PCMO RRAM-based neurons, dense neural networks can be designed.

The work is partially funded by the Office of Principal Scientific Advisor, DST, and Ministry of Electronics and Information technology (MeitY), India.

978-1-6654-0685-7/21/$31.00 ©2021 IEEE

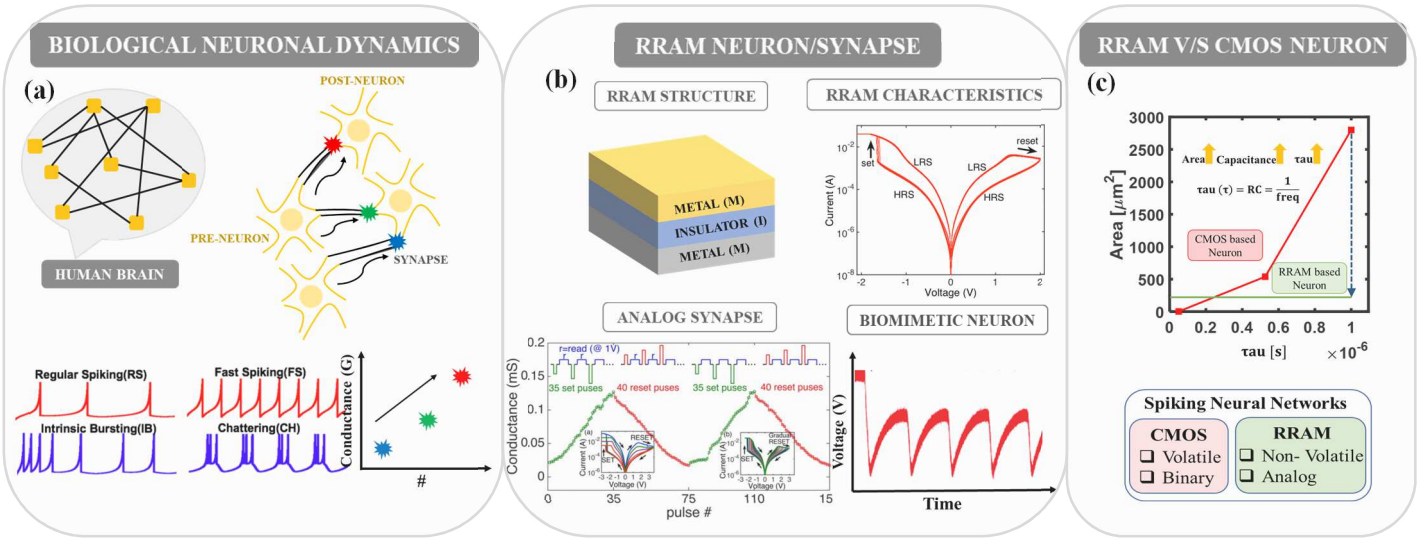

Fig. 1. (a) Biological Neuronal Dynamics. Different neurons are interconnected via synapses. Analog conductance of the synapse determines the strength of the connection. Various neuron patterns are emulated by Izhikevich Model. (b) RRAM device schematic and its DCIV characteristics. RRAM exhibit both synapse and neuron characteristics [4,8]. (c) CMOS-based neurons need large capacitors (area penalty) at a biological timescale. NVM-based neuron are nanoscale even for large ($\tau = RC$) [9]. For biological SNNs, RRAMs offer non-volatile memory and analog synapse capability with area efficiency.

However, prior to hardware implementation of SNNs, it is essential to analyze the SNNs in a simulation environment, which would require simple yet comprehensive computational models. Previously, extensive research has been done with RRAM device switching models [10-12]. In [13], a reaction drift model is introduced with PCMO RRAM devices. In this paper, we propose a modified physics-based device model written in Verilog-A. The proposed model captures the current transport, electrothermal dynamics, and trap density modulation (trap generation-recombination) of the PCMO based RRAM device. The model has been calibrated with the experimental neuronal dynamics of IF neuron, and Izhikevich dynamics. Such models enable the design of large neural networks and study ensemble dynamics.

II. Device Details

The PCMO RRAM has a metal-insulator-metal (MIM) structure, with Tungsten (W) as the top reactive electrode and Platinum (Pt) as the non-reactive bottom electrode, as shown in Fig.2(a). The device is fabricated on a Si/SiO2/Ti substrate. The PCMO RRAM provides multiple analog conductance levels, which are controlled by the trap density of the device. The trap density can be decreased by the SET process and increased by the RESET process. In the SET (RESET) process, a negative (positive) voltage pulse is applied to the top W electrode of the RRAM. A redox reaction occurs at the reactive top electrode-PCMO interface, where oxygen ions are pushed into (out of) the PCMO material and trap density decreases (increases).

