A Time-based Sensing Scheme for Multi-level Cell (MLC) Resistive RAM

John Reuben
Chair of Computer Science 3 - Computer Architecture
Friedrich-Alexander-University Erlangen-Nurnberg
Erlangen, Germany
johnreubenp@gmail.com

Dietmar Fey
Chair of Computer Science 3 - Computer Architecture
Friedrich-Alexander-University Erlangen-Nurnberg
Erlangen, Germany
dietmar.fey@fau.de

***Abstract*—The quest to increase memory density in Resistive Random Access Memory (RRAM) has motivated researchers to store more bits/cell by implementing Multi-Level Cell (MLC) or multi-bit RRAM. Implementing multiple states narrows the distance between states, making sensing of MLC RRAM a challenging task. In this paper, we present a circuit which senses the state of a MLC by converting the current drawn from the cell to voltage pulses, where the number of pulses is proportional to the current's magnitude. The circuit distinguishes between the states by the relative current's magnitude and hence does not require an absolute reference. Simulations in IHP's 130 nm CMOS technology confirmed fast (single step) sensing while tolerating appropriate variations in the sensed resistance. The proposed circuit is also area efficient when compared to conventional parallel sensing approach.**

***Index Terms*—read, multilevel cell (MLC), schmitt-trigger, sense amplifier, Resistive RAM, sensor, memristor**

I. Introduction

Resistive Random Access Memory (RRAM) is an emerging Non-Volatile Memory (NVM) with increasing applications in memory and logic circuits [1], [2]. The fundamental device in RRAM is a Metal-Insulator-Metal structure which can store data as resistance of a conductive filament formed in the insulator [3]. The conductive filament can be grown (Low Resistance State (LRS)) and broken (High Resistance State (HRS)) under voltage stress, enabling writing and erasing of data. The quest to increase memory density has motivated researchers to store more bits/cell by implementing Multi-Level Cell (MLC) or multi-bit RRAM. In RRAM, MLC is implemented by varying the compliance current, or by varying the voltage, or by varying the programming pulse widths [4]. Among these three methods, implementing MLC by varying the compliance current is the most viable method to implement MLC in RRAM [5]. The RRAM is integrated in series with a transistor and the compliance current is varied by varying the gate voltage of the transistor in a 1 Transistor-1 Resistor (1T1R) configuration. In this way, a single HRS and multiple LRS (corresponding to different compliance currents) are implemented and, the physical phenomenon is believed to be the formation and subsequent widening of the conductive filament with increasing compliance current (Fig.1-(a)). Demonstration of MLC in RRAM by varying the compliance current can be found in [6]–[9] (read-out currents of some of these MLC RRAMs is listed in Table I).

To 'read' the data from a MLC RRAM, we need a sensing methodology to convert the resistance to a digital data, which is the focus of this paper. Based on the architecture, sensing methodology can be either sequential or parallel. In a sequential approach to MLC sensing, a single Sense Amplifier (SA) is used and numerous comparisons are made by varying the reference quantity (voltage or current) sequentially, resulting in the identification of the cell resistance [10]. The parallel approach uses numerous sense amplifiers and compares the read quantity with the reference quantity simultaneously, similar to a flash ADC. The former approach has less hardware complexity but incurs latency, while the latter achieves high speed sensing at the cost of hardware. From another perspective, the sensing methodology for resistive memories can be voltage-mode or current-mode. In voltage-mode sensing, the bit-line (BL) is pre-charged to a voltage and then the word-line (WL) is activated. Depending on the RRAM cell's resistance, the BL voltage changes (either marginally if HRS or drastically if LRS) and the change is captured by comparing with a reference voltage in voltage-mode SA. In current-mode sensing, a small voltage is applied across the RRAM cell and the induced current is drawn out and compared with the reference current in a current-mode SA. A detailed review of both the schemes and the challenges faced in sensing can be found in [11]. All these sensing techniques need an absolute reference voltage/current for comparison and generating multiple references (V_{REF}/I_{REF}) adds overhead to the sensing circuitry. In sequential sensing, the references have to be generated and also compared with the quantity to be sensed using a control circuit. In this paper, we present a circuit which senses the state of a MLC by converting the current drawn from the memory cell to voltage pulses, where the number of pulses is proportional to the current's magnitude. The sensor delineates the states by the relative current magnitude and hence, does not require any reference current. The circuit and the simulations results are presented in the following section.

