

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

# A 1Mb RRAM Macro with 9.8ns Read Access Time Utilizing Dynamic Reference Voltage for Reliable Sensing Operation

Junjie Mu, *Member, IEEE*, Lu Lu, *Member, IEEE*, Ju Eon Kim, *Member, IEEE*, Byungkwon An, *Graduate Student Member, IEEE*, Vishal Sharma, *Member, IEEE*, Arya Jagath Lekshmi, *Member, IEEE*, Putu Andhita Dananjaya, Weng Hong Lai, Wen Siang Lew, Tony Tae-Hyoung Kim, *Senior Member, IEEE*

**Abstract**—Resistive RAM (RRAM) has emerged as a promising candidate for the next generation of non-volatile memories (NVMs) due to its low write voltage and compact area that is compatible with CMOS technology. In this work, we propose a 1Mb macro consisting of two 512Kb sub-arrays, with the macro area reduced by implementing a common source-line (SL) structure. A voltage-mode sense amplifier (VSA) is designed to overcome the challenge of the low ratio between high-resistance and low-resistance states (R-ratio). Two columns of replica cells situated in the center of the RRAM array are used to generate the reference voltages, which exhibit behavior that closely tracks the changes in the bit-line (BL) voltage even under PVT variations. Additionally, power-gating and common-mode feedback circuits are implemented in the SA to save power and improve sensing speed. The test chip, fabricated in 40nm CMOS technology, occupies a core area of 0.996mm<sup>2</sup>. Compared to the prior work, the array density (5.58Mb/mm<sup>2</sup>) improves by 1.12×. The read access time is 9.8ns with a read precharge voltage of 0.3 V.

**Index Terms**—RRAM, non-volatile memory, R-ratio, voltage-mode sense amplifier, replica cells, read access time.

## I. INTRODUCTION

FLASH memory has long been the dominant non-volatile memory (NVM) technology due to high-density data storage capabilities until at least the 2X-nm generation [1-2]. However, the intrinsic physical characteristics of flash memory, such as high leakage current and slow write speed, impose significant limitations on its further development beyond certain technical nodes. As a result, there is a growing demand for the exploration and creation of novel NVM

technologies that can achieve high density, low power, and fast operation.

Several emerging NVMs such as magnetoresistive random-access memory (MRAM) [3-4], ferroelectric RAM (FERAM) [5], phase-change RAM (PCRAM) [6-7], and resistive RAM (RRAM) [8-18] have shown promise in terms of improved write speed and reduced power consumption compared to embedded flash memory. However, FERAM faces challenges of limited scalability, while PCRAM suffers from high energy consumption for write operation. RRAM has gained significant attention among these emerging NVM technologies due to its favorable characteristics, such as low operating voltage (<3V), fast switching speed (<10ns), compact area (<10nm), and excellent compatibility with CMOS technology [2]. These advantages make RRAM an appealing candidate for high-density and high-speed NVM applications.

In this work, we propose a 1Mb RRAM macro based on a 1T1R structure, aiming to achieve rapid read access to the RRAM. A two-stage voltage-mode sense amplifier (VSA) with power-gating and common-mode feedback circuits is implemented to reduce power and improve read speed. Besides, replica columns are utilized to generate reference voltage, enabling dynamic tracking of the bit-line (BL) voltage across various PVT variations even at a low read voltage.

## II. BACKGROUND

### A. RRAM Bitcell and Basic Operations

RRAM is a metal-insulator-metal (MIM) structure with the ability to switch between a high-resistance state (HRS) and a low-resistance state (LRS), depending on the direction of the current passing through it. The 1T1R bitcell comprises one

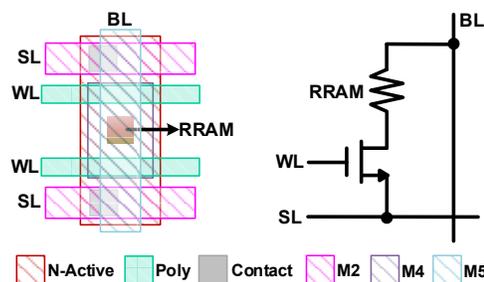


Fig. 1. RRAM bitcell layout (left) and schematic (right).