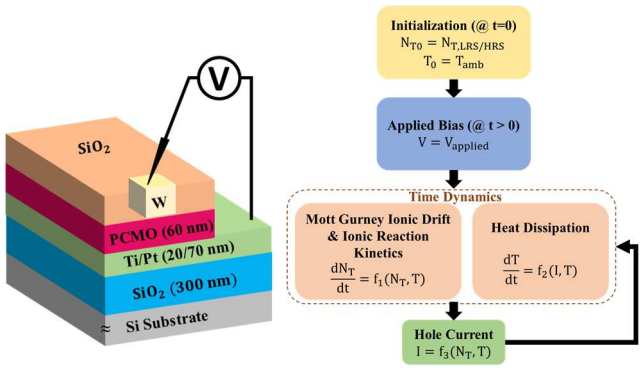

Fig. 2. (a) Device Schematics of PCMO based RRAM (b) Flowchart describing current computation in Verilog-A model. Current is a function of temperature and trap density.

III. Device Model

To model the behavior of the RRAM, it is important to implement the current transport, electrothermal dynamics, and trap generation-recombination dynamics. Each of these phenomena is described below.

A. Current Model

The current through PCMO RRAM follows Space Charge Limited Current (SCLC) mechanism [14-15]. The following are equations are used to model the current-voltage-trap density relationship,

$$I = I_{Ohmic} + I_{SCLC} \tag{1}$$

$$I_{Ohmic} = qA\mu N_v \left(\frac{T}{T_{amb}}\right)^{3/2} e^{\left(-\frac{q\phi_B}{kT}\right)} \left(\frac{V}{L}\right) \tag{2}$$

$$I_{SCLC} = A\mu\epsilon_o\epsilon_{PMO} \left(\frac{N_v}{N_T}\right)\left(\frac{T}{T_{amb}}\right)^{3/2} e^{\left(-\frac{qE_{trap}}{kT}\right)} \left(\frac{V^2}{L^3}\right) \tag{3}$$

B. Thermal Model

The PCMO RRAM exhibits self-heating, which facilitates the resistive-switching within the device [16]. The device temperature is computed using the following Fourier heat equation,

$$\frac{T - T_{amb}}{R_{th}} + C_{th}\frac{dT}{dt} = IV \tag{4}$$

TABLE I. MODEL PARAMETER VALUES

Model	Parameter Information		
	Symbol	***Quantity***	***Value***
Current Model	μ	Mobility	17.5 cm^2/V-s
	ϕ_B	Barrier Height	NA[a]
	ϵ_{PMO}	Dielectric Constant	30
	N_v	Effective Density of States	8.16 x 10^{19} cm^{-3}
	E_{trap}	Trap Level	0.06 eV
	N_T	Trap Density	NA[b]
Device Specification	L	Length	60 nm
	A	Area	1 x 1 um^2
Thermal Model	T_{amb}	Ambient Temperature	300 K
	R_{th}	Thermal Resistance	19006 K / W
	C_{th}	Thermal Capacitance	0.26 pJ / K
Reaction Drift Model	k_{eq}	Reaction Rate	$\sim(0.1/(\mu m^4))^3$
	n	Traps per anion [13]	2
	f	Calibration factor	1-10

[a.] Depends on the trap density, as per [14]

[b.] Depends on the conductance level

C. Reaction Drift Model

The trap density modulates the conductance of the device within the device [13]. On the application of a positive (negative) bias to the top electrode, trap generation (recombination) occurs at the reactive top electrode followed by transport (drift) of vacancies into the bulk under the influence of an electric field. With trap generation (recombination), trap density increases (decreases) within the device. This, in turn, leads to a lower (higher) current through the device, and hence a lower (higher) conductance level. The trap density modulation is modeled using the following equations referred from [13],

$$\frac{dN_T}{dt} = \frac{k_{eq} v_{drift}}{nLN_T^n} \quad (5)$$

Here, N_T is the trap density within the device, k_{eq} is reaction rate, v_{drift} is drift velocity, and L is device thickness.

The above equations are modeled in Verilog-A and simulated using the Analog Design Environment (ADE) of Cadence Virtuoso. The model is initialized with an ambient temperature T_{amb} (300K), a constant input voltage, and an initial trap density which corresponds to the initial conductance level according to the experiments. Equations (1) to (5) are solved self-consistently. Fig. 2(b) demonstrates the flowchart of the model, and Table I shows the parameters used in the model.

IV. MODEL DETAILS

In the presented model, to capture all the experimental features for the set of parameters, a calibration factor is proposed, and (5) is multiplied by the factor. The factor adjusts the amount of conductance change. The reasoning and justification for the proposed methodology are presented below.

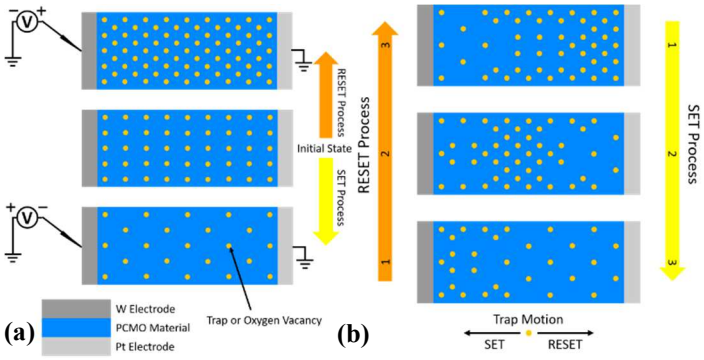

Fig. 3. (a) Uniform trap density profile (b) Non-uniform trap density within the bulk PCMO material.

A. Assumption of the Reaction Drift Model

The reaction drift model presented in [13] and (5) assumes that the trap density profile $N_T(x)$ and the electric field profile E(x) are uniform throughout the length of the device. Further, either the SET or RESET process modulates the trap density uniformly throughout the device (Fig. 3(a)). During the SET process, as the overall (or effective) trap density decreases, the trap density at the interface reduces, increasing the rate of change of trap density (5). During the RESET process, as the overall (or effective) trap density increases, the trap density at the reactive electrode interface increases as well. This reduces the rate of change of trap density, as per (5).

B. Experimental Evidence for the Trap Density Profile

Recently, a 3 Terminal PCMO based RRAM was proposed, which investigated the region of the resistance change of the RRAM [17]. After a resistive switching process, the region of the resistance change, or the region where the trap density has been modulated, lies near the non-reactive electrode, away from the reactive electrode. This implies that $N_T(x)$ and E(x) are non-uniform.

C. Simplified Uniform Trap Profile Approximation

Fig. 3(b) shows a representation of the non-uniform trap density profile. As shown in [17], a cluster of traps would be present near the non-reactive Pt electrode, away from the reactive W electrode.

When we apply a negative bias to the reactive W electrode, the cluster of traps will start to move towards the W electrode. Once the cluster of traps reaches the reactive electrode, the traps are consumed, lowering the overall (or effective) N_T of the device. The N_T at the reactive electrode interface may be higher than the effective N_T as the transport of traps towards the interface (increasing trap density) will be in competition with the annihilation of traps owing to reaction (decreasing trap density). This may give rise to a different dN_T/dt than observed in the reaction drift model.

When we apply a positive bias to the reactive W electrode, the traps are generated at the reactive electrode interface, increasing the overall (or effective) N_T of the device. These newly generated traps travel under the electric field towards the non-reactive Pt electrode and accumulate there, which leads to

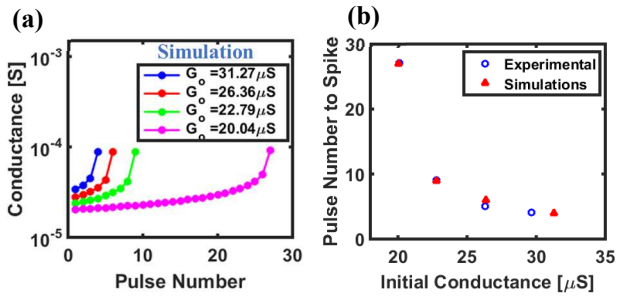

Fig. 4. Verilog-A simulation with consecutive pulses of equal amplitude and duration are applied to the device with different initial conductance. (G_o). (a) Conductance vs Pulse No. to achieve a single spike (b) No. of pulses required to get a spike.

a cluster of traps at the Pt electrode. Since the traps generated at the interface are removed simultaneously, the dN_T/dt would be different than predicted in the reaction drift model.

Considering the non-uniformity of the trap density and electric field profile, (5) is multiplied by a calibration factor '*f*' to alter the rate of change of trap density, which contains the effect of non-uniform trap density. The factor *f* is constant with time, initialized at the start of the simulations for different (V_{app}, $N_{T,initial}$). In [14], it is shown that a non-uniform trap density profile does not lead to significant asymmetry in the IV characteristics of the device. Thus, an effective uniform trap density modeled with factor '*f*' is a sufficient approximation. The presented model matches the experimental frequency vs. V_{app} characteristics of the IF neuron. It also exhibits the experimentally demonstrated decreasing spike time with increasing conductance characteristics.

V. Simulation Results

The model is initialized to N_T as per the initial conductance level, T = 300K, and V = V_{app}. Factor *f* is calibrated for each simulation, for which results are presented below:

A. Effect of Initial Conductance on Spiking Frequency

For different initial conductance values (or trap densities), a train of voltage pulses with (V_{app}, Pulse Width) = (-2.4 V, 150 ns) is applied to the device. Once the RRAM reaches a particular conductance level, a spike is issued. As shown in Fig. 4(a), with the increase in initial conductance (G_o) of the device, the number of pulses required to elicit a spike decreases. The model is calibrated to match the experimental data for different conductance levels, and the results are shown in Fig. 4(b).

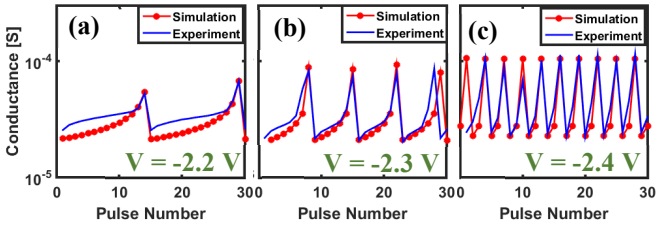

Fig. 5. Integrate and Fire (IF) Operation: Comparison of Experimental and Simulated results for SET voltages (a) V = -2.2, (b) V = -2.3, (c) V = -2.4 Pulse Width = 150 ns were applied. The Verilog-A simulations accurately captures the spiking behavior observed in the experiments.

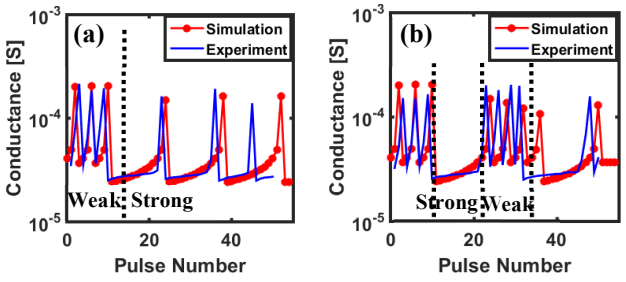

Fig. 6. Izhikevich Operation: Experimental & Simulated results of (a) Intrinsic Bursting and (b) Chattering Neuron. A SET pulse of -2.2 V for spike and a strong or weak RESET operation depending upon the desired behavior.

B. Integrate and Fire Behavior (IF)

Fig. 5 shows the calibrated simulation results to exhibit IF behavior for three different voltage trains of V_{app} = -2.2V, -2.3V, -2.4V, and a Pulse Width of 150 ns. With each successive pulse in the pulse train, the trap density of the device is integrated, raising the conductance level of the device, imitating the 'integrate' operation. Once the conductance level exceeds a threshold value, the conductance change is abrupt, mimicking a fire operation. Here, a RESET pulse is applied to reduce the conductance to a low level, imitating the 'fire' operation. The voltage pulse train is applied again to elicit a spike. Higher applied voltage integrates more trap density; hence the conductance change with each successive pulse is higher, and the spike time is low. For lower voltages, the amount of trap density integrated is comparatively less, which leads to a smaller spiking time.

C. Izhikevich Dynamics

Fig. 6 shows the calibrated simulation results to exhibit Izhikevich dynamics of Chattering Behavior (CH) and Intrinsic Bursting Behavior (IB). To exhibit a burst of spikes, a small RESET voltage (Weak RESET) is applied to reduce the conductance level to a moderate value, rather than a small conductance value (via a Strong RESET), and a SET pulse is applied immediately. This process is repeated until the desired number of spikes are reached, then a Strong RESET is performed.

VI. Conclusion

To summarize, we demonstrated IF neuron behavior and different neuronal spiking pattern with the proposed computational Verilog-A model of PCMO RRAM-based device. The calibrated simulation results show that the model successfully captures the device dynamics and hence will enable an efficient circuit simulator environment to develop SNNs. With SNNs build on comprehensive models, solutions to different intricate problems can be predicted with higher accuracy.