978-1-7281-2769-9/19/$31.00 ©2019 IEEE

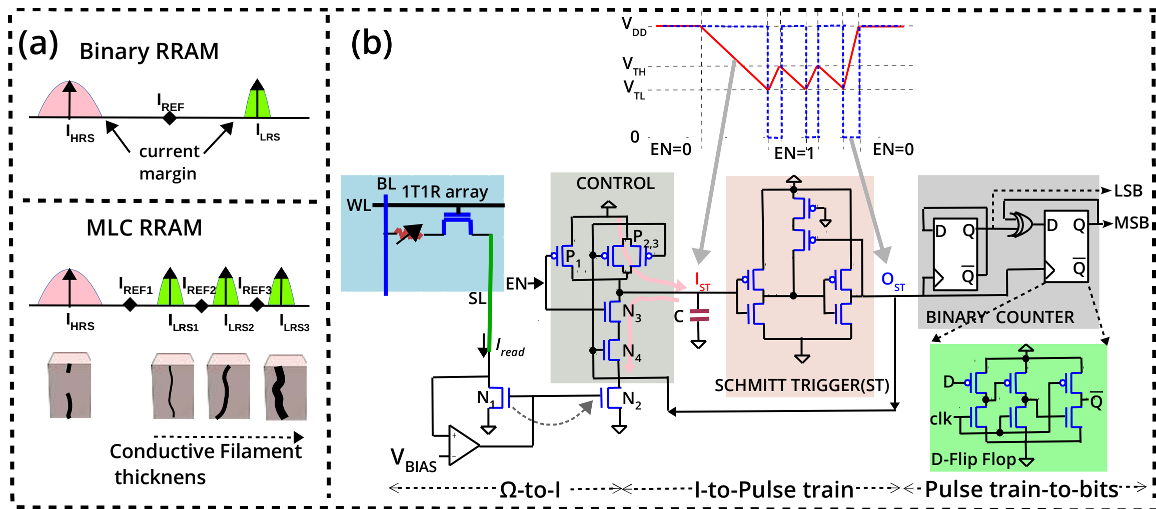

Fig. 1. (a) Multiple states in RRAM: HRS corresponds to a broken filament while LRS1,2,3 corresponds to filaments of different thickness. (b) Sensing methodology: The current drawn from the memory array (I_{read}) is used to discharge the capacitor. A Schmitt-Trigger circuit (with lower threshold voltage V_{TL} and upper threshold voltage V_{TH}) converts the current to equivalent pulses and a binary counter counts the pulses to evaluate the state of the MLC RRAM (for 8 states, a three bit counter can be used). I_{ST} denotes input node of ST and O_{ST} denotes output node of ST

TABLE I

A SAMPLE OF RECENTLY FABRICATED MLC RRAMS WITH THEIR MEDIAN RESISTIVE STATES. INSTEAD OF THE LRS/HRS IN Ω, THE READ-OUT CURRENT AT 0.2 V IS REPRODUCED FROM PUBLISHED WORKS

Device	HRS	LRS1	LRS2	LRS3	LRS4	LRS5	LRS6	LRS7	Ref
HfO_2 (2-bit/cell)	3μA	20μA	30μA	40μA					[6]
TaO_2 (3-bit/cell)	1.5μA	30μA	50μA	80μA	100μA	150μA	200μA	300μA	[8]
TaO_2 (3-bit/cell)	1μA	10μA	20μA	40μA	50μA	70μA	100μA	120μA	[9]

II. PROPOSED SENSING METHODOLOGY

A. Principle

In the first stage of sensing, the RRAM's resistance is converted to a current which flows in the N_1-N_2 current mirror, following the approach of [12]. In a 1T-1R configuration, this is implemented by activating the WL and applying a small voltage (typically <= 0.2 V so that the cell's state is not disturbed) across the cell. As depicted in Fig.1-(b), the op-amp biases the drain of N_1 at a constant voltage, V_{BIAS} to ensure that N_1 is in saturation (feedback bias [12]). Therefore, transistor pair $N_1 - N_2$ acts as a current-mirror and I_{read} will be mirrored in N_2 and is available for sensing[1]. This I_{read} is used to discharge the capacitor which is pre-charged to V_{DD} (when *SENSE ENABLE* signal (EN) is low, the capacitor is charged through P_1 to V_{DD}). When EN goes high, sensing starts. The capacitor discharges from V_{DD} at a rate proportional to I_{read}. However, when the voltage at I_{ST} goes below V_{TL}, O_{ST} goes low and stops the discharging process (N_4 is OFF). The capacitor at I_{ST} starts charging (through $P_{2,3}$) till it reaches V_{TH}. When I_{ST} reaches V_{TH}, O_{ST} goes high and this triggers the discharging of I_{ST} (N_4 is ON). This discharging and charging repeats as long as EN is high. This periodic discharging and charging of the capacitor results in a pulse train at O_{ST} (a discharge followed by a charge constitutes a single negative pulse). Since the discharging time is proportional to I_{read}, a higher current will generate more pulses. By carefully choosing the capacitor value and EN time period, currents of increasing magnitude can be converted to increasing number of pulses. The resulting pulses are then converted to bits using a synchronous binary counter clocked with the pulse train.

