ENERGY EFFICIENT IN-MEMORY COMPUTING ARCHITECTURE WITH ENHANCED POLYMORPHIC CIRCUIT

A PROJECT REPORT

Submitted by

PAVITHRA B **513120106063**

SHRUTHI B K **513120106078**

VINITHA K **513120106095**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

THANTHAI PERIYAR GOVERNMENT INSTITUTE OF TECHNOLOGY, VELLORE-02.

ANNA UNIVERSITY: CHENNAI 600 025

MAY 2024

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"ENERGY EFFICIENT IN-MEMORY COMPUTING ARCHITECTURE WITH ENHANCED POLYMORPHIC CIRCUIT"** is the bonafide work of **"SHRUTHI B K – 513120106078"** who carried out the project work under my supervision.

SIGNATURE
Dr. S. LETITIA, M.E., Ph.D.,
HEAD OF THE DEPARTMENT,
PROFESSOR,

Department of Electronics and Communication Engineering,
Thanthai Periyar Govt. Institute of Technology, Vellore-632002.

SIGNATURE
Mrs. M. JANANI, M.E.,
PROJECT CO-ORDINATOR& SUPERVISOR,
ASSISTANT PROFESSOR,
Department of Electronics and Communication Engineering,
Thanthai Periyar Govt. Institute of Technology, Vellore-632002.

Project work viva voce held on _____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of task would be incomplete without the mention of the people who made it possible, their constant guidance and encouragement crowned our effects with success.

Our profound thanks to our Principal **Dr. P. K. PALANI, B.E (HONS)., M.E., Ph.D.,** Thanthai Periyar Government Institute of Technology, Vellore, for his scholarly guidance and valuable motivation throughout the career.

We articulate our special honor to Professor **Dr. S. LETITIA, M.E.,Ph.D.,** Head of the Department, Electronics and Communication Engineering, Thanthai Periyar Government Institute of Technology, Vellore, for her valuable suggestions and moral support to complete the project in an efficacious way.

We extend our heartfelt thanks to our esteemed Faculty Advisor **Dr. R. DHANALAKSHMI, M.E., Ph.D.,** Associate Professor, Department of Electronics and Communication Engineering, for her invaluable guidance and unwavering support. Her expertise has been instrumental in shaping our project's success, and we are deeply grateful for her dedication.

We hereby express our deep sense of gratitude to our Guide / Project coordinator **Mrs. M. JANANI, M.E.,** Assistant Professor, Department of Electronics and Communication Engineering, for her expert guidance and encouragement throughout the project and her continuous support to complete the project in a successful manner.

Also, we express our sincere thanks to all the teaching and non-teaching staff members of Electronics and Communication Engineering department for their encouragement in making our project successful.

ABSTRACT

To accomplish In-Memory Computing (IMC) by implementing all fundamental two-input Boolean functions, a hybrid memory architecture based on a novel array of SRAM and Resistive Random-Access Memory (RRAM) cells is proposed. It is possible to set up the SRAM array as a dual-purpose component. The Proposed architecture can be utilized as an SRAM array to store data for high-performance applications. It can also be set up to read data from RRAMs and carry out In-memory computations as a sense amplifier (SA-SRAM).The Proposed 11T cell's leakage power consumption is also decreased by 67.05% and 35.15%, respectively, when compared with 6T and 10T SRAM cells. Furthermore, the energy consumption of our design in application areas like image processing is significantly less than the well-known comparative in-memory architecture solution.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BJT	-	Bipolar Junction Transistor
BL	-	Bit Line
BLB	-	Bit Line Bar
CMOS	-	Complementary Metal Oxide Semiconductor
DRAM	-	Dynamic Random Access Memory
DWL	-	Divided Word Line
EDA	-	Electronic Design Automation
FINFET	-	FIN Field Effect Transistor
GDI	-	Gate Diffusion Input
HRS	-	High Resistance State
HSNM	-	Hold Static Noise Margin
HV	-	High Voltage
I/O	-	Input Output
IC	-	Integrated Chip
IG FINFET	-	Independent Gate FINFET
IMC	-	In Memory Computing
INV	-	Inverter
LRS	-	Low Resistance State
LV	-	Low Voltage
MD	-	Mode Decoder
MMA	-	Main Memory Array
MMRD	-	MMA Row Decoders
MOSFET	-	Metal Oxide Semiconductor Field Effect Transistor
MRAM	-	Magnetic Random Access Memory
MTJ	-	Magnetic Tunnel Junction

MV	-	Medium Voltage
OCR	-	Optical Character Recognition
PDP	-	Power Delay Product
PMU	-	Polymorphic Unit
RBL	-	Read Bit Line
RCA	-	Ripple Carry Adder
READ-EN	-	Read Enable
ReBL	-	Resistive Bit Line
RED	-	Read Enable Decoder
RRAM	-	Resistive Random Access Memory
RSNM	-	Read Static Noise Margin
RWL	-	Read Word Line
SA SRAM	-	Sense Amplifier SRAM
SAU	-	Sense Amplifier Unit
SCD	-	Source Column Decoders
SLD	-	Source Line Decoders
SNM	-	Static Noise Margin
SoC	-	System on Chip
SPICE	-	Simulation Program with Integrated Circuit Emphasis
SRAM	-	Static Random Access Memory
SRD	-	Source Row Decoder
WBL	-	Write Bit Line
WL	-	Word Line
WSNM	-	Write Static Noise Margin
WWL	-	Write Word Line

CHAPTER 1
INTRODUCTION

With the increasing need for high-performance processing applications such as artificial intelligence, neural networks, search engines, biological systems, and image processing, the Von Neumann architecture faces a critical challenge known as the memory wall. In this case, transferring a large amount of data back and forth between a separate processor and memory leads to severe energy consumption, high latency, and I/O congestion.

Therefore, one of the most promising approaches for alleviating these challenges is In-Memory Computing (IMC), which provides computing capability inside memory. IMC can perform simple computational tasks to reduce memory-processor data transfers. In other words, memory cells are considered to accomplish normal read-write operations and perform simple logical computations within the memory to bypass the Von Neumann bottleneck.

Dynamic Random-Access Memory (DRAM) and Static Random-Access Memory (SRAM) are used for volatile in-memory computing design, and emerging technologies such as Magnetic RAM (MRAM) and Resistive Random Access Memory (RRAM) are used in the nonvolatile area. It is worth mentioning that the advantages of using volatile memories are related to widespread usage and the read/write operation speed. Moreover, these memories can implement short-term memory in Hybrid SRAM/RRAM IMC architecture based on a Reconfigurable SRAM Sense Amplifier. However, high power consumption in DRAMs and large area occupation in SRAMs may deter these designs.

In contrast, low power dissipation, dense integration, and nonvolatility are the significant privileges of emerging technologies for the IMC paradigm. Meanwhile, these memories face high read/write delay and reliability issues due to low endurance.

A novel hybrid SRAM/RRAM-based In-Memory Computing architecture capable of performing all Boolean logic functions is proposed. The SRAM array is designed by benefiting from the proposed reconfigurable 11T SRAM cell capable of reading data from RRAM-based main memory in sense-amplifier mode and performing all boolean logic functions besides the data storage ability in the SRAM mode. Moreover, the proposed SRAM cell has a high static noise margin beside the free half-select issue, which is critical in memory design.

1.1 FINFETS Vs MOSFETS

FinFETs, or Fin Field-Effect Transistors, represent a significant evolution from traditional MOSFETs. Unlike MOSFETs, which have a planar structure, FinFETs feature a three-dimensional fin-like architecture, allowing for better control of current flow.

This design results in improved performance, lower power consumption, and enhanced scalability, crucial for advanced semiconductor technology.

Additionally, FinFETs exhibit superior electrostatic control, enabling higher transistor density on a chip and thus enabling more powerful and efficient electronic devices.

While MOSFETs have dominated the semiconductor industry for decades, FinFETs have emerged as a promising alternative, particularly for cutting-edge applications in fields like artificial intelligence, high-performance computing, and mobile devices.

1.2 FIN FIELD EFFECT TRANSISTOR [FINFET]:

FinFET (Fin shaped FET) is a Field Effect Transistor (FET) device structure. A FinFET is fabricated in a silicon layer overlying an insulating layer with the device extending from insulating layer as a fin. In order to provide enhanced drive current and effectively suppressed short channel effects, double gates are provided over the sides of the channel in FinFET.

A FinFET technology provides the higher drive current for a given transistor footprint, hence higher speed, lower leakage current, hence lower power consumption.

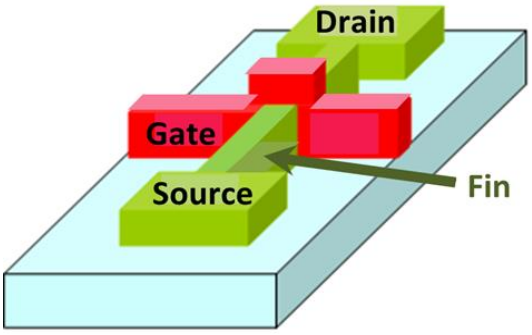


Figure 1.1 Double Gate FINFET

1.2.1 NEED FOR FINFET

- Since the fabrication of MOSFET, the channel length of the device has been shrinking constantly so as to fabricate compact and fast devices.
- The shorter section of the gate electrode is known as the length and the longer section is called the width.

- As the channel length of a MOSFET reduces, the short-channel effects increase. The short-channel effects are attributed to two physical phenomena:
  - The limitation imposed on electron drift characteristics in the channel.
  - The modification of the threshold voltage due to the channel length.

1.2.2 CHARACTERISTICS OF FINFETS

The three-dimensional structure of FinFET provides many advantages over MOSFET.

- A large number of transistors can be incorporated into a single chip. FinFET technology is suitable for IC fabrication, as it has higher scalability for the given footprint area than MOSFET'S.
- As chips are downsized, transistors also shrink. This compactness brings the drain and source closer and reduces the gate control over the channel carriers.
- This type of short-channel effect can cause serious issues in MOSFETs. The presence of fins gives FinFETs better short-channel behavior.

1.3 RESISTIVE RANDOM-ACCESS MEMORY (RRAM)

- RRAM is a two-terminal device comprising an oxide layer sandwiched between two metal layers. By applying a voltage across its terminals, RRAM’s resistance can change between the low-resistance state (LRS) and the high-resistance state (HRS).
- Despite RRAM being a memory element, its influence has extended beyond memory design to logic circuits and computing systems.

- Furthermore, the use of RRAM in next-generation nonvolatile memories has been touted due to its near-zero leakage power consumption, low read/write voltage, fast switching speed, and excellent scalability.

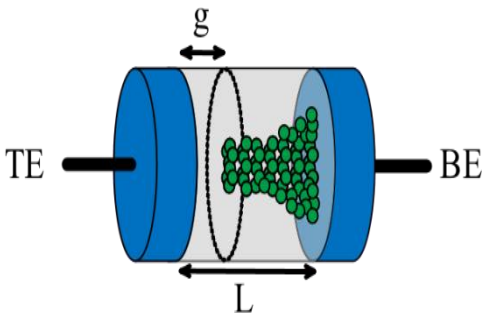


Figure 1.2 Physical structure of RRAM.

- Applying different voltages to the top (TE) and the bottom (BE) metal electrodes can form/rupture oxygen vacancy inside the oxide layer (OL), considered a conductive filament (CF).
- A positive voltage should be applied to RRAM to transmit from HRS to LRS (SET). Alternatively, transmitting from LRS to HRS (RESET) can be achieved using a negative voltage.
- It is worth mentioning that the RRAM device can be fabricated on top of the transistors.

CHAPTER 2

LITERATURE REVIEW

J. T. McDonald et al. (2016) discusses about that None of the designs has benefited from the fast read/write of the volatile DRAM (or SRAM) designs and the low power consumption and dense integration of nonvolatile emerging technologies (RRAM/MRAM). Using a single circuit to implement multiple functions reduces chip area and forms a primary barrier to security issues like reverse engineering or IC counterfeiting. These features are provided by polymorphic circuits.

A. Agrawal et al. (2018) proposed an 8TSRAM cell for in-memory computing capable of performing simple Boolean computations. Although the latency is reduced because of the separated read/write operations and in-memory Boolean logic calculations, write power consumption and reliability concerns in half-selected cells are significant issues in this design. Moreover, this design is not suitable for performing complex Boolean logic functions.

Z. Lin et al. (2020) implemented an 8T SRAM cell with three ports for four operands to perform complex Boolean functions. In one cycle, this design can perform three Boolean operations (XOR/XNOR, AND/NAND, and OR/NOR). However, the static noise margin (SNM) of the SRAM cells is considerably reduced at the nanoscale. Moreover, the area overhead and volatility are the critical bottlenecks of this design.

S. Angizi et al. (2020) Proposed nonvolatile hybrid MTJ/CMOS logic gates capable of performing Boolean logic functions for the in-memory computation . Although this design shows acceptable performance compared to benchmarks, the delay in performing Boolean computation is considerably higher than in CMOS-based designs. Moreover, this design is not suitable for applications that need real-time computations.

S. Sheikhfaal and R. F. Demara (2020) proposed a spintronic/CMOS memory, including volatile DRAM and nonvolatile MRAM cells, to cope with the requirements of real and none real-time applications. However, this design is unsuitable for complex calculations. Moreover, this design's writing energy and delay are considerