Two-Direction In-Memory Computing Based on 10T SRAM With Horizontal and Vertical Decoupled Read Ports

Zhiting Lin, *Senior Member, IEEE*, Zhiyong Zhu, Honglan Zhan, Chunyu Peng, Xiulong Wu, Yuan Yao, Jianchao Niu, and Junning Chen

***Abstract*— In-memory computing establishes a new and promising computing paradigm aimed at solving problems caused by the von Neumann bottleneck. It eliminates the need for frequent data transfer between the memory and processing modules and enables the parallel activation of multiple lines. However, vertical data storage is generally required, increasing the implementation complexity for the SRAM writing mode. This article proposes a 10-transistor (10T) SRAM to omit vertical data storage and improve the stability of in-memory computing. A cross-layout of the word line enables arrays with multirow or multicolumn parallel activation to perform vector logic operations in two directions. In addition, the novel horizontal read channel allows matrix transposition. By reconfiguring the data lines, sense amplifiers, and multiplexing read ports, the proposed SRAM can be regarded as a content-addressable memory (CAM), and its symmetry provides selectable data search by column or by row according to the application that easily fits the SRAM storage mode without additional data adjustments. A proposed self-termination structure aims to decrease search energy consumption by approximately 38.5% at 0.9 V at the TT process corner. To verify the effectiveness of the proposed design, a 4 Kb SRAM was implemented in 28-nm CMOS technology. The read margin of the proposed 10T SRAM cell is three times higher than that of the conventional 6-transistor cell. At 0.9 V, logic operations can be performed at approximately 300 MHz, and binary CAM search operations are achieved at approximately 260 MHz with around 1 fJ of energy consumption per search/bit.**

***Index Terms*— Binary content-addressable memory (CAM) (BCAM), in-memory computing, logic operation, SRAM, ternary CAM (TCAM), von Neumann bottleneck.**

Manuscript received October 21, 2020; revised January 4, 2021 and February 17, 2021; accepted February 17, 2021. Date of publication March 9, 2021; date of current version August 26, 2021. This article was approved by Associate Editor Vivek De. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB2202602, in part by the State Key Program of the National Natural Science Foundation of China under Grant 61934005, in part by the National Natural Science Foundation of China under Grant 62074001, in part by the National Science and Technology Major Project under Grant 2017ZX01028–101-003, and in part by the Joint Funds of the National Natural Science Foundation of China under Grant U19A2074. *(Corresponding author: Chunyu Peng.)*

The authors are with the School of Electronics and Information Engineering, Anhui University, Hefei 230601, China (e-mail: cyupeng@ahu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSSC.2021.3061260.

Digital Object Identifier 10.1109/JSSC.2021.3061260

I. Introduction

COMPUTING platforms are mostly based on the von Neumann architecture, which is implemented by separating memory banks and computing elements. The development of artificial intelligence, machine learning, and similar technologies has changed society in many aspects. These changes include the introduction of applications such as speech recognition, facial recognition, and autonomous driving. However, such data-intensive applications impose a burden on the von Neumann architecture. For example, the frequent data exchange between the memory and computing modules leads to increased energy consumption and reduced throughput of the entire computing process due to the limited bandwidth.

As a computing paradigm, that may resolve the von Neumann bottleneck, in-memory computing eliminates the need for large-scale data moving, and as such is attracting widespread attention. A machine learning classifier has been constructed using the standard 6-transistor (6T) SRAM array [1]. In [2], a 10T SRAM cell was used to perform dot products. In [3], a multifunctional memory inference processor was proposed to support various machine learning algorithms.

To implement in-memory computing based on SRAM, a multirows parallel reading technique has been widely adopted [1], [3]–[5]. For this technique, data were stored in columns, increasing the implementation complexity for the SRAM writing mode due to the required processes for additional data moving. Once the data are written, it is difficult to read out by word. The increasing complexity also increases the power consumption and delay. Furthermore, when multiple lines are activated simultaneously, then data destruction and excessive voltage drop in bitline (BL) may occur.

To overcome these challenges, this study proposes a 10T SRAM. The advantages of the proposed structure are as follows.

1) The 10T SRAM mode has two read modes available across rows and columns. These modes allow in-memory matrix transposition, logic operations, and content-addressable memory (CAM) data search without additional data movement.
2) The 10T SRAM cell has horizontal and vertical decoupled read ports, which improves the computation stability (i.e., the stability of stored data during computation)

and the amount of data that can be simultaneously accessed.

3) A self-termination structure dynamically detects the voltage drop and decreases the energy consumption.

The remainder of this article is organized as follows. Section II introduces related work on in-memory computing. Section III details the SRAM mode in the proposed memory structure as well as matrix transposition, in-memory logic operations, and the CAM data search. Section IV provides the measurement results using the proposed design. Finally, the discussion is provided in Section V. The conclusions are provided in Section VI.

II. Related Work

Logic operations form the basis of computer operation and are indispensable for in-memory computing. In this section, we investigate the related work on logic operations.

An improved architecture called an X-SRAM based on an 8-transistor SRAM cell is available for in-memory logic computing [6]. Functioning with peripheral circuits, X-SRAM realizes in-memory vector logic operations including NAND, NOR, and XOR. A 4 + 2-transistor SRAM (cross-coupled inverters + two read access transistors) using deeply depleted channel transistor technology has been proposed [4], where an *n*-well replaces the writing access transistor to write data. In addition, a reconfigurable sense amplifier (SA) enables in-memory AND, OR, and XOR logic operations. Moreover, this SRAM can be reconstructed as either a binary CAM (BCAM) or a ternary CAM (TCAM) for search operations. An SRAM cell for Internet-of-Things applications has been introduced to perform logic operations including AND, OR, and XOR [7].

With the improving monolithic 3-D integration technology, 3-D-SRAM cells have achieved high stability and fast data access while performing in-memory logic operations such as AND/NAND, OR/NOR, and XOR/XNOR [8]. A 9-transistor 3-D computing-in-memory SRAM cell with two layers has been devised by placing a classical 6T SRAM cell and three additional transistors on the bottom and top layers, respectively [9]. The two layers enable bitwise logic operations in one cycle.

The high speed, low power consumption, and excellent durability of emerging nonvolatile memory technologies, such as resistive random access memory (RRAM) [10]–[12] and magnetic random access memory (MRAM) [13]–[16], have fostered in-memory computing and other applications. However, such technologies are still under development.

CAM cells enable parallel search [8], [17]–[20], which is essential for various applications. Previous studies have considered implementing a CAM search using 6T SRAM cells. In [21], a 6T SRAM was reconfigured as either a BCAM or a TCAM, to enable bitwise AND/NOR logic operations and parallel data readout. However, the stability was compromised in this architecture. In [5] and [21], [22], word lines are activated as the search lines, which require the data being searched to be stored in a column. This requires the data to be transposed and then written into the SRAM array. However, storing data in columns is complex to implement with the conventional SRAM writing mode. Consequently, dual-period storage technology was adopted at the expense of additional power consumption, data moving, and delays.

To achieve efficient in-memory computing without the above-mentioned problems, this study considers a flexible symmetric structure that can be configured in different working modes according to the application. This structure supports two-direction in-memory CAM search, thus, the data being searched can be written in rows, similar to that in conventional SRAM, without the need for data transposition.

III. Proposed SRAM and Its Operations

A. Overview of the Proposed SRAM Circuit

This section details the structure of the proposed 10T SRAM, which is shown in Fig. 1. Fig. 1(a) shows the overall architecture. Fig. 1(b) and (c), respectively, depict the 10T SRAM cell and its layout. The cell has three main parts: two cross-coupled inverter memory modules, two access transistor write paths, and two decoupled read paths.

The write word line (WWL) controls the gates of the access transistors N1 and N2. The column-sharing word line (CWL) controls the gate of transistor M1, and storage node Q controls the gate of transistor M2 on the left read path. The row-sharing word line (RWL) controls the gate of transistor M3, and storage node QB controls the gate of transistor M4 on the right read path. CWL and RWL are placed vertically and horizontally in the layout, respectively. Unlike the conventional 6T SRAM cell, two additional horizontal lines (RL and RR) are added to constitute the read paths. The proposed cell avoids read–write conflicts, and transistors M1 to M4 can be set to the minimum size for reducing the cell area overhead.

B. SRAM Mode and In-Memory Matrix Transposition

Table I and Fig. 1(b) show that the data on bit lines BL/BLB are stored when the WWL is activated during writing, similar to the conventional 6T SRAM write method. While reading, two read modes exist. First, the data can be accessed using the row-read mode, as shown in Fig. 2(a), where BLB1/BLB2, RR, and RWL1 are initialized as 1, 0, and 0, respectively. When RWL1 changes to 1, if the storage node Q is 1, similar to Q1, BLB1 is disconnected from the ground; otherwise, similar to Q2, BLB2 is connected to the ground. Therefore, BLB1 voltage remains high, but BLB2 voltage drops. As a result, the data of the same row can be read out through the single-ended SAs of BLBs. Second, the data in a column can be accessed using the column-read mode, as shown in Fig. 2(b), where RL1/RL2, BL, and CWL1 are initialized as 1, 0, and 0, respectively. When CWL1 changes to 1, if the storage node Q is 1, similar to Q1, RL1 is connected to the ground; otherwise, similar to Q2, RL2 is disconnected from the ground. Therefore, RL1 voltage remains high, but RL2 voltage drops. Then, the data in a column can be read out through the single-ended SAs of RLs.