Manuscript received XX; revised XX; This work was supported by the RIE2020 A\*STAR AME IAF-ICP under Grant I1801E0030. This brief was recommended by XX. (Corresponding author: Tony Tae-Hyoung Kim)

Junjie Mu, Lu Lu, Ju Eon Kim, Byungkwon An, Vishal Sharma, Arya Jagath Lekshmi, and Tony Tae-Hyoung Kim are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail: junjie003@e.ntu.edu.sg; Liu010@e.ntu.edu.sg; je01911@gmail.com; e190209@e.ntu.edu.sg; vishalfzd@gmail.com; aryalekshmi.jagath@ntu.edu.sg; thkim@ntu.edu.sg).

Putu Andhita Dananjaya, Weng Hong Lai, and Wen Siang Lew are with the School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore (e-mail: putu.ad@ntu.edu.sg; wenghong.lai@ntu.edu.sg; wensiang@ntu.edu.sg).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>

Digital Object Identifier XX

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

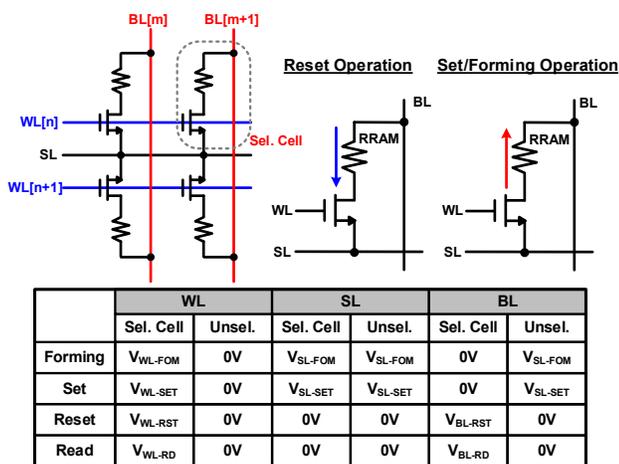


Fig. 2. RRAM array structure and basic operations.

NMOS transistor and an RRAM connected in series as shown in Fig. 1. In the configuration, the access transistor serves to select and control the RRAM by enabling the word-line (WL), while the RRAM works as a storage device by switching the state between HRS (“0”) and LRS (“1”). To reduce the area overhead of individual bitcell, we have implemented a split gate and common source-line (SL) design [10], resulting in a compact bitcell area of  $0.085\mu\text{m}^2$ . The voltage drift due to the resistance at the SL terminal can be mitigated by wider metal tracks.

Fig. 2 illustrates the fundamental operations performed on the RRAM array. Prior to programming, a forming operation is conducted on fresh cells to induce soft breakdown of the dielectric by applying high forming voltages ( $V_{WL-FOM}$  and  $V_{SL-FOM}$ ) on WL and SL and grounding the BL. It is a one-time process to switch the RRAM devices from the initial HRS to LRS. For unselected cells within the same row, an equally high voltage ( $V_{SL-FOM}$ ) is applied to the BL, ensuring that no voltage difference is imposed across the unselected RRAM devices and they remain in their initial states. The transition from HRS to LRS, corresponding to writing “1”, is referred to as “Set”. Similar to the forming operation, a bias voltage of reduced magnitude in comparison to the forming voltage is applied to the SL. The opposite process (from LRS to HRS), known as

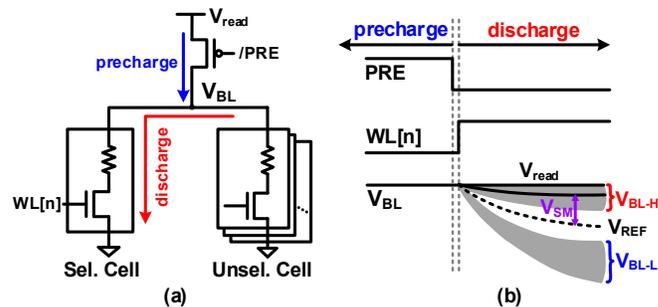


Fig. 3. (a). Simplified read operation in one column and (b). read operation waveform.

“Reset”, is implemented by applying a write voltage to the BL. Simultaneously, the BLs of unselected cells are grounded. During the read operation, the BL is precharged with a low voltage to avoid unwanted partial resetting of the RRAM and state flipping.

### B. Challenges for Reading RRAM at Low Voltage

The scaling of device sizes in RRAM technology has resulted in increased variations in both write time and cell resistance [15]. To improve production yields and reduce the required bias voltages for write operations, efforts have been made to increase the low resistance of RRAM ( $R_L$ ), resulting in low R-ratio ( $R_H/R_L$ ) and RRAM macro design challenges, such as small sensing margin, low read voltage, and slow read speed. The current-mode sense amplifiers (CSAs) have demonstrated faster read compared to VSAs at normal operation voltage [16]. While the CSAs are challenged by insufficient sense margin and requirement for large voltage headroom at lower read voltage due to the reduced R-ratio.

Fig. 3 shows the operation of VSA. During the precharge phase, the BL is charged to the read voltage ( $V_{read}$ ). Following the activation of the access transistor, a current flows through the RRAM, inducing a voltage drop across the BL depending on the state of the selected RRAM. The RRAM in HRS results in a minor voltage drop ( $V_{read}-V_{BL-H}$ ), while the LRS corresponds to a significant voltage swing. For conventional

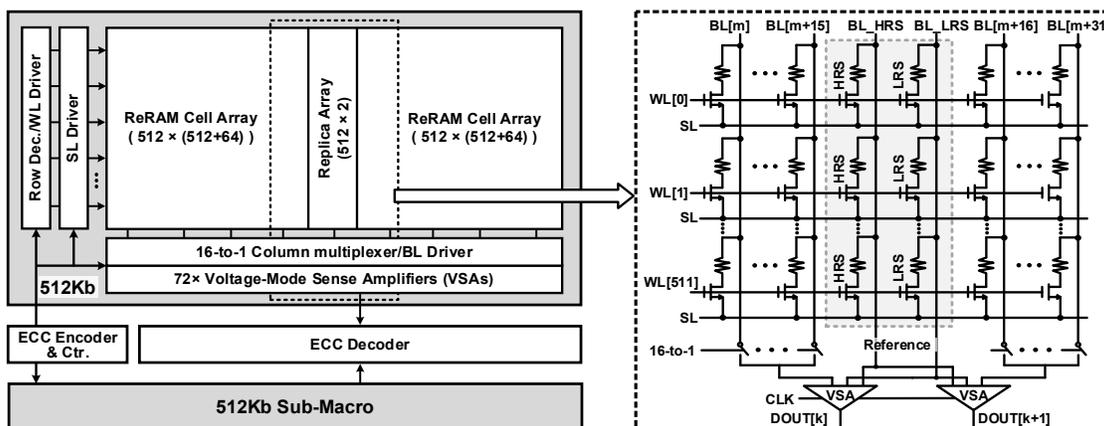


Fig. 4. Overall architecture of the 1Mb RRAM macro.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