[1]In conventional current-mode sensing, this mirrored current is compared with a reference current (I_{REF}) in a sense amplifier. Four states require three comparisons with $I_{REF1}, I_{REF2}, I_{REF3}$ and the corresponding control circuit to orchestrate it, Fig. 1-(a)

B. Circuit design and simulation results

In this section, we describe the design methodology of the sensing circuit of Fig. 1 in IHP's 130 nm CMOS technology (V_{DD} = 1.2 V). The op-amp used to bias the current mirror is a classical two-stage miller-compensated op-amp. V_{BIAS} of 0.8 V was used at the input of op-amp to bias the drain of N_1. Since the SL is held at 0.8 V, 1 V was applied at the BL to read from the RRAM cell. When WL is activated, the voltage across the 1T1R cell (BL–SL) is 0.2 V and I_{read} was 3/20/30/40 μA, depending on the programmed state (The RRAM was programmed to different states using the modified Stanford-PKU RRAM model presented in [5]). The drawn current is mirrored and N_2 will sink I_{read} when connected to V_{DD}. In this manner, the current to be sensed is separated from the memory array's influence (wire parasitic, array size *etc*) and the design of the sensing circuit is independent of the array size, enabling easy portability.

Above the N_2 transistor is a NAND-like CMOS structure which acts as the control circuit for the discharging and

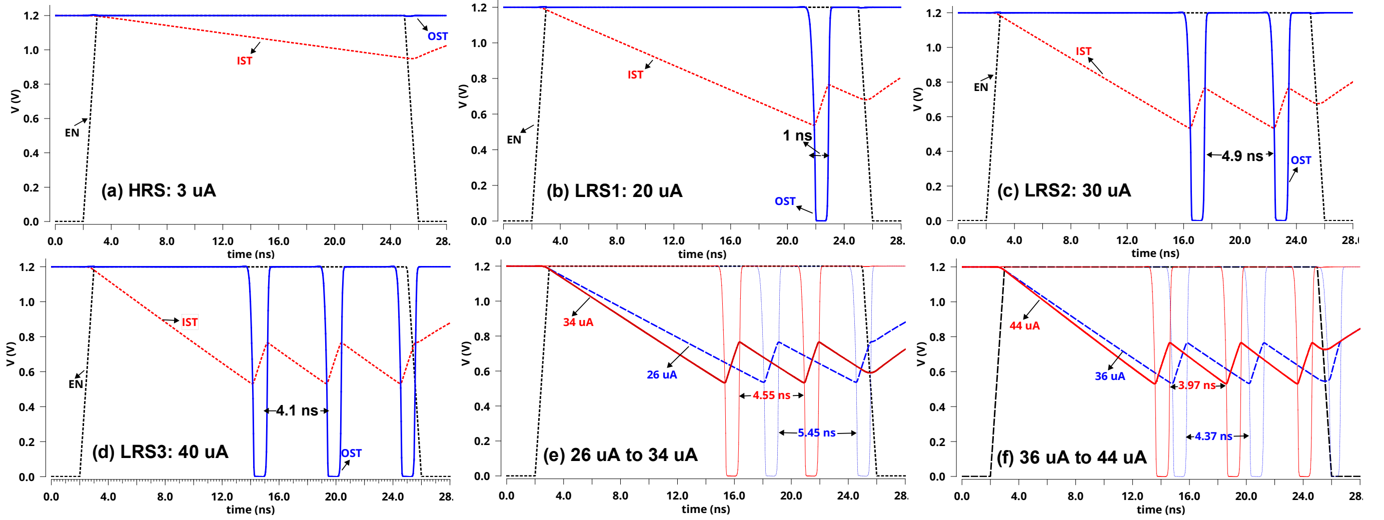

Fig. 2. Simulation results of the proposed circuit in IHP's 130 nm CMOS technology: Current of 3 μA does not produce any pulse (a), while currents of 20, 30 and 40 μA produce one, two and three pulses, respectively (b-d). Since the flip-flop of Fig. 1 is negative edge-triggered, the output of binary counter is '11' for HRS and '00', '01' and '10' for LRS1, LRS2 and LRS3 (e) Currents between 26 μA and 34 μA result in two pulses, tolerating ± 4 μA variations at LRS2 (f) Currents between 36 μA and 44 μA result in three pulses, tolerating ± 4 μA variations at LRS3 and similar tolerance to variations at HRS and LRS1 were verified (not plotted)

charging of the capacitor. Only when both EN and O_{ST} are high, it allows discharging of the capacitor (through N_3, N_4). Throughout the MLC sensing phase, EN must be high. Even when EN is high, if O_{ST} is low, the capacitor is allowed to charge (N_4 is OFF) through $P_{2,3}$. It must be noted that the discharging time is the crucial time which determines the number of pulses. This is because the capacitor is discharged by I_{read}, while it is charged (towards V_{DD}) at a constant time, independent of I_{read}. This necessitates a relatively shorter charging time so as to make the total period of the pulse (discharging+ charging time) proportional to I_{read}. The two PMOS transistors ($P_{2,3}$) serve this purpose and further, their (W/L) was made (450/130) nm to drastically shorten the charging time. All other transistors in Fig. 1 are sized normally, $i.e$ (150/130) nm.