As a basic and important operation, matrix transposition is generally realized by data reading, moving, and writing back in a complicated operation with high power consumption. In contrast, transposition is easily implemented in the proposed SRAM cell using a unique column-read mode. Assuming

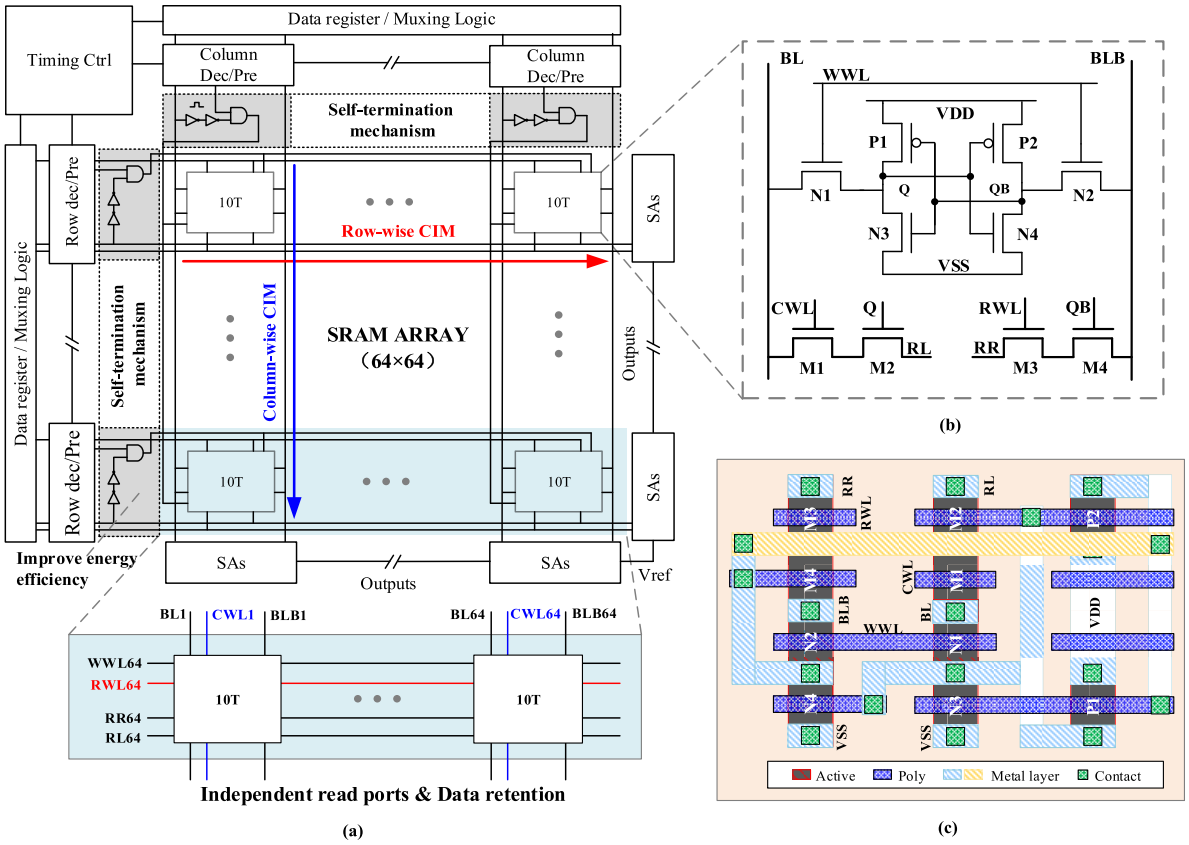


Fig. 1. (a) Overall SRAM architecture. (b) Schematic and (c) layout of the 10T cell.

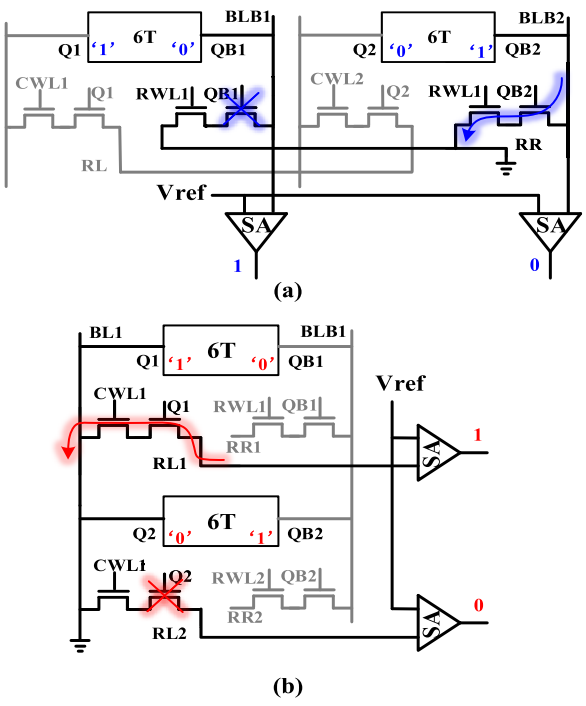


Fig. 2. Two SRAM read modes. (a) Row read. (b) Column read.

that the matrix elements have been written row by row into the array as in conventional SRAM writing, BL, RL, and CWL are initialized as 0, 1, and 0, respectively. When CWL changes to 1, the data of the same column can be read out through the RLs by SAs as shown in Table I. Compared with the conventional read–write method, in-memory matrix transposition prevents intense data moving and substantially improves efficiency. Additionally, the separated read port allows data isolation in the storage node from the external circuit, improving the stability of the SRAM cell.

C. In-Memory Logic Operations

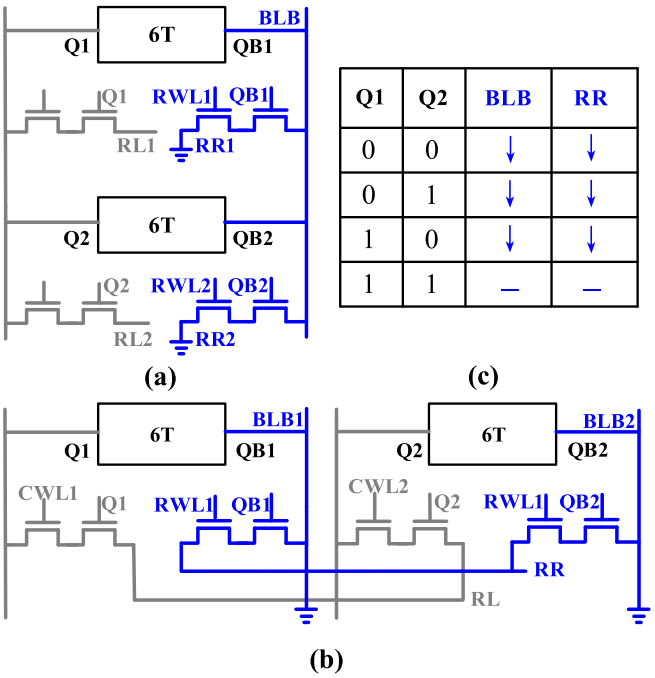
1) In-Memory Logic and Operation: The bitwise AND operation between two bits in different rows is shown in Fig. 3(a). According to the data in storage nodes Q1 and Q2, the output can be obtained from the voltage at BL BLB. Specifically, when RR1, RR2, and BLB are initialized as 0, 0, and 1, then RWL1 and RWL2 change to 1. If both Q1 and Q2 are 1, the BLB voltage remains at the high state because the corresponding paths from BLB to the ground are in an OFF state. In contrast, if Q1 or Q2 is 0, at least one path from BLB to ground is in an ON state, and the BLB is discharged resulting in output 0. This process of logic operations is similar to that of the SRAM read operation. Although Fig. 3(a) only shows the row AND operation of two operands in two rows, the operation of multiple operands in more rows can be activated in a similar manner. That is because the RWLs of two or more rows can be enabled simultaneously.

Similarly, the bitwise AND operation between two words in different columns is also allowed. For example, Fig. 3(b)

TABLE I
DETAILED OPERATION TABLE

Mode		Signal	WWL	CWL	RWL	BL	BLB	RL	RR	Group of SAs	
										First	Second
SRAM	Read	Row	L	L	H	floating	Pre(H)	floating	L	-	BLB
		Column	L	H	L	L	floating	Pre(H)	floating	RL	-
	Write		H	L	L	H(write 1) L(write 0)	L(write 1) H(write 0)	floating	floating	-	-
	Hold		L	L	L	H	H	H	H	-	-
CIM	AND	Row	L	L	H	floating	Pre(H)	L	L	-	BLB
		Column	L	L	H	L	L	floating	Pre(H)	RR	-
	OR	Row	L	H	L	Pre(H)	floating	L	L	-	BL
		Column	L	H	L	L	L	Pre(H)	floating	RL	-
	BCAM	Row	L	L^1/H^2	H	H^1/L^2	L^1/H^2	Pre(H)	Pre(H)	RL/RR	-
		Column	L	H	L^1/H^2	Pre(H)	Pre(H)	H^1/L^2	L^1/H^2	-	BL/BLB
	TCAM	Row	L	L^1/H^2	H	H^1/L^2	L^1/H^2	Pre(H)	Pre(H)	RL/RR	-
		Column	L	H	L^1/H^2	Pre(H)	Pre(H)	H^1/L^2	L^1/H^2	-	BL/BLB