REFERENCES

[1] W. Maass, "Networks of Spiking Neurons : The Third Generation of Neural Network Models," Neural Networks, vol. 10, no. 9, pp. 1659–1671, December 1997.

[2] J.H.B. Wijekoon, and P. Dudek. "Compact silicon neuron circuit with spiking and bursting behaviour." Neural Networks vol. 21, no. 2-3, pp. 524-534, 2008.

[3] E. M. Izhikevich, "Which model to use for cortical spiking neurons?," in IEEE Transactions on Neural Networks, vol. 15, no. 5, pp. 1063-1070, September. 2004

[4] N. Panwar, B. Rajendran and U. Ganguly, "Arbitrary Spike Time Dependent Plasticity (STDP) in Memristor by Analog Waveform Engineering," IEEE Electron Device Letters, vol. 38, no. 6, pp. 740-743, June 2017.

[5] S. Lashkare, N. Panwar, P. Kumbhare, B. Das and U. Ganguly, "PCMO-Based RRAM and NPN Bipolar Selector as Synapse for Energy Efficient STDP", in IEEE Electron Device Letters, vol. 38, no. 9, pp. 1212-1215, September 2017.

[6] S. Lashkare, S. Chouhan, T. Chavan, A. Bhat, P. Kumbhare, and U. Ganguly, "PCMO RRAM for Integrate-and-Fire Neuron in Spiking Neural Networks," IEEE Electron Device Lett., vol. PP, no. 99, p. 1, February 2018.

[7] S. Lashkare, S. Chouhan, A. Bhat and U. Ganguly, "General Izhikevich Dynamics In PR0.7 CA0.3 MNO3 RRAM Neuron," 2020 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA), pp. 48-49, September 2020.

[8] S. Lashkare, P. Kumbhare, V. Saraswat and U. Ganguly, "Transient Joule Heating-Based Oscillator Neuron for Neuromorphic Computing," in IEEE Electron Device Letters, vol. 39, no. 9, pp. 1437-1440, September 2018.

[9] T. Chavan, S. Dutta, N. R. Mohapatra and U. Ganguly, "Band-to-Band Tunneling Based Ultra-Energy-Efficient Silicon Neuron," in IEEE Transactions on Electron Devices, vol. 67, no. 6, pp. 2614-2620, June 2020.

[10] A. Sawa, T. Fujii, M. Kawasaki, and Y. Tokura, "Hysteretic current–voltage characteristics and resistance switching at a rectifying Ti⁄Pr 0.7 Ca 0.3 Mn O 3 interface," Applied Physics Letters, vol. 85, no. 18, pp.4073-4075, November 2004.

[11] A. Baikalov, Y.Q. Wang, B.Shen, B.Lorenz, S. Tsui, Y.Y. Sun, Y.Y. Xue, and C. Chu, "Field-driven hysteretic and reversible resistive switch at the Ag–Pr 0.7 Ca 0.3 MnO 3 interface," Applied Physics Letters, vol. 83, no. 5, pp.957-959, July 2003

[12] B. Hajri, M. M. Mansour, A. Chehab and H. Aziza, "Oxide-based RRAM models for circuit designers: A comparative analysis," 2017 12th International Conference on Design & Technology of Integrated Systems In Nanoscale Era (DTIS), pp. 1-6, May 2017

[13] V. Saraswat et al., "Reaction-Drift Model for Switching Transients in $Pr_{0.7}Ca_{0.3}MnO_3$-Based Resistive RAM," in IEEE Transactions on Electron Devices, vol. 67, no. 9, pp. 3610-3617, September 2020.

[14] I. Chakraborty, A. K. Singh, P. Kumbhare, N. Panwar and U. Ganguly, "Materials parameter extraction using analytical models in PCMO based RRAM," 2015 73rd Annual Device Research Conference (DRC), pp. 87-88, August 2015.

[15] M.A., Lampert," Volume-controlled current injection in insulators", Reports on Progress in Physics, vol. 27, no. 1 , p.329,1964.

[16] N. Panwar, A. Khanna, P. Kumbhare, I. Chakraborty, and U. Ganguly, "Self-heating during submicrosecond current transients in $Pr_{0.7}Ca_{0.3}MnO_3$-based RRAM", IEEE Trans. Electron Devices, vol. 64, no. 1, pp. 137-144, January 2017.

[17] S.Lashkare, V. Saraswat, and U. Ganguly. "Understanding the region of resistance change in Pr0. 7Ca0. 3MnO3 RRAM." ACS Applied Electronic Materials, vol. 2, no. 7, pp. 2026-2031, June 2020.