This periodic discharging and charging of the capacitor at I_{ST} is converted to a pulse train by the ST circuit. To minimize hardware, we chose the six-transistor ST circuit of [13]. This compact ST has a fixed V_{TL} of 0.53 V and V_{TH} of 0.76 V in 130 nm technology. The capacitor charges from 0.53 V to 0.76 V in approximately 1 ns, thereby producing a negative pulse of ns duration. Since the binary counter is clocked with this ns wide negative pulse, we need a Flip-flop (FF) which can operate with GHz clock. We chose the extended true single-phase clocked (E-TSPC) FF presented in [14] which is a negative edge-triggered FF capable of operating in GHz frequencies [15], [16].

The value of C and the EN time period (T_{EN}) must be determined judiciously, and, the following observations must be taken into account to ensure accurate sensing in-spite of variations in the RRAM's resistance/I_{read}.

1) The spacing between I_{HRS} and the first LRS, I_{LRS1} is greater than the spacing between neighboring LRS in most MLC RRAMs (Fig. 1-(a)). Hence, the circuit can be designed to produce no pulse for HRS, while LRS1, LRS2 and LRS3 will produce one, two and three pulses, respectively.
2) The amount of RRAM variations tolerated by the sensing circuit will be maximum if the circuit is designed to switch from producing n pulse to n+1 pulse midway between two low resistance states.

We shall describe the design of the proposed MLC sensing circuit to distinguish between the four states of the MLC RRAM manufactured at IHP. Therefore, the circuit must be designed to differentiate between 3 μA (HRS) and 20 μA, 30 μA and 40 μA (three LRS).

$$I = \frac{dQ}{dt} = \frac{d}{dt}(C.V) = C.\frac{dV}{dt} \tag{1}$$

Since the current which discharges the capacitor C is a constant current, I_{read}, the rate at which capacitor voltage decreases is a constant in a given sensing period, given by

$$\frac{dV}{dt} = \frac{I_{read}}{C} \tag{2}$$

making the voltage across the capacitor

$$V_C(t) = V_{DD} - \left(\frac{I_{read}}{C}\right).t \tag{3}$$

When I_{read} is 3 μA, C should not discharge below V_{TL} of the ST for one EN period $i.e$ not even a single negative pulse.

$$V_{DD} - \left(\frac{I_{read}}{C}\right).T_{EN} > V_{TL} \tag{4}$$

When I_{read} is 40 μA, C should discharge and charge three times in one EN period, producing three negative pulses and T_{EN} must be long enough to accommodate them.

$$(T^{dis}_{V_{DD}-V_{TL}} + 2.T^{dis}_{V_{TH}-V_{TL}} + 3.T^{ch}_{V_{TL}-V_{TH}}) < T_{EN} \tag{5}$$

where $T^{dis/ch}_{x-y}$ is the time for the voltage across the capacitor to discharge/charge from voltage x to y. From Eq. 3,

$$T^{dis}_{V_{DD}-V_{TL}} = \left(\frac{C}{I_{read}}\right)(V_{DD} - V_{TL}); \quad (6)$$

$$T^{dis}_{V_{TH}-V_{TL}} = \left(\frac{C}{I_{read}}\right)(V_{TH} - V_{TL}) \quad (7)$$

and the charging time[2] is given by,

$$T^{ch}_{V_{TL}-V_{TH}} = R_{PMOS}.C.ln\left(\frac{V_{DD} - V_{TL}}{V_{DD} - V_{TH}}\right) \quad (8)$$

where R_{PMOS} is the ON resistance of the $P_{2,3}$. Therefore, for a given I_{read}, Eq 4 and 5 can be used to derive C and T_{EN} to satisfy them. Once the boundary conditions (3 μA and 40 μA) are satisfied, it can be verified that the intermediate states, 20 μA and 30 μA will produce one and two negative pulses, respectively. This is because, from Eq. 6, the discharging time is a strong function of I_{read} (other parameters C,V_{TH},V_{TL} fixed) and currents less than 40 μA will have larger discharge time and consequently less pulses in the same period, T_{EN}. To accommodate maximum RRAM variations, the derived C and T_{EN} can be fine-tuned so that the circuit transitions from producing one pulse to two pulses around 25 μA, and from two pulses to three pulses around 35 μA. For IHP's MLC RRAM, C of 0.5 pF and T_{EN} of 24 ns satisfy the requirements to sense the four states with maximum tolerance to variations. Simulation results are plotted in Fig. 2-(a-d). Further, to investigate the tolerance to RRAM variations, I_{read} was varied around the mean current of a state and the output is plotted in Fig. 2-(e,f). From fig. 2-(e,f), one can verify that the sensor tolerates $I_{read} \pm 4$ μA variations. This was achieved by carefully choosing C and T_{EN} to transition from producing n to $n+1$ pulse midway between neighboring I_{LRS} *i.e* the sensing circuit was engineered to transition from one to two pulse approximately at 25 μA, and, from two to three pulse at 35 μA. The conversion of pulse train to bits by the binary counter was also verified by simulations.

III. Improving the Proposed Sensing Methodology by Eliminating the Passive Capacitor

A. Principle

The sensing circuit of Fig. 1 requires a pF capacitor which will be difficult to implement in CMOS - a precise pF capacitance is difficult to design and also occupies more area. Since the sensing circuit has to be area optimized, we replaced the passive capacitor of Fig. 1 with a MOSFET capacitance and redesigned the circuit as depicted in Fig. 3. Since the input capacitance of the MOSFET in 130 nm will be in fF, the current has to be scaled down accordingly to have a similar circuit operation. This is achieved in two stages: first, the magnitude of the read current is reduced by four by proportionately reducing the $READ$ voltage applied across

[2] derived from the observation that a capacitor initially charged to V_{ini} charges towards V_{DD} with the instantaneous voltage, $V_{DD} + (V_{ini} - V_{DD}).e^{\frac{-t}{RC}}$

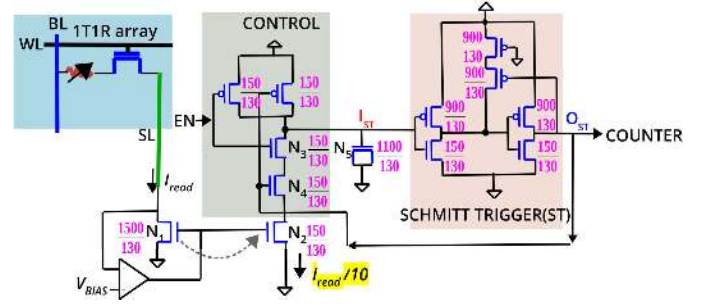

Fig. 3. The sensing circuit proposed in Fig. 1 is modified by replacing it's passive capacitor with a MOSFET's gate capacitance N_5. I_{read} is reduced by reducing the voltage across the RRAM cell and further divided by 10 to achieve a small current which can discharge the fF capacitance of the MOSFET in ns time .

the RRAM cell. This is accomplished by applying 0.85 V to the BL when using V_{BIAS} of 0.8 V, resulting in only 50 mV across the cell (as opposed to 200 mV used in Fig. 1). For IHP's RRAM, this amounts to a read-out current of 0.75 μA (HRS), 5 μA (LRS1), 7.5 μA (LRS2) and 10 μA (LRS3) for the four states. Next, the resulting I_{read} is further scaled by the N_1–N_2 current mirror to $\frac{I_{read}}{10}$ $i.e$ $(\frac{W}{L})_{N_1}$= 10 × $(\frac{W}{L})_{N_2}$. With this current scaling, the capacitance at node I_{ST} can be charged/discharged in ns duration so as to enable the pulses to be captured by the counter at the output of the ST.

B. Circuit design and simulation results

We shall design the circuit of Fig. 3 to sense IHP's MLC RRAM in 130 nm CMOS technology. The capacitance to be charged/discharged is the MOSFET's gate capacitance and hence fixed by the technology. We have little freedom to change the (W/L) and it is set to (1100/130) nm to get a higher capacitance. Therefore, C is fixed and T_{EN} is the only parameter of Equations 4 and 5, which can be designed to sense a given I_{read}. Another undesirable effect of the improved circuit is that the charging time is also considerably reduced due to the reduced capacitance at node I_{ST}. Since the time to charge (from V_{TL} to V_{TH}) determines the width of the negative pulse, the charging time must be long enough (hundreds of ps or a fraction of a ns) to produce a negative pulse wide enough to act as the clock of the counter. To achieve this, the $(\frac{W}{L})_{PMOS}$ was made 6 × $(\frac{W}{L})_{NMOS}$ in the Schmitt Trigger (Fig. 3). Such a sizing of the transistors increased the V_{TH} of the ST, thereby producing a negative pulse of 0.25 ns duration. With the transistors sized as shown in Fig. 3, the circuit was able to produce three negative pulse for LRS3 and no pulse at all for HRS when T_{EN} was 10 ns. Simulation results are plotted in Fig. 4: (a)-(d). Further, to investigate the tolerance to RRAM variations, I_{read} was varied around the mean current of a state and the output is plotted in Fig. 4: (e,f). The circuit is able to tolerate ± 0.75 μA around the mean current of the state.

C. Significance of results

The tolerance to RRAM's variations is reduced in the MOS capacitor circuit $i.e$ ± 0.75 μA as opposed to ± 4 μA. This

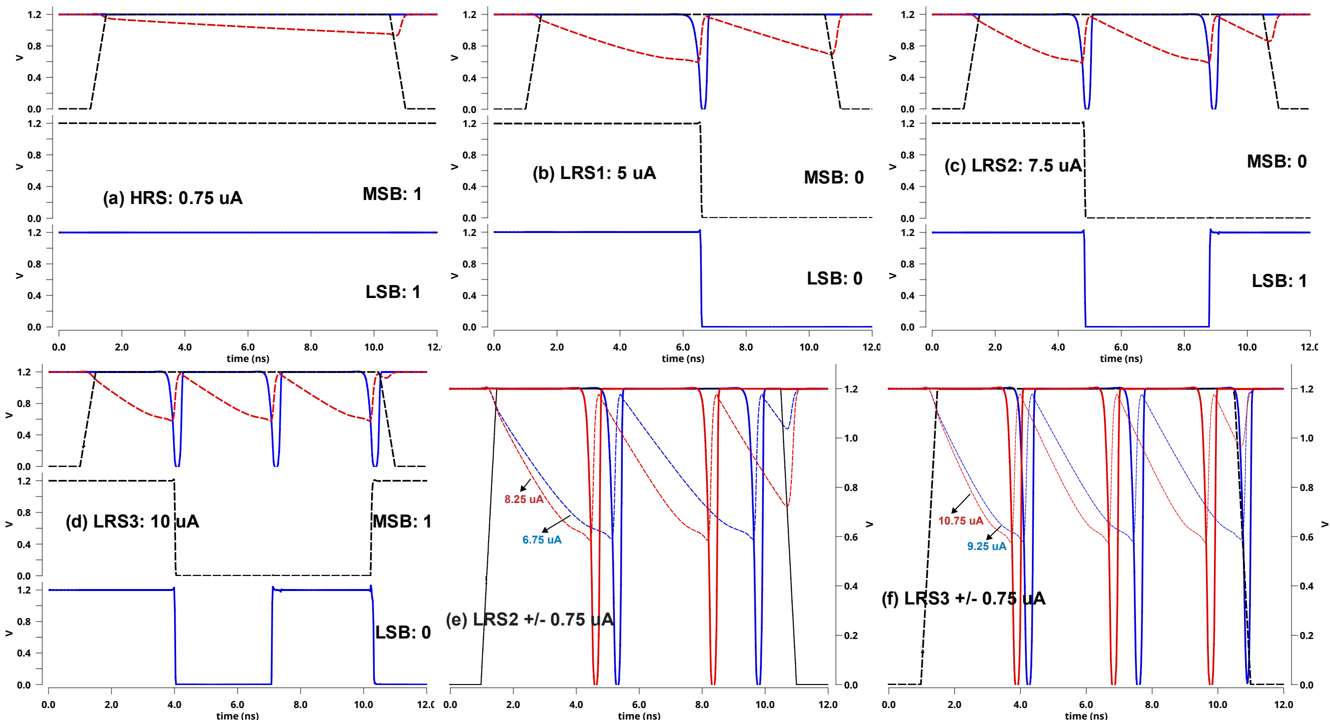

Fig. 4. Simulation results of the capacitor-less circuit (Fig.3) in IHP's 130 nm CMOS technology: Current of 0.75 μA does not produce any pulse (a), while currents of 5, 7.5 and 10 μA produce one, two and three pulses, respectively (b-d). (e) Currents between 6.75 μA and 8.25 μA result in two pulses, tolerating $\pm$ 0.75 μA variations at LRS2 (f) Currents between 9.25 μA and 10.75 μA result in three pulses, tolerating $\pm$ 0.75 μA variations at LRS3

TABLE II
VARIATION TOLERANCE OF THE PROPOSED SENSING CIRCUIT (WITH MOS CAPACITOR)

State/ Mean resistance	Variation tolerated	Percentage
HRS (66.6 KΩ)	$>$ 24.6 KΩ	62%
LRS1 (10 KΩ)	8.69 KΩ–11.76 KΩ	13.1%
LRS2 (6.6 KΩ)	6.06 KΩ–7.4 KΩ	8.2%
LRS3 (5 KΩ)	4.65 KΩ–5.4 KΩ	7%

is because the margin between neighboring LRS states is reduced to 2.5 μA and EN time period is also less, producing closely spaced pulses (Fig.4). However, a tolerance of $\pm$ 0.75 μA is still reasonable since the spacing between neighbouring LRS states is only 2.5 μA. In Table II, the variations in the programmed resistive state is expressed as a percentage of the mean resistance of the state and analyzed for each state. The tolerance % is different for each LRS because the same current tolerance ($\pm$ 0.75 μA) translates as different resistance tolerance since the mean resistance is different. Therefore, the proposed sensing methodology can tolerate 13.1% variations at LRS1 while only 7% variations at LRS3. Interestingly, this reduced tolerance at lower resistances is not a disadvantage in RRAM technology since, the lower the resistance of RRAM cell, less the variations it exhibits. This is because at lower resistance (which is achieved by using a higher compliance current), the increased number of oxygen vacancy defects present in the filament form a well-defined conductive path, thereby exhibiting less variation [4]. For example, for a $TiN/Ti/HfO_x/TiN$ MLC RRAM studied in [17], the variation at LRS1, LRS2 and LRS3 are 12.6%, 3.2% and 2.4 % respectively. For the same device, the variation at HRS is 20.5% and the higher variation at HRS in RRAM technology is attributed to the stochastic nature of filament rupture [18]. In the proposed sensing method, we are able to accommodate the increased variation at HRS (upto 62%) because we assigned the HRS state to zero pulse and the other three LRS states to increasing number of pulses. Therefore, the sensing circuit is able to well tolerate the variations which occur in practical MLC RRAMs.

Table III compares the hardware requirements and speed of the proposed sensing scheme (with MOS capacitor) with conventional schemes. Parallel sensing scheme of [19] requires 1M Ω resistors which will be difficult to fabricate in CMOS. In contrast, the circuit of Fig. 3 does not have any passive element and will occupy less area than the sensing circuit of [19]. Further, our sensor scales well from 2-bit to 3-bit MLC by requiring only one additional flip-flop and logic gate (for 3-bit counter). The serial approach will require 7 I_{REF} [20], while the parallel approach will require 7 op-amps and resistors [19] to sense 3-bits/cell.

IV. CONCLUSION

We have proposed a time-based sensing circuit for MLC RRAM which achieves a trade-off between the sequential and

TABLE III
COMPARISON OF SENSING SCHEMES FOR 2 BITS/CELL

Scheme	Hardware Complexity	Latency	Ref
Sequential	Single SA (10 transistor) + 3 I_{REF} and control circuitry to switch between the I_{REF}	3 steps	[20]
Parallel	3 op-amps + 3 resistors + 3 diodes	1 step	[19]
Time-based	1 op-amp + 37 transistors (13 transistors in Fig.3 + 24 transistors in counter)	1 step	this work

parallel sensing mechanisms conventionally used for multi-level memories. The proposed scheme is faster than sequential approach since it senses in a single step and does not require multiple comparisons. The proposed sensing scheme requires less hardware than parallel sensing which uses multiple operational amplifiers in parallel. The time-based sensing circuit tolerates RRAM variations in accordance with the sensed resistance $i.e$ it tolerates more variations at HRS and less and less variations at lower resistances. Such variation tolerance aligns well with RRAM technology whose MLC cells exhibit more variations at HRS. Tolerance to CMOS process variations were studied and found to be reasonable, but the circuit is sensitive to transistor mismatch, which needs to be improved (current-mirror formed by N_1-N_2). The proposed sensing circuit does not require an absolute reference, which conventional Sense-Amplifier based read techniques employ, obviating the need to generate many precise current/voltage references on-chip.

V. ACKNOWLEDGMENT

This work was funded by Deutsche Forschungsgemeinschaft (DFG)- IMBRA (Project number 389549790)

REFERENCES

[1] J. Reuben, R. Ben-Hur, N. Wald, N. Talati, A. Ali, P.-E. Gaillardon, and S. Kvatinsky, “Memristive logic: A framework for evaluation and comparison,” in *Power And Timing Modeling, Optimization and Simulation (PATMOS)*, September 2017, pp. 1–8.

[2] J. Reuben, R. Ben-Hur, N. Wald, N. Talati, A.Ali, P.-E. Gaillardon, and S. Kvatinsky, “A taxonomy and evaluation framework for memristive logic,” in *Handbook of Memristor Networks*, A. Adamatzky, L. Chua, and G. Sirakoulis, Eds. Springer International Publishing, 2019.

[3] N. Talati, R. Ben-Hur, N. Wald, A. Haj-Ali, J. Reuben, and S. Kvatinsky, *mMPU—A Real Processing-in-Memory Architecture to Combat the von Neumann Bottleneck*. Singapore: Springer Singapore, 2020, pp. 191–213.

[4] A. Prakash and H. Hwang, “Multilevel cell storage and resistance variability in resistive random access memory,” *Physical Sciences Reviews*, vol. 1, no. 6, pp. –, 2016.

[5] J. Reuben, D. Fey, and C. Wenger, “A modeling methodology for resistive ram based on stanford-pku model with extended multilevel capability,” *IEEE Transactions on Nanotechnology*, vol. 18, pp. 647–656, 2019.

[6] E. Prez, C. Zambelli, M. K. Mahadevaiah, P. Olivo, and C. Wenger, “Toward reliable multi-level operation in rram arrays: Improving post-algorithm stability and assessing endurance/data retention,” *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 740–747, 2019.

[7] A. Prakash, D. Deleruyelle, J. Song, M. Bocquet, and H. Hwang, “Resistance controllability and variability improvement in a taox-based resistive memory for multilevel storage application,” *Applied Physics Letters*, vol. 106, no. 23, p. 233104, 2015.

[8] A. Prakash, J. Park, J. Song, J. Woo, E. Cha, and H. Hwang, “Demonstration of low power 3-bit multilevel cell characteristics in a taox-based rram by stack engineering,” *IEEE Electron Device Letters*, vol. 36, no. 1, pp. 32–34, Jan 2015.

[9] S. H. Misha, N. Tamanna, J. Woo, S. Lee, J. Song, J. Park, S. Lim, J. Park, and H. Hwang, “Effect of nitrogen doping on variability of taox -rram for low-power 3-bit mlc applications,” *ECS Solid State Letters*, vol. 4, no. 3, pp. P25–P28, 2015.

[10] C. Calligaro, V. Daniele, R. Gastaldi, A. Manstretta, and G. Torelli, “A new serial sensing approach for multistorage non-volatile memories,” in *Records of the 1995 IEEE International Workshop on Memory Technology, Design and Testing*, 1995, pp. 21–26.

[11] W.-S. Khwa, D. Lu, C.-M. Dou, and M.-F. Chang, *Emerging NVM Circuit Techniques and Implementations for Energy-Efficient Systems*. Cham: Springer International Publishing, 2019, pp. 85–132.

[12] W. Bae, K. J. Yoon, T. Song, and B. Nikolic, “A variation-tolerant, sneak-current-compensated readout scheme for cross-point memory based on two-port sensing technique,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 12, pp. 1839–1843, Dec 2018, doi: 10.1109/TCSII.2018.2868460.

[13] Z. Wang, “Cmos adjustable schmitt triggers,” *IEEE Transactions on Instrumentation and Measurement*, vol. 40, no. 3, pp. 601–605, June 1991.

[14] X. P. Yu, M. A. Do, W. M. Lim, K. S. Yeo, and J. . Ma, “Design and optimization of the extended true single-phase clock-based prescaler,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 11, pp. 3828–3835, Nov 2006.

[15] J. Reuben, Z. V. Mohammed, and H. M. Kittur, “Low power, high speed hybrid clock divider circuit,” in *2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT)*, March 2013, pp. 935–941.

[16] J. Reuben, H. M. Kittur, and M. Shoaib, “A novel clock generation algorithm for system-on-chip based on least common multiple,” *Computers & Electrical Engineering*, vol. 40, no. 7, pp. 2113 – 2125, 2014.

[17] A. Prakash, J. Park, J. Song, S. Lim, J. Park, J. Woo, E. Cha, and H. Hwang, “Multi-state resistance switching and variability analysis of hfox based rram for ultra-high density memory applications,” in *2015 International Symposium on Next-Generation Electronics (ISNE)*, May 2015, pp. 1–2.

[18] E. Ambrosi, A. Bricalli, M. Laudato, and D. Ielmini, “Impact of oxide and electrode materials on the switching characteristics of oxide reram devices,” *Faraday Discuss.*, vol. 213, pp. 87–98, 2019.

[19] Y. Yilmaz and P. Mazumder, “A drift-tolerant read/write scheme for multilevel memristor memory,” *IEEE Transactions on Nanotechnology*, vol. 16, no. 6, pp. 1016–1027, Nov 2017.

[20] B. Q. Le, A. Grossi, E. Vianello, T. Wu, G. Lama, E. Beigne, H. . P. Wong, and S. Mitra, “Resistive ram with multiple bits per cell: Array-level demonstration of 3 bits per cell,” *IEEE Transactions on Electron Devices*, vol. 66, no. 1, pp. 641–646, Jan 2019.