[1]Search bit 1 [2]Search bit 0 - Unused H:VDD L:GND Pre:Pre-charge



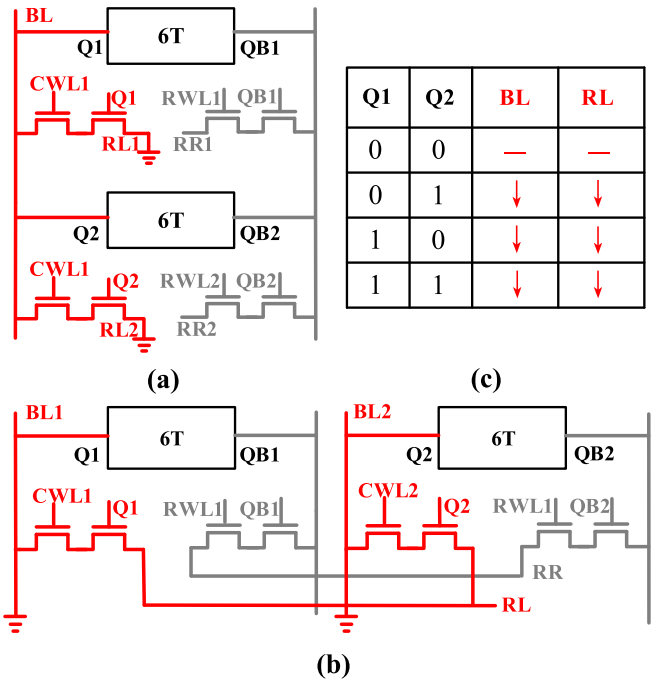
Q1	Q2	BLB	RR
0	0	↓	↓
0	1	↓	↓
1	0	↓	↓
1	1	–	–

Fig. 3. (a) Logic row AND operation between two row bits. (b) Logic column AND operation between two column bits. (c) Discharge conditions of BLB and RR for different Q1 and Q2.

shows the AND operation between two column bits. First, the BLB1, BLB2, and RR are initialized as 0, 0, and 1. The BLBx of the unselected columns that are not involved in the AND operation are pre-charged to high. Then, when the RWL1 changes to 1, the AND operation starts. If both Q1 and Q2 are 1, the path from RR to the ground is disconnected and thus the RR voltage remains high. In contrast, if Q1 or Q2 is 0, at least one path from RR to the ground is formed, and the RR is discharged. Finally, the AND operation result of two bits between different columns can be read out through the single-ended SA of RR. When the RR voltage is discharged to an extremely low value and the QBs of unselected columns are high, the snake current BLBx-M1M2-RR may appear and affect the result of the logic operation. To avoid this situation, a self-termination structure is proposed as detailed in Section III-D4, which can be quickly turn OFF the corresponding RWL.

2) In-Memory Logic or operation: Fig. 4(a) illustrates the bitwise OR operation between two bits in different rows. In particular, the RL1, RL2, and BL are set to 0, 0, and 1 respectively. The RLx of the unselected rows that are not involved in the OR operation is pre-charged to high. Next, the CWL1 is enabled to start the logic OR operation. As shown in Fig. 4(c), for Q1 = 0 and Q2 = 0, the BL voltage remains high, which indicates the result of the OR operation is 0. For Q1 = 1 or Q2 = 1, the BL voltage drops, which indicates the result of the OR operation is 1. The row OR operation can be applied to more than two operands in different rows. When the Qs of the unselected rows are high, the BL voltage may be affected by the RLx of the unselected rows because all CWLs are turned ON. However, because the RLx has been pre-charged to high, this situation can only happen when the BL voltage is too low. To avoid this situation, the self-termination structure is also applied to the row-wise OR operation to turn OFF the corresponding CWL.

The bitwise OR operation between two bits in different columns is shown in Fig. 4(b). The output can be identified



Q1	Q2	BL	RL
0	0	—	—
0	1	↓	↓
1	0	↓	↓
1	1	↓	↓

Fig. 4. (a) Logic row OR operation between two row bits. (b) Logic column OR operation between two column bits. (c) Discharge conditions of BL and RL for different Q1 and Q2.

by whether line RL is discharged or not. Assuming that both Q1 and Q2 are 0, the RL voltage remains at the high state because the path from the RL to ground is in the state OFF, and the output of the SA is 0, making the result of operation OR between Q1 and Q2 to be 0 state. When Q1 or Q2 is 1, the discharge path of RL is activated, and the output of the SA is a 1. By activating the CWLs of multiple columns simultaneously, the data in two or more columns can be selected to perform the logic OR operation.

D. In-Memory CAM Operations

The CAM can be divided into BCAM and TCAM according to the search accuracy. In BCAM, an accurate match can be achieved by comparing the data stored in the memory array and the search data. In contrast, TCAM adds bit state X, whose value is irrelevant. Solutions for both BCAM and TCAM are presented in this article with the NOR-type match lines [23]. In the CAM mode, the cross-coupled inverters are used for storage, and two independent decoupled read paths corresponding to two match lines are used for comparison. The matching results are obtained by combining the two match lines, which differ from the conventional structure with a single match line.

Due to the symmetry of the proposed circuit, BL/BLB and *RL*/RR can be used as search lines or match lines with different configurations. Thus, the proposed SRAM cell structure can be configured for the CAM search in two directions, namely, rows and columns. The signals of the proposed circuit for different operations are shown in Fig. 5. The first four cycles show the writing and row reading of data bits 0 and 1.

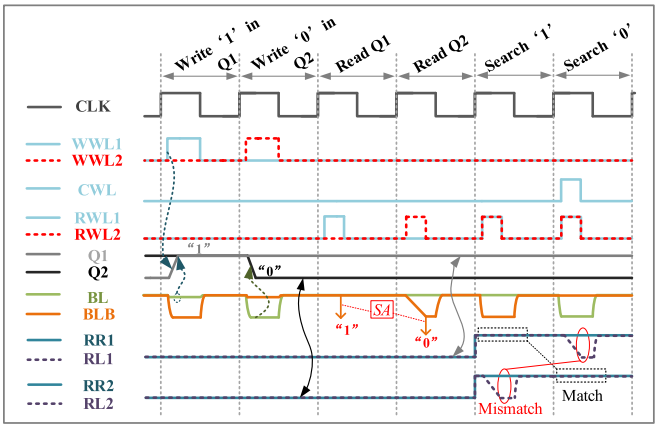


Fig. 5. Timing diagrams of writing, row reading, and BCAM row-wise search.

The fifth and sixth cycles show the row-wise search of bits 0 and 1, respectively.

1) BCAM Row-Wise Search: Row-wise search means that the search data are compared with each datum stored in a row of the memory array, and the address of the internal matched data is returned. In this case, BL/BLB is used as the search line SL/SLB, and *RL*/RR is used as the match line ML/ML'.

Considering a single BCAM cell, RL and RR are first precharged to the high state. If the search datum is 1, BL and BLB are set to 1 and 0, respectively. Otherwise, their values are switched. When RWL is set to 1 and CWL is set to 1 or 0 according to the corresponding SL (BL) search data, the comparison is activated.

When storage node Q is 1, M2 and M4 are turned ON and OFF, respectively, as shown in Fig. 1(b). Assuming that the search datum is 1, match lines RL and RR remain at the high state, which indicates a hit in the operation. In contrast, when storage node Q is 0, the left path containing M1 and M2 is turned OFF, and the right path containing M3 and M4 is turned ON. As a result, RR is discharged, and a mismatch between the data is recognized. When the search datum is 1, i.e., when BL (SL) is 1, RL is not discharged through BL regardless of mismatch or match, only RR is discharged through BLB in case of mismatch. Therefore, CWL can be set to 0 to reduce energy consumption and mitigate interference between the columns. The search for datum 0 is similar to that for datum 1, but the CWL is turned ON to form the discharge path (*RL*-M1M2-BL) in case of mismatch.

As BCAM works in parallel mode, a 4 × 4 simplified array, as that shown in Fig. 6, is used to illustrate the proposed BCAM row-wise search. Only the data in the third row match with the search data, 1010, and the match lines of other rows are discharged. When the search datum and stored datum are 0, the search lines (e.g., SLB2 and SLB4) are 1 and connect with match line ML1'. If the ML1' voltage is very low, SLB2 and SLB4 charge it. To prevent this charging, a self-termination structure is proposed and outlined later in this article.

2) BCAM Column-Wise Search: The symmetry of the proposed SRAM structure removes the constraint of conventional one-way search, enabling column-wise search. This means

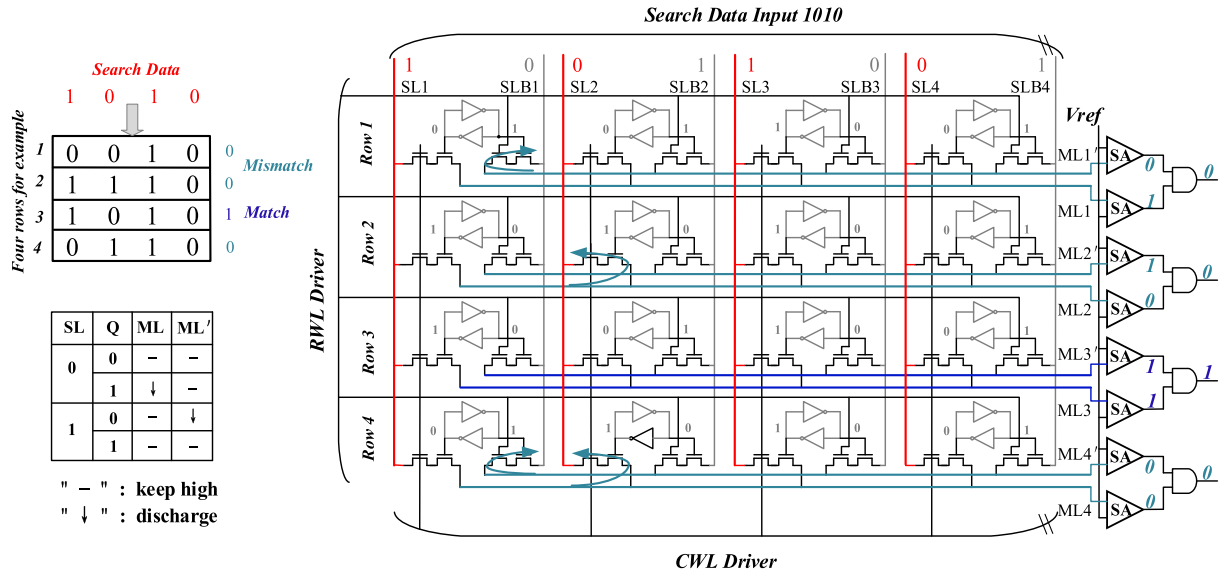


Fig. 6. Example of BCAM row-wise search in 4 × 4 simplified array, where only the data at row three match.

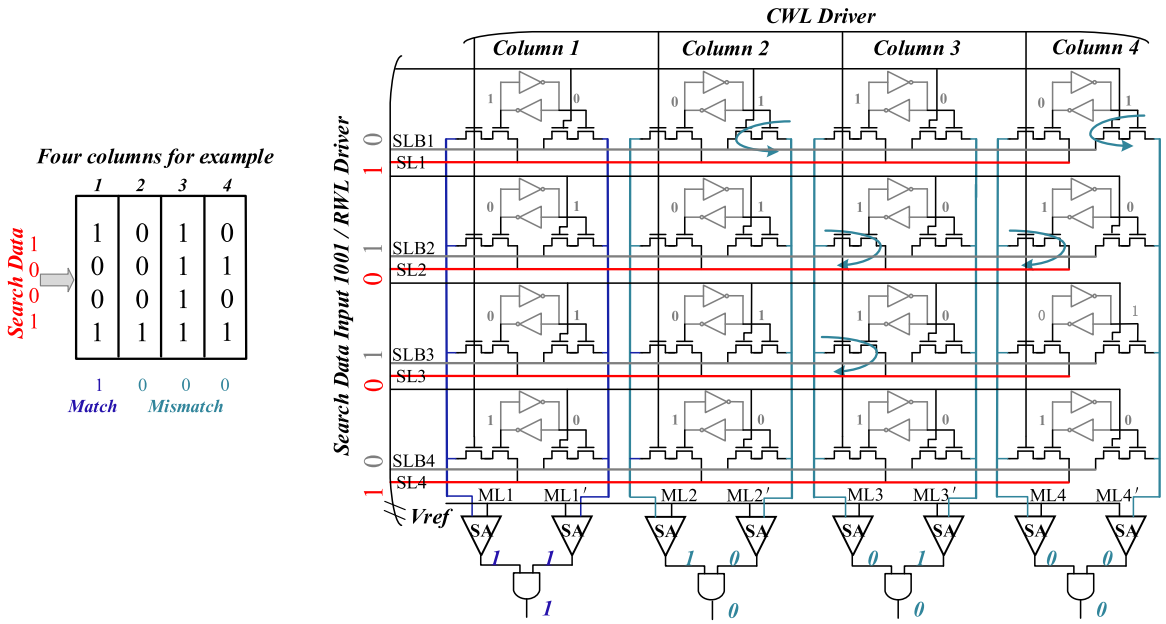


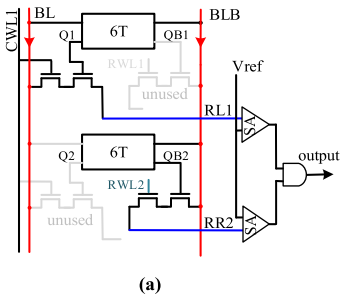
Fig. 7. Example of BCAM column-wise search in 4 × 4 simplified array, where only the data at column one match.

search data are compared with data stored in each column of the memory array, and the address of the matching data is returned. BL/BLB are used as match lines ML/ML', and *RL*/RR are used for data input. Hence, the functions are switched compared with row-wise search. For column-wise search, both BL and BLB are pre-charged to the high state. If the search datum is 1, RL and RR are set to 1 and 0, respectively, otherwise their values are switched to the opposite state.

When CWL is set to 1 and RWL is set to 1 or 0 according to the corresponding SLB (RR) search data, the comparison is activated. When storage node Q is 1, as shown in Fig. 1(b), M2 and M4 are turned ON and OFF, respectively. For search datum 1, match lines BL and BLB remain at the high state, which indicates a hit in the operation. In contrast, when storage node Q is 0, the left path containing M1 and M2 is turned OFF, and the right path containing M3 and M4 is turned ON. As a result, BLB is discharged through RR, and a mismatch is recognized. When the search datum is 0, i.e., when SL (*RL*) is 0 and SLB (RR) is 1, BLB is not discharged through RR regardless of mismatch or match, only BL is discharged through RL in case of mismatch. Therefore, RWL can be set to 0 to reduce energy consumption and mitigate interference between the rows.

In Fig. 7, a 4 × 4 array illustrates column-wise search. Only the data in the first column match with the search data, 1001.

3) TCAM Search: TCAM is different from BCAM because, besides states 0 and 1, undetermined state X is also considered. To represent 0, 1, and X, two SRAM cells are used to construct



BL/BLB	Status	Q1	Q2	RL1	RR1	output
0/1	0	0	0	H	H	1
	X	0	1	H	H	1
	-	1	0	L	H	0
	1	1	1	L	H	0
1/0	0	0	0	H	L	0
	X	0	1	H	H	1
	-	1	0	H	L	0
	1	1	1	H	H	1

H : keep high
L : discharge

(b)

Fig. 8. (a) Example of TCAM row-wise one-bit search. (b) Discharge of RL1 and RR2 for different combinations of Q1 and Q2.

a TCAM block, whose implementation for row-wise search is shown in Fig. 8(a). Nodes Q1 and Q2 being 0 represents state 0, nodes Q1 and Q2 being 1 represent status 1. As shown in Fig. 8(b), for Q1 = 0 and Q2 = 1, the match lines RL1 and RR2 will not discharge regardless of whether the search data is 0 or 1. Hence, it matches with both 1 and 0 of the search data. However, the status of Q1 = 1 and Q2 = 0 does not have this feature and is not allowed in the TCAM mode. Thus, node Q1 being 0 along with node Q2 being 1 represents status X.

For row-wise search, match lines RL1 and RR2 are first pre-charged to the high state. If the search datum is 1, BL and BLB are set to 1 and 0, respectively. Otherwise, BL and BLB are set to 0 and 1, respectively. As shown in Fig. 8(b), if the search data are the same as the stored data, match lines RL1 and RR2 remain at the high state. When the stored state is X, RL1 and RR2 remain at the high state, that is, matching is always returned.

Compared with row-wise search, the TCAM cell for column-wise search consists of two neighboring 10T cells in the same row. Column-wise search proceeds similar to row-wise search, and hence its details are omitted for brevity.

4) Energy Saving Technology: To ensure reliability and to decrease energy consumption, a self-termination structure is proposed. This structure automatically generates the RWL (CWL) signal to control the state of M3 (M1) by detecting the voltage of the match line. As shown in Fig. 9, the structure consists of two asymmetric inverters and a logic gate. The inverters dynamically detect the voltage of match line RR. When the voltage remains at a high level, the self-termination mechanism is idling, that is, RWL follows the RWL' signal. Once the match line is discharged and reaches the flip voltage of the inverter, a cutoff signal is generated. As a result, the pulsewidth of RWL is shortened compared with the width of the RWL' signal. Correspondingly, the discharge path of RR is turned OFF in advance. With a proper flip voltage for the asymmetric inverter, the voltage drop of the match line can be quickly sensed and controlled. Therefore, the power consumption caused by the discharge of the match line can be effectively reduced.

IV. Evaluation Results

The proposed SRAM memory was simulated and fabricated using the TSMC 28-nm complementary metal oxide semiconductor (CMOS) technology with a nominal VDD of 0.9 V.

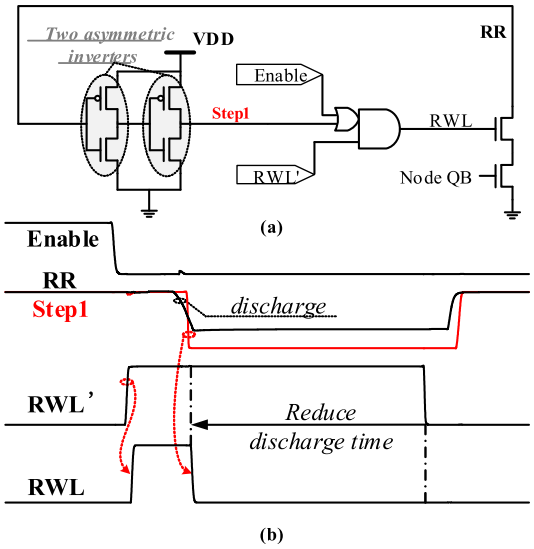


Fig. 9. (a) Schematic of self-termination structure. (b) Timing diagram.

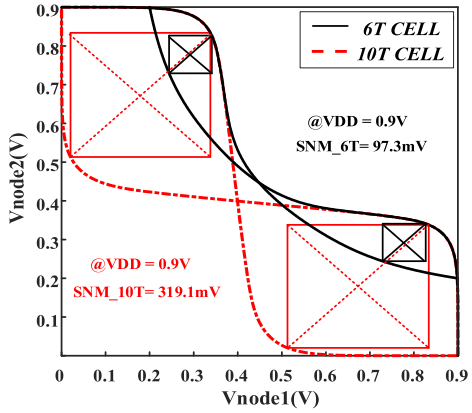


Fig. 10. Comparison of butterfly curves from read static noise margin of proposed 10T SRAM and conventional 6T SRAM cells.

The memory array size was 64 × 64 to obtain a 4 Kb array. Fig. 10 shows the read static noise margin (SNM) of the proposed 10T SRAM cell and conventional 6T SRAM cell at room temperature 27 °C. The read SNM of the proposed 10T cell is 319.1 mV at VDD = 0.9 V, and the conventional 6T SRAM cell shows a read SNM of 97.3 mV at VDD = 0.9 V. The read SNM of the decoupled 10T cell is 3.27 times more than that of conventional 6T cell at the same VDD.

In the BCAM mode, the decoupled read ports were used for search. The search SNM Monte Carlo simulations of different test temperatures, i.e., 27 and 100 °C, were conducted by using ADE XL in Cadence. Fig. 11 shows the results of the 8k-point Monte Carlo simulations at VDD = 0.9 V. Fig. 11(a) reveals that the proposed 10T cell produces 317.8 mV mean search SNM with a standard deviation of 7.2 mV at 27 °C and TT process corner. When the temperature goes to 100 °C, the results indicate that the 10T cell has 295.7-mV mean search SNM with a standard deviation of 6.7 mV at the TT corner in Fig. 11(b). The stability of the cell node decreases as temperature increases, and the decrease in standard deviation

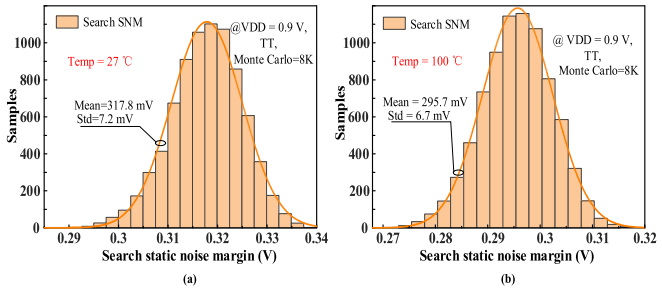


Fig. 11. Search SNM in BCAM mode at (a) 27 °C and (b) 100 °C.

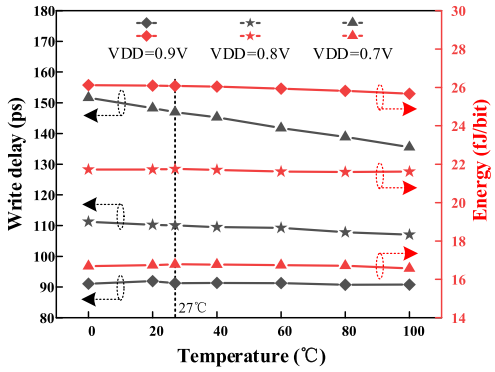


Fig. 12. Post-layout simulation of write delay and energy across temperature (0–100 °C).

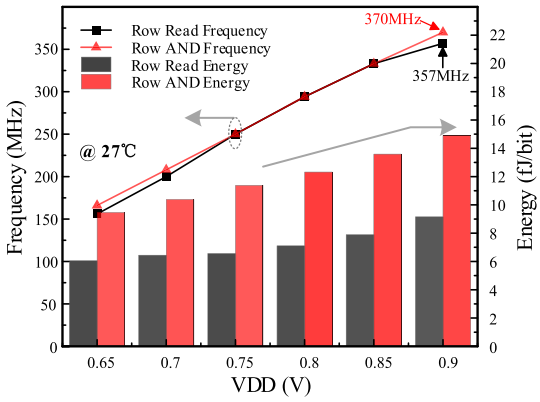


Fig. 13. Simulated frequency and energy consumption of the row AND operation between two words and the row read.

shows that the distribution of search SNM is more concentrated and the influence of the process is relatively weakened.

The delay and energy consumption of the write operation from 0 to 100 °C is shown in Fig. 12. The time interval from WWL voltage rising edge (i.e., 50% of VDD) to cell node voltage establishment (i.e., 90% of VDD) [24] is defined as the write delay. In general, the write delay and energy consumption are less affected by the operating temperature. The lower is the supply voltage, the higher is the write delay and the lower is the write energy consumption; these observations are consistent with the predictions.

Although the row AND operation is similar to the row read operation, the difference is that multiple RWLs are activated in row AND operation. Fig. 13 shows the simulated frequency and

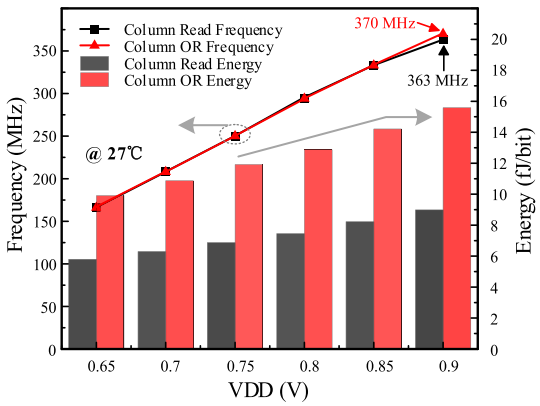


Fig. 14. Simulated frequency and energy consumption of the column OR operation between two words and the column read.

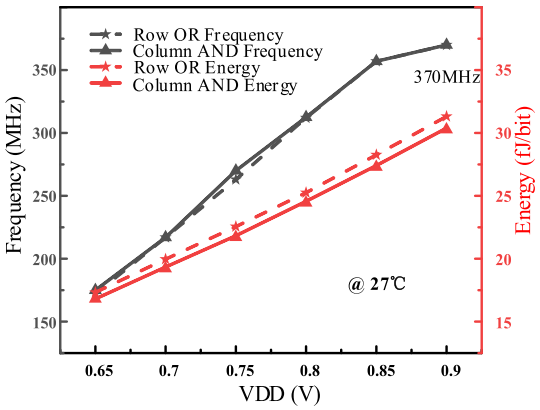


Fig. 15. Simulated frequency and energy consumption of the row OR and column AND operation.

energy consumption of the AND operation between two words in two rows and row read operation at 27 °C. The frequency of both row AND operation is similar to the frequency of row read operation, and the row AND operation consumes approximately twice the energy of that consumed by the row read. This is because the operands of the rowAND operation are two words.

Fig. 14 shows the simulated frequency and energy consumption of the logic column OR operation and column read under different supply voltages (VDD from 0.65 to 0.9 V) at 27 °C. As the supply voltage decreases, the frequency and energy consumption of both operations are reduced. Again, their frequency curves are similar, but the energy consumption of the OR operation is approximately twice that of the column read. Considering the symmetry of the proposed SRAM and the characteristics of different logic operations, the row OR operation is similar to the column AND operation. The simulated frequency and energy consumption of both row OR operation and column AND operation at 27 °C are shown in Fig. 15. The operands are in two rows (row OR) or two columns (column AND). The black curves show that the frequency increases with voltage, as expected. At VDD = 0.9 V, the maximum frequency of the row OR operation is 370 MHz, which is equal to the maximum frequency of the column AND operation.

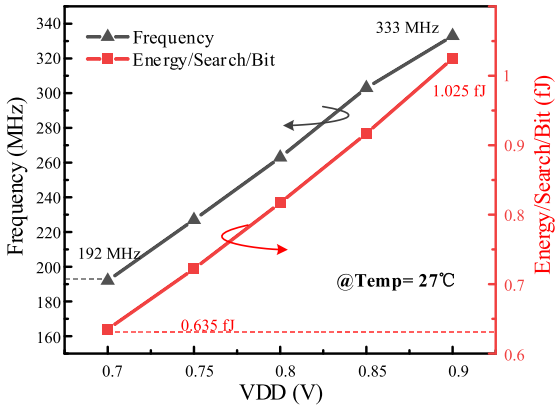


Fig. 16. Simulated frequency and energy consumption of the BCAM row-wise search.

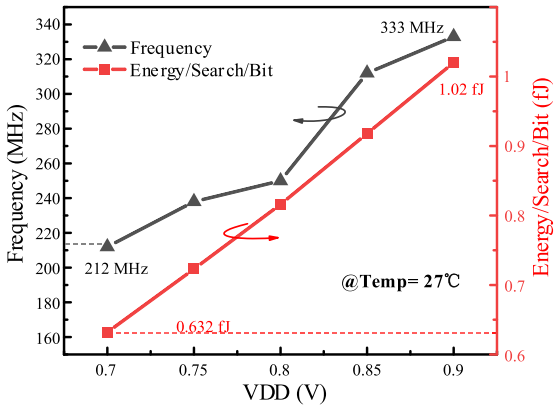


Fig. 17. Simulated frequency and energy consumption of the BCAM column-wise search.

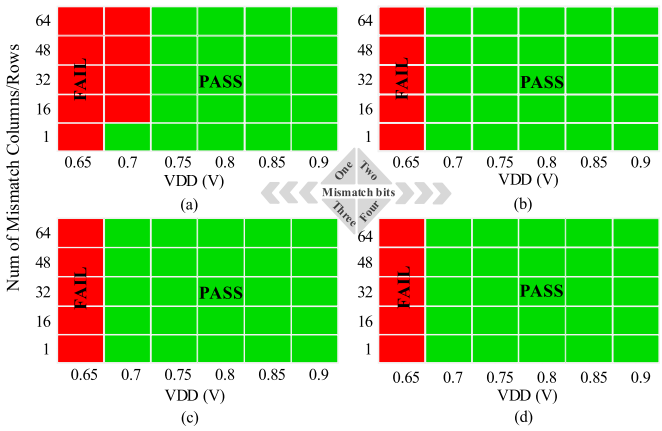


Fig. 18. Measured results of BCAM search for different supply voltage during (a) one, (b) two, (c) three, and (d) four-bit mismatch.

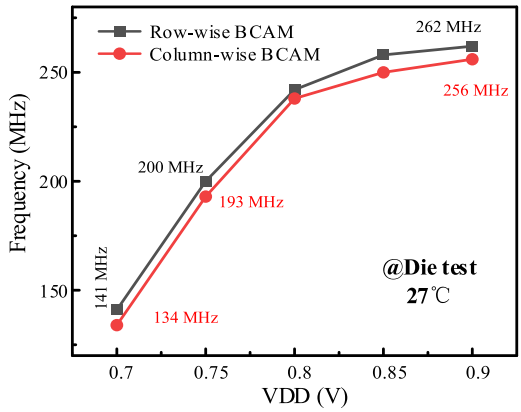


Fig. 19. Measured frequency of the BCAM row-and column-wise search.

The energy consumption of row OR operation is slightly higher than that of column AND operation over the operating voltage range.

Figs. 16 and 17 show the simulated frequency and energy of the BCAM row-and column-wise search operations, respectively, for different supply voltages. In fact, there is a possibility that there is only one bit mismatch in a row, and the discharge speed of the match line is relatively slow; thus, the pulsewidth of CWL/RWL should be sufficiently large to ensure the accuracy of matching results. In addition, when all the search results are matched, all match lines remain high, and the energy overhead of the search is at its lowest. When all search results are mismatched, the energy overhead of the search is the highest. A half-mismatching data pattern is reasonable to evaluate the energy consumption. Therefore, the frequency is obtained with a one-bit mismatch in the worst case, and the energy is simulated with a half-mismatching data pattern [4]. During one-bit row-wise (column-wise) mismatch testing, the search data are all 1 (bits 1 to 64). One bit of the stored data in each row (column) is set to 0, and the remaining bits are stored as 1. During half-mismatching testing, half of the search data are 0 (bits 1 to 32) and the other are 1 (bits 33 to 64). The stored data are 0 and 1 interleaved; that is, the data of each row (column) during row-wise (column-wise) search are as follows: bits 1 and 16 are 0; bits 17 to 32 are 1; bits 33 to 48 are 0; and bits 49 to 64 are 1. With a few exceptions, the frequency and energy consumption are approximately linear with the supply voltage. At a supply voltage of 0.9 V and temperature of 27 °C, the maximum frequencies of the BCAM row-and column-wise search are 333 MHz, and the energy consumption values are 1.025 and 1.02 fJ per search/bit, respectively. The minimum energy consumption is 0.635 fJ at 195 MHz and 0.7 V for row-wise search (Fig. 16) and 0.632 fJ at 212 MHz and 0.7 V for column-wise search (Fig. 17).

Fig. 18 shows the chip test results for the BCAM given by the number of mismatching columns or rows and according to the supply voltage. BCAM search errors are indicated by red marks, whereas correct search results are indicated by green marks. When a one-bit mismatch occurs between the search word and stored words, the voltage drop of the match line reduces as the supply voltage decreases, and match errors occur. When a multibit mismatch occurs between the search word and stored words or the supply voltage increases, the probability of search failure decreases.

Fig. 19 shows the measured operational frequency of the BCAM search at room temperature of 27 °C. The frequency

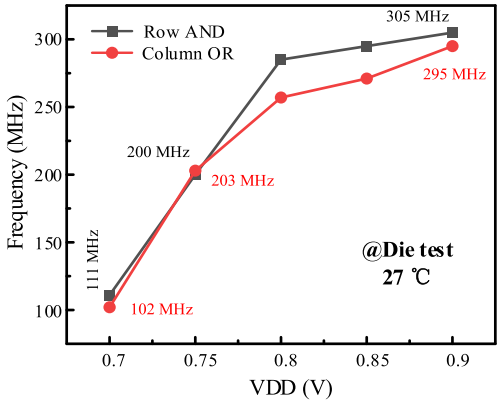
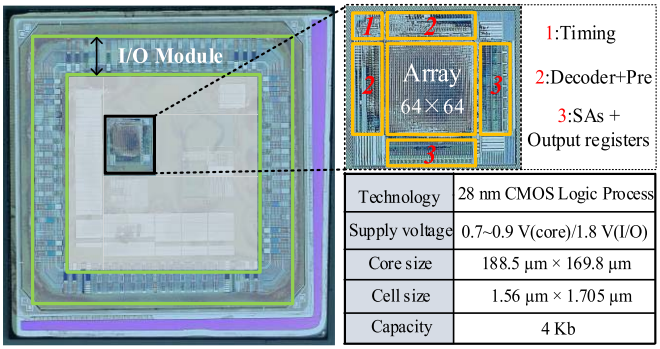


Fig. 20. Measured frequency of logic row AND and column OR operations.



Technology	28 nm CMOS Logic Process
Supply voltage	0.7~0.9 V(core)/1.8 V(I/O)
Core size	188.5 μm × 169.8 μm
Cell size	1.56 μm × 1.705 μm
Capacity	4 Kb

Fig. 21. Die photograph of fabricated SRAM.

curves of the BCAM row-and column-wise search operations are similar, reaching maximum frequencies of 262 and 256 MHz, respectively, at a supply voltage of 0.9 V, and frequencies of 141 and 134 MHz, respectively, at a supply voltage of 0.7 V. Fig. 20 shows the measured operating frequency of the logic in-memory operations at 27 °C. The operating frequency decreases as the supply voltage decreases. The maximum frequencies of the AND and OR operations are 305 and 295 MHz, respectively, at a supply voltage of 0.9 V. Parasitic parameters result in the frequency measurements of both the BCAM operation and the logic operations being lower than the simulation results. This is as expected, however, the difference is reasonable. Fig. 21 shows the die photograph of the fabricated SRAM memory.

V. Discussion

A. Area Cost and Structure Characteristics

The memory array plays an important role in the area of SRAM; thus, some in-memory computing SRAM architectures [1], [4], [21] use 6T cell to compress the area overhead. However, the 6T SRAM cell [21] degrades the read noise margins and exhibits poor performance at low voltages. The proposed SRAM has made some sacrifices in the cell area in pursuit of the reliability. The dedicated read ports provide reliable read operation with no interfering current to the cell storage node, making the read SNM equivalent to the hold SNM.

TABLE II
COMPARISONS WITH OTHER WORKS

		This work		[4]	[21]	[5]	[25]	[26]
Technology		28 nm		55 nm DDC	28 nm FDSOI	65 nm	65 nm	28 nm
Cell type		10T		4+2T	6T	8T	10T	(6+3× 2)T
Array size		64 × 64		128 × 128	64 × 64	128 × 128	128 × 128	256 × 128
Supply voltage (V)		0.9		0.8	1	1.2	1.2	1
BCAM	Freq. (MHz)	262(0.9 V)	256(0.9 V)	270(0.8 V)	370(1 V)	813(1.2 V)	500(1.2 V)	–
	Energy (fJ/bit)	1.025(0.9 V)	1.02(0.9 V)	0.45 (0.8 V)	0.6 (1 V)	0.85 (1.2 V)	0.77 (1.2 V)	–
		0.635(0.7 V)	0.632(0.7 V)					
Logic	Freq. (MHz)	~300(0.9 V)		230(0.8 V)	–	793(1.2 V)	–	2200(1 V)
	Energy (fJ/bit)	~15(0.9 V)		24.1(0.8 V)	–	~31(1.2 V)	–	23.8(1 V)
		~12.5(0.8 V)				~22.5(1 V)		
		~10.5(0.7 V)				16.6(0.8 V)		
Search mode		①	②	①	②	②	①	–
Function		SRAM/CAM/Logic/ Matrix transpose		SRAM/ CAM/Logic	SRAM/ CAM/Logic	SRAM/ CAM/Logic	BCAM	Logic/ Add

① Row-wise search
② Column-wise search
DDC, deeply depleted channel; FDSOI, fully depleted silicon on insulator

Table II shows the comparison between the proposed SRAM and some previous works. In [4], the logic and the CAM operations were limited to one-direction, and there may be snaking currents during operations that need to be overcome. In [21], double word lines were adopted to realize the CAM search operation, however, it can only complete column-wise search with the searched words only stored in columns. Similarly, the 8T cell in [5] also has the above features, and the writing of the searched data required additional costs. In this work, the SRAM cell has a better noise margin compared with the cells in [4] and [21] and also provides the SRAM unique bidirectional in-memory computing characteristics. In terms of logic operation, the proposed design exhibits lower energy consumption compared with previous works. The total logic energy consumption consists of the energy consumption of decoding, word lines drive, signal lines pre-charging and discharging, and sensitive amplification. The energy consumption per bit (fJ/bit) refers to the total energy consumption of the logic operation divided by the word length.

For matrix transposition, using a 4 × 4 matrix as an example and assuming that the matrix elements are 0 or 1 and are stored in 4 × 4 SRAM array. Then, a complete transposition operation using the conventional method, includes three stages, namely, the reading of 4 rows data (four read cycles), the processing of data by the arithmetic unit, and the writing of the generated matrix (four write cycles). The transposed matrix is obtained after the row and column elements of the original matrix are exchanged. According to the column read operation of the proposed SRAM, the 4 × 4 matrix transposition is easily realized through the reading of the data in four columns data (four read cycles), and the transposed matrix does not need to be stored separately. Therefore, the matrix transposition of the proposed SRAM omits data processing and writing after transposition, thereby reducing energy consumption and improving efficiency.

Both the typical 10T BCAM cell and the proposed cell comprise a 6T SRAM cell and a matching circuit comprising four NMOS transistors. The former contains two horizontal

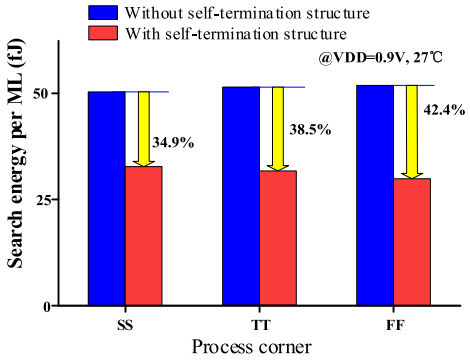


Fig. 22. Simulation results of energy consumption of BCAM search with/without self-termination structure across different corners (TT, SS, FF) at VDD = 0.9 V at 27 °C.

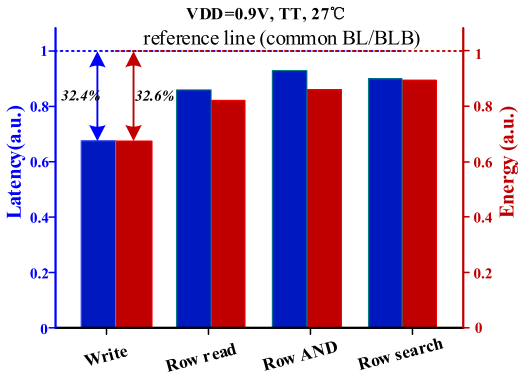


Fig. 23. Postlayout simulation comparison results of the latency and energy between the common shared BL/BLB and the separated BL/BLB at 0.9 V at 27 °C.

lines WL and ML, and four vertical lines BL/BLB and SL/SLB. The latter has four horizontal lines WWL, RWL, RL, and RR, and three vertical lines BL/BLB and CWL. In terms of layout design and wiring complexity, both of them are very similar. Although we are currently unable to obtain the specific layout information of commercial cells, their area overhead can be reasonably predicted to be very close.

In this work, the two-direction operation in the schematic design has good symmetry; however, in the back-end design, there are non-ideal parameters, such as different metal layer wiring, resulting in different parasitic capacitance and resistance. In addition, to meet design rules of 28 nm CMOS technology, the layout design of the horizontal and vertical SAs are different. There are also non-ideal design factors in the power supply. All the above factors may cause slight differences between row-and column-operation speeds.

B. Self-termination Structure

A self-termination structure is proposed in this article. Fig. 22 shows the simulated energy comparisons between the conventional BCAM search without self-termination structure and the proposed BCAM search operation using self-termination structure. The results at the same VDD = 0.9 V and temperature of 27 °C under different process corners (i.e., TT, SS, and FF) are illustrated in the figure. At the SS corner, the energy consumption of the match line (i.e., ML of Fig. 22) with a self-termination structure is 34.9% lower than that of the conventional match line. In addition, at the corners of TT and FF, the energy consumption improvement of the match line are 38.5% and 42.4%, respectively. The self-termination structure is beneficial in reducing the energy consumption of the CAM search.

C. Bitline Design

Two possible SRAM cell BL/BLB designs exist. The first one is that the BL/BLB are used in both writing and logic/search operations. The second one is that two sets of bit lines are used for different operations. The write-BL/BLB are dedicated to writing, and the compute-BL/BLB are used for logic/search operations. Fig. 23 shows the post-layout simulation results of the energy and latency of these two designs. The latency and energy for the various operations in the first design are normalized to 1, indicated by the dashed lines. The second design with separated BL/BLB is better than the first design with common shared BL/BLB in terms of speed and energy consumption. However, the cell with separated BL/BLB is more complicated in layout design, particularly placement and routing, and its cell area is 1.1 times than that of the proposed cell area with common shared BL/BLB.

D. Search Robustness

The match lines VDD level may be affected by the charge share of each capacitance of intermediate node between M1 (M3) and M2 (M4). In the BCAM row-wise search, when Q is 0, RL is not affected because M2 is turned OFF, and the voltage of an intermediate node between M3 and M4 is pulled up through M4 during the pre-charge phase. When Q is 1, the node voltage between M1 and M2 is pulled up through M2 during the pre-charge phase, and the node voltage between M3 and M4 is unknown. Similarly, in the BCAM column-wise search, only when Q is 0 can the match line BL be affected by the intermediate node capacitance between M1 and M2. To solve this problem and improve the robustness of the structure, the strategy is to activate RWL/CWL in advance during the pre-charge phase to refresh the voltage of the intermediate node between M1 (M3) and M2 (M4). The reliability of this strategy has been functionally verified.

Owing to the reduction of the device feature size and the advancements in manufacturing technologies, metal interconnections are increasingly becoming thinner, thus, electromigration (EM) [27] has become a serious problem affecting chip life, especially below 28 nm. To mitigate the effect of EM, semiconductor foundries have imposed restrictions on interconnections in design rules, such as length, width, and minimum spacing. To reduce EM in the layout design, we widened the width of power, clock, and individual signal

wires (RR/*RL*, BL/BLB) and used higher metal layers to increase the current-carrying capacity.

VI. Conclusion

A 10T SRAM cell with two decoupled read ports is proposed. The cell effectively isolates the connection between the storage node and the peripheral circuit and ensures data stability. The read static noise margin of the cell is 319.1 mV at VDD of 0.9 V and temperature of 27 °C, representing a threefold improvement over the conventional 6T SRAM cell. The search SNM is approximately 300 mV over the operating temperature (from 27 to 100 °C). The proposed search auxiliary circuit, a self-termination structure, can dynamically detect the match line voltage and relatively adjust the word lines RWL/CWL pulsewidth to reduce energy consumption (approximately 38.5% at a typical corner of TT). The minimum size of the transistors of the read ports is achieved to reduce the area overhead of the proposed structure. The symmetry of the proposed cell and the novel horizontal and vertical double word lines, can activate multiple rows or columns in parallel, and enable two-direction in-memory computing without additional data moving. The proposed SRAM cell can perform the CAM search, memory matrix transposition, and AND and OR operations. The 64 × 64 BCAM two-direction search achieves operation at approximately 260 MHz for 0.9 V supply while consuming around 1 fJ per search/bit. Logical operations between two 64-bit words can be performed at approximately 300 MHz for 0.9-V supply.

References

[1] J. Zhang, Z. Wang, and N. Verma, "A machine-learning classifier implemented in a standard 6T SRAM array," in *Proc. IEEE Symp. VLSI Circuits (VLSI-Circuits)*, Honolulu, HI, USA, Jun. 2016, pp. 1–2, doi: 10.1109/VLSIC.2016.7573556.

[2] A. Biswas and A. P. Chandrakasan, "CONV-SRAM: An energy-efficient SRAM with in-memory dot-product computation for low-power convolutional neural networks," *IEEE J. Solid-State Circuits*, vol. 54, no. 1, pp. 217–230, Jan. 2019, doi: 10.1109/JSSC.2018.2880918.

[3] M. Kang, S. K. Gonugondla, A. Patil, and N. R. Shanbhag, "A multi-functional in-memory inference processor using a standard 6T SRAM array," *IEEE J. Solid-State Circuits*, vol. 53, no. 2, pp. 642–655, Feb. 2018.

[4] Q. Dong *et al.*, "A 4 + 2T SRAM for searching and in-memory computing with 0.3-V V_{DDmin}," *IEEE J. Solid-State Circuits*, vol. 53, no. 4, pp. 1006–1015, Apr. 2018, doi: 10.1109/JSSC.2017.2776309.

[5] Z. Lin *et al.*, "In-memory computing with double word lines and three read ports for four operands," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 28, no. 5, pp. 1316–1320, May 2020, doi: 10.1109/TVLSI.2020.2976099.

[6] A. Agrawal, A. Jaiswal, C. Lee, and K. Roy, "X-SRAM: Enabling in-memory Boolean computations in CMOS static random access memories," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 12, pp. 4219–4232, Dec. 2018.