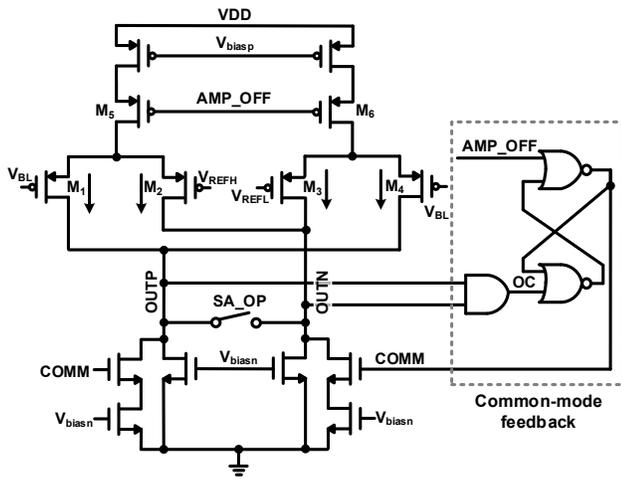


Fig. 5. Schematic of differential-difference amplifier with power gating and common-mode feedback.

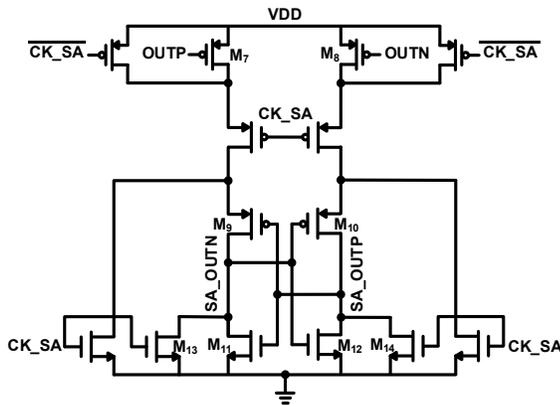


Fig. 6. Schematic of StrongARM latch.

VSAs, the optimal reference voltage is the midpoint  $((V_{BL-H} + V_{BL-L})/2)$  of the voltage drop across BL for the high and low resistive states. The sensing margin ( $V_{SM}$ ) is expressed as follows, where  $V_{BL}$  is the voltage level of the selected BL:

$$V_{SM} = V_{BL} - \frac{V_{BL-H} + V_{BL-L}}{2} \quad (1)$$

However, due to process variations, the distribution of  $V_{BL-H}$  and  $V_{BL-L}$  exhibits a wider range, resulting in a reduction in the sensing margin. Hence, generating an accurate reference voltage poses a significant challenge. In this work, the reference voltage is generated through a pair of replica cells that can dynamically track the behavior of BLs even under PVT variation.

### III. PROPOSED RRAM MACRO

#### A. Overall Architecture

Fig. 4 (left) shows the overall architecture of the 1Mb RRAM macro, comprising two 512Kb sub-macro, an error checking code (ECC) encoder, and an ECC decoder. Each 512Kb sub-macro consists of a RRAM subarray, a row decoder, a column decoder, WL/SL/BL drivers, and  $72 \times$  VSAs. Note that only one sub-macro is activated for each operation to reduce power overhead. The RRAM subarray consists of  $512 \times 1152$  RRAM

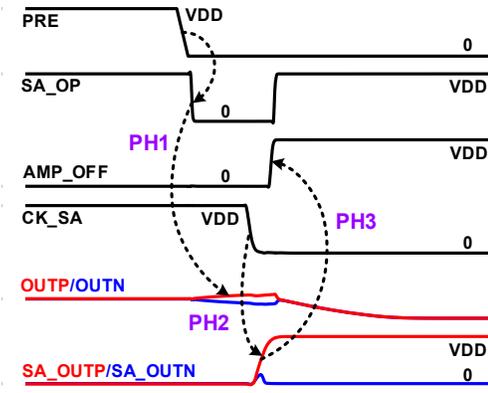
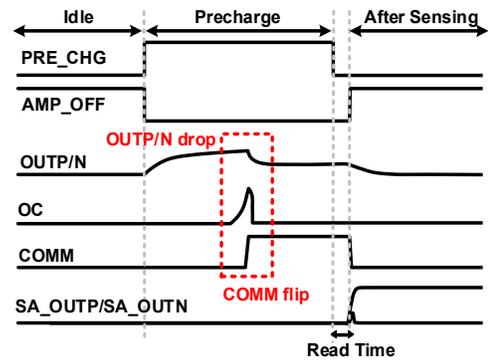


Fig. 7. Waveform of sense amplifier.