[7] Y. Zhang, L. Xu, Q. Dong, J. Wang, D. Blaauw, and D. Sylvester, "Recryptor: A reconfigurable cryptographic cortex-M0 processor with in-memory and near-memory computing for IoT security," *IEEE J. Solid-State Circuits*, vol. 53, no. 4, pp. 995–1005, Apr. 2018, doi: 10.1109/JSSC.2017.2776302.

[8] S. R. Srinivasa *et al.*, "ROBIN: Monolithic-3D SRAM for enhanced robustness with in-memory computation support," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 7, pp. 2533–2545, Jul. 2019, doi: 10.1109/TCSI.2019.2897497.

[9] F.-K. Hsueh *et al.*, "TSV-free FinFET-based monolithic $3D^{+}$-IC with computing-in-memory SRAM cell for intelligent IoT devices," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2017, pp. 12.6.1–12.6.4, doi: 10.1109/IEDM.2017.8268380.

[10] Z.-R. Wang *et al.*, "Efficient implementation of Boolean and full-adder functions with 1T1R RRAMs for beyond von Neumann in-memory computing," *IEEE Trans. Electron Devices*, vol. 65, no. 10, pp. 4659–4666, Oct. 2018, doi: 10.1109/TED.2018.2866048.

[11] M. Moreau *et al.*, "Reliable ReRAM-based logic operations for computing in memory," in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr. (VLSI-SoC)*, Verona, Italy, Oct. 2018, pp. 192–195, doi: 10.1109/VLSI-SoC.2018.8644780.

[12] Z. Yang and L. Wei, "Logic circuit and memory design for in-memory computing applications using bipolar RRAMs," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Sapporo, Japan, May 2019, pp. 1–5, doi: 10.1109/ISCAS.2019.8702555.

[13] F. Parveen, Z. He, S. Angizi, and D. Fan, "HielM: Highly flexible in-memory computing using STT MRAM," in *Proc. 23rd Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jeju, South Korea, Jan. 2018, pp. 361–366, doi: 10.1109/ASPDAC.2018.8297350.

[14] W. Kang, H. Wang, Z. Wang, Y. Zhang, and W. Zhao, "In-memory processing paradigm for bitwise logic operations in STT–MRAM," *IEEE Trans. Magn.*, vol. 53, no. 11, pp. 1–4, Nov. 2017, doi: 10.1109/TMAG.2017.2703863.

[15] S. Jain, A. Ranjan, K. Roy, and A. Raghunathan, "Computing in memory with spin-transfer torque magnetic RAM," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 3, pp. 470–483, Mar. 2018, doi: 10.1109/TVLSI.2017.2776954.

[16] Y. Zhang *et al.*, "Time-domain computing in memory using spintronics for energy-efficient convolutional neural network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 3, pp. 1193–1205, Mar. 2021, doi: 10.1109/TCSI.2021.3055830.

[17] S. K. Maurya and L. T. Clark, "A dynamic longest prefix matching content addressable memory for IP routing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 6, pp. 963–972, Jun. 2011, doi: 10.1109/TVLSI.2010.2042826.

[18] T.-L. Tsai, J.-F. Li, C.-L. Hsu, and C.-T. Sun, "Testing of in-memory-computing 8T SRAMs," in *Proc. IEEE Int. Symp. Defect Fault Tolerance VLSI Nanotechnol. Syst. (DFT)*, Noordwijk, The Netherlands, Oct. 2019, pp. 1–4, doi: 10.1109/DFT.2019.8875487.

[19] X. Wang, Y. Yang, and M. Shang, "A novel content addressable memory based on hybrid memristor-CMOS architecture," in *Proc. 37th Chin. Control Conf. (CCC)*, Wuhan, China, Jul. 2018, pp. 8502–8506, doi: 10.23919/ChiCC.2018.8482628.

[20] S. Srinivasa, W.-H. Chen, Y.-N. Tu, M.-F. Chang, J. Sampson, and V. Narayanan, "Monolithic-3D integration augmented design techniques for computing in SRAMs," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Sapporo, Japan, May 2019, pp. 1–5, doi: 10.1109/ISCAS.2019.8702536.

[21] S. Jeloka, N. B. Akesh, D. Sylvester, and D. Blaauw, "A 28 nm configurable memory (TCAM/BCAM/SRAM) using push-rule 6T bit cell enabling logic-in-memory," *IEEE J. Solid-State Circuits*, vol. 51, no. 4, pp. 1009–1021, Apr. 2016, doi: 10.1109/JSSC.2016.2515510.

[22] H.-C. Chen, J.-F. Li, C.-L. Hsu, and C.-T. Sun, "Configurable 8T SRAM for enbling in-memory computing," in *Proc. 2nd Int. Conf. Commun. Eng. Technol. (ICCET)*, Nagoya, Japan, Apr. 2019, pp. 139–142, doi: 10.1109/ICCET.2019.8726871.

[23] V. R. Datti and P. V. Sridevi, "Performance evaluation of content addressable memories," in *Proc. 7th Int. Conf. Rel., Infocom Technol. Optim. (Trends Future Directions) (ICRITO)*, Noida, India, Aug. 2018, pp. 596–598, doi: 10.1109/ICRITO.2018.8748808.

[24] B. Wang, T. Q. Nguyen, A. T. Do, J. Zhou, M. Je, and T. T.-H. Kim, "Design of an ultra-low voltage 9T SRAM with equalized bitline leakage and CAM-assisted energy efficiency improvement," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 2, pp. 441–448, Feb. 2015, doi: 10.1109/TCSI.2014.2360760.

[25] A. T. Do, C. Yin, K. S. Yeo, and T. T.-H. Kim, "Design of a power-efficient CAM using automated background checking scheme for small match line swing," in *Proc. ESSCIRC (ESSCIRC)*, Bucharest, Romania, Sep. 2013, pp. 209–212, doi: 10.1109/ESSCIRC.2013.6649109.

[26] W. Simon, J. Galicia, A. Levisse, M. Zapater, and D. Atienza, "A fast, reliable and wide-voltage-range in-memory computing architecture," in *Proc. 56th Annu. Design Autom. Conf.*, Las Vegas, NV, USA, Jun. 2019, pp. 1–6.

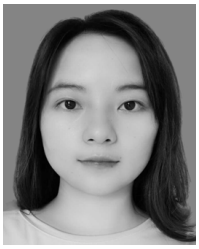
[27] S. Majji, T. R. Patnala, M. Valleti, C. S. Pasumarthi, S. Kothapalli, and S. R. Karanam, "A study on the comprehensive analysis of electro migration for the nano technology trends," in *Proc. 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Coimbatore, India, Mar. 2020, pp. 898–901, doi: 10.1109/ICACCS48705.2020.9074328.

Zhiting Lin (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electronics and information engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2004 and 2009, respectively.

From 2015 to 2016, he was a Visiting Scholar with the Engineering and Computer Science Department, Baylor University, Waco, TX, USA. In 2011, he joined the Department of Electronics and Information Engineering, Anhui University, Hefei, Anhui, where he is currently a Professor. He has authored or coauthored about 50 articles. He holds over 20 Chinese patents. His research interests include pipeline analog-to-digital converters and high-performance static random access memory.

Zhiyong Zhu received the B.S. degree in electric information engineering from Anhui University, Anhui, China, in 2018, where he is currently pursuing the M.S degree in circuits and systems.

His current research interest includes static random access memory.



Honglan Zhan received the B.S. degree in electrical engineering and automation from Hefei normal University, Hefei, China, in 2018. She is currently pursuing the M.S degree in integrated circuit engineering with Anhui University, Anhui, China.

Her current research includes high-performance static random access memory and in-memory computing.

Chunyu Peng received the B.S. degree in communications engineering and the M.S. degree in circuits and systems from Anhui University, Anhui, China, in 2010 and 2013, respectively, where he is currently pursuing the Ph.D. degree in microelectronics and solid-state electronics.

He is currently an Assistant Lecturer in microelectronics and solid-state electronics with Anhui University. His research interests include signal processing, analog IC design, and high-performance memory technology.

Xiulong Wu received the B.S. degree in computer science from the University of Science and Technology of China (USTC), Hefei, China, in 2001, and the M.S. and Ph.D. degrees in electronic engineering from Anhui University, Hefei, in 2005 and 2008, respectively.

From 2013 to 2014, he was a Visiting Scholar with the Engineering Department, The University of Texas at Dallas, Richardson, TX, USA. He is currently a Professor with Anhui University. He has authored or coauthored about 60 articles. He holds over ten Chinese patents. His research interests include high-performance static random access memory and mixed-signal ICs.

Yuan Yao received the B.S. degree in integrated circuit design and integration system from the Qingdao University of Science and Technology, Qingdao, China, in 2018. He is currently pursuing the M.S degree in integrated circuit engineering with Anhui University.

His research interests includes static random access memory and in-memory computing circuits.

Jianchao Niu received the B.S. degree in electrical information engineering from Hefei University, Hefei, China, in 2018. He is currently pursuing the M.S degree in integrated circuit engineering with Anhui University, Anhui, China.

His current research includes high-performance static random access memory and in-memory computing.

Junning Chen received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 1993.

From September 1993 to July 1996, he held a Post-Doctoral Position with the CAD Laboratory, Fudan University, Shanghai, China. He was the Principal Investigator with the National Science and Technology Major Projects. He became a Professor with Anhui University, Hefei, China, in 1996. His research interests include VLSI design and the physics and processes related to semiconductor devices.