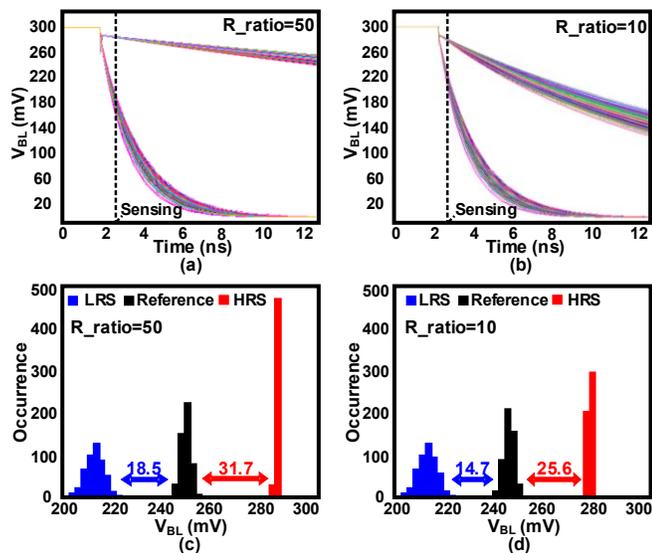
Operation	AMP_OFF	OUTP/OUTN	OC	COMM
Idle	High	Near ground	Low	Low
Precharge	Low	$< V_{DD} -  V_{thp} $	Low	Low (hold)
	Low	$> V_{DD} -  V_{thp} $	High	High
After sensing	High	$< V_{DD} -  V_{thp} $	Low	Low

Fig. 8. Operation and waveform of the common-mode feedback circuit.

cells, with two columns dedicated to replica cells that generate the reference voltages, and 128 columns reserved for storing parity bits. The replica cells are positioned at the center of the RRAM array to minimize variations. The 72 bits (64-bit data and 8-bit for parity check) of data are read out in a single cycle through the SAs using a 16-to-1 multiplexer.

Fig. 4 (right) demonstrates the diagram between the RRAM array and SAs. A 16-column RRAM array is connected to the input of the SA through a 16-to-1 multiplexer. The other inputs of the SA are the reference voltages generated by two columns of replica cells located in the center of the RRAM array, which are shared among 72 SAs. The first column of the two adjacent replica cells is programmed to HRS, while the other column is programmed to LRS. Consequently, these two columns of replica cells generate respective reference voltages, represented as  $V_{REFH}$  and  $V_{REFL}$ . During the read operation, the reference cells in the selected row are activated. The BLs voltages in the replica cells gradually decrease with a similar changing trend as

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <



**Fig. 9.** Simulated transient BL voltages during the read operation when the R-ratio is (a) 50 and (b) 10. (c) and (d) statistics the distribution of BL and reference voltages.

the BLs in the RRAM cells, attributable to the identical dimensions of the RRAM and NMOS transistors in both the replica and RRAM cells. Note that the two columns of replica cells are independent of the other BLs and do not affect the voltage variation of BLs. Then the SAs compare the voltage levels of the input signals and subsequently generate an output based on this comparison.

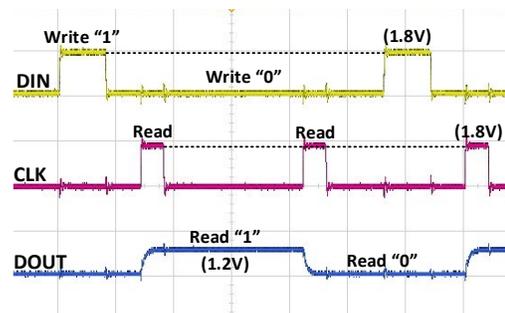
### B. Proposed Two-Stage Voltage-Mode Sense Amplifier

The proposed VSA consists of a differential-difference amplifier (the first stage) and a StrongARM latch (the second stage), as shown in Fig. 5 and Fig. 6. The differential-difference amplifier with four input transistors ( $M_1$ - $M_4$ ) compares the BL voltage with reference voltages and amplifies the voltage difference. The StrongARM latch comprising a differential input pair ( $M_7$ - $M_8$ ) and two cross-coupled pairs ( $M_9$ - $M_{12}$ ) produces rail-to-rail outputs in response to the polarity of input difference ( $V_{OUTP}$ - $V_{OUTN}$ ). The operation of the SA is divided into three phases, as depicted in Fig. 7.

In Phase-1, after discharge, signal  $SA\_OP$  drops to 0, enabling the amplifier to detect the voltage difference between the  $V_{REFH}/V_{REFL}$  and  $V_{BL}$ . For the selected bitcell in the LRS, the current following  $M_3$  and  $M_4$  are similar because their input nodes are connected to low-resistance RRAM. The current flowing through  $M_1$  exceeds that flowing through  $M_2$  since  $V_{BL}$  is lower than  $V_{REFH}$ . Then, the voltage at the  $OUTP$  node is higher than that at the  $OUTN$  node, resulting in a voltage difference expressed by equation (2), where  $A_v$  is the voltage gain of the amplifier.

$$V_{out1} = A_v((V_{BL} - V_{REFH}) - (V_{REFL} - V_{BL})) \quad (2)$$

In contrast, when the RRAM cells are in the HRS, the voltage level at the  $OUTP$  node is lower compared to that at the  $OUTN$  node. Note that PMOS transistors are utilized as input



**Fig. 10.** Sample measure waveforms of the 1Mb RRAM macro.

transistors owing to the low precharge read voltage (300mV) generated by a voltage divider consisting of diode-connected MOSFETs.

In Phase-2, the outputs ( $SA\_OUTN$  and  $SA\_OUTP$ ) initially discharge to the ground through  $M_{13}$  and  $M_{14}$ . Following this, when the signal  $CK\_SA$  goes to ground, the cross-coupled pair amplifies the input difference and pulls the output  $SA\_OUTP$  up (or down) to VDD (or VSS) if  $OUTP$  is higher (or lower) than  $OUTN$ .

The first-stage amplifier still consumes power due to the static current after detection. In Phase-3, a power-gating circuit is implemented to save power by disconnecting the current path from the supply voltage to the ground. An additional pair of PMOS transistors ( $M_5$ - $M_6$ ) is introduced between the current source transistor and four input transistors. Once the second-stage latch produces the outputs, the signal  $AMP\_OFF$  rises to "1", initiating the disconnection of the current path.

The amplitude of output voltages in the first stage affects the sensing time of the second stage. To address this issue, a common-mode feedback circuit is designed, as illustrated in Fig. 5, in which feedback logic is incorporated to regulate the output voltage range by controlling common-mode input ( $COMM$ ). Fig. 8 presents the operation and waveform of the common-mode feedback circuit. In the idle state (when  $AMP\_OFF$  is high), the common mode voltage tends toward the ground, causing  $COMM$  to become low. During the precharge phase, as the common-mode output voltage rises to the threshold ( $V_{DD} - |V_{thp}|$ ),  $COMM$  flips to high voltage. As a result, the current branch controlled by  $COMM$  is activated, leading to a drop in the output voltage, thereby improving the sensing speed of the second-stage comparator. After sensing, the SR latch maintains the  $COMM$  state to prevent oscillation caused by lower output common-mode voltages.

## IV. EVALUATION

Fig. 9 (a) and (b) illustrate the transient BL voltage during the read operation, considering R-ratio values of 50 and 10, respectively, based on 500 Monte Carlo simulations. The BL is initially precharged to 300mV. Subsequently, when the WL is enabled, the BL discharges to a voltage determined by the selected RRAM resistance. In the case of high resistance, the voltage drop on the BL is minimal, while for low resistance, the BL voltage

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

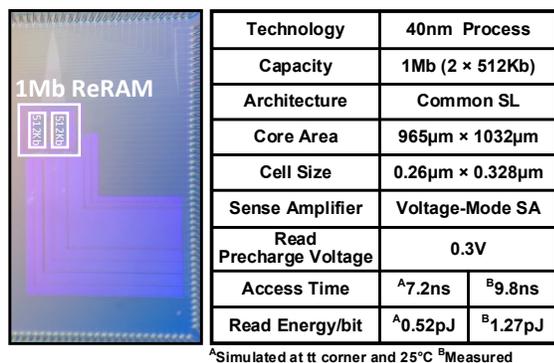


Fig. 11. Die micrograph and summary table.

TABLE I  
COMPARISON TABLE

	ISSCC'18 [10]	ISSCC'19 [10]	VLSI'20 [9]	This work
Technology	40nm	22nm	22nm	40nm
Architecture	Common SL	Split SL	Split SL	Common SL
Cell Size (µm <sup>2</sup> )	0.0848	0.027	0.0257	0.085
Capacity	11Mb	3.6Mb	13.5Mb	1Mb
Density (Mb/mm <sup>2</sup> )	5	10.1	10.24 (excluding analog)	5.58
Read Time (ns)	9ns@0.26V	<5ns@0.7V <10ns@0.5V	6.5ns@0.7V	9.8ns@0.3V

gradually approaches “0”. As the R-ratio decreases, the gap between the BL voltages at two different resistance values gradually narrows. The distribution of BL voltage and reference voltage is statistically analyzed. When the R-ratio is 50, the read margins for the low and high resistances are calculated as 18.5mV and 31.7mV, respectively. As the R-ratio decreases to 10, these margins reduce to 14.7mV and 25.6mV, respectively. By increasing the sensing time, the read margins expand at the expense of slower reading speeds, thereby creating a trade-off between access time and read margin.

Fig. 10 shows the waveforms for RRAM in write and read modes. When input data is “1”, the selected RRAM is set to LRS, resulting in a “1” as the readout data. Likewise, when writing “0”, the RRAM is reset to HRS, and the readout data becomes “0” when the CLK signal is enabled. Fig. 11 showcases the die photograph and summarizes the testchip, revealing that the 1Mb RRAM macro fabricated using 40nm technology occupies a core area of 0.996mm<sup>2</sup>. Table I shows the comparison with prior RRAM macros.

## V. CONCLUSION

In this paper, a 1Mb RRAM macro comprising two 512Kb sub-macro is proposed, which is fabricated using a 40nm process with logic-process compatible RRAM devices. The split gate and common SL are implemented to compact the area of each bitcell, realizing a high array density of 5.58Mb/mm<sup>2</sup>, which results in a 1.12× improvement compared to prior work [10]. The proposed sense amplifier with dynamic reference voltage achieves an access time of 9.8ns with a read precharge voltage of 0.3V.

## REFERENCES

- [1] H. Akinaga, et al., “Resistive Random Access Memory (ReRAM) Based on Metal Oxides,” *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2237-2251, Dec. 2010.
- [2] H.-S. P. Wong, et al., “Metal-oxide RRAM,” *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951-1970, Jun. 2012.
- [3] M. Durlam et al., “A 1-Mbit MRAM based on 1T1MTJ bit cell integrated with copper interconnects,” *IEEE Journal of Solid-State Circuits*, vol. 38, no. 5, pp. 769-773, May 2003.
- [4] Y. Iwata et al., “A 16Mb MRAM with FORK writing scheme and burst modes,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2006, pp. 138-139.
- [5] M. Qazi et al., “A low-voltage 1Mb FeRAM in 0.13µm CMOS featuring time-to-digital sensing for expanded operating margin in scaled CMOS,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2011, pp. 208-210.
- [6] G. D. Sandre et al., “A 90 nm 4 Mb embedded Phase-Change memory with 1.2 V 12 ns read access time and 1 MB/s write throughput,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2010, pp. 268-269.
- [7] Y. N. Hwang et al., “MLC PRAM with SLC write-speed and robust read scheme,” in *Symp. VLSI Tech. Dig. Tech. Papers*, Jun. 2010, pp. 201-202.
- [8] P. Jain et al., “A 3.6Mb 10.1Mb/mm<sup>2</sup> Embedded Non-Volatile ReRAM Macro in 22nm FinFET Technology with Adaptive Forming/Set/Reset Schemes Yielding Down to 0.5V with Sensing Time of 5ns at 0.7V,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)*, Feb. 2019, pp. 212-214.
- [9] C.-C. Chou et al., “A 22nm 96KX144 RRAM Macro with a Self-Tracking Reference and a Low Ripple Charge Pump to Achieve a Configurable Read Window and a Wide Operating Voltage Range,” in *IEEE Symposium on VLSI Circuits*, Jun 2020, pp. 1-2.
- [10] C-C. Chou, et al., “An N40 256K×44 embedded RRAM macro with SL-precharge SA and low-voltage current limiter to improve read and write performance,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)*, Feb. 2018, pp. 478-480.
- [11] Y. Hayakawa, et al., “Highly reliable TaOx ReRAM with centralized filament for 28-nm embedded application,” in *Symp. VLSI Tech. Dig. Tech. Papers*, Jun. 2015, pp. 14-15.
- [12] S.-S. Sheu et al., “A 4 Mb embedded SLC resistive-RAM macro with 7.2 ns read-write random access time and 160 ns MLC-access capability,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2011, pp. 200-201.
- [13] M.-F. Chang et al., “Area-efficient embedded resistive RAM (ReRAM) macros using logic-process Vertical-Parasitic-BJT (VPBJT) switches and read-disturb-free temperature-aware current-mode read scheme,” *IEEE J. Solid-State Circuits*, vol. 49, no. 4, pp. 908-916, Apr. 2014.
- [14] M.-F. Chang et al., “Low VDDmin Swing-Sample-and-Couple Sense Amplifier and Energy-Efficient Self-Boost-Write-Termination Scheme for Embedded ReRAM Macros Against Resistance and Switch-Time Variations,” *IEEE Journal of Solid-State Circuits*, vol. 50, no. 11, pp. 2786-2795, Nov. 2015.
- [15] M.-F. Chang et al., “A Low-Voltage Bulk-Drain-Driven Read Scheme for Sub-0.5 V 4 Mb 65 nm Logic-Process Compatible Embedded Resistive RAM (ReRAM) Macro,” *IEEE Journal Solid-State Circuits*, vol. 48, no. 9, pp.2250-2258, Sept. 2013.
- [16] W. Otsuka et al., “A 4 Mb conductive-bridge resistive memory with 2.3 GB/s read-through and 216 MB/s program throughput,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2011, pp. 210-211.
- [17] M.-F. Chang et al., “A 0.5 V 4 Mb logic-process compatible embedded resistive RAM (ReRAM) in 65 nm CMOS using low voltage current mode sensing scheme with 45 ns random read time,” in *IEEE Int. Solid State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2012, pp. 434-435.
- [18] X.-Y. Xue et al., “A 0.13 m 8 Mb logic based CuSiO resistive memory with self-adaptive yield enhancement and operation for power reduction,” in *Proc. Symp. VLSI Tech.*, Jun. 2012, pp. 42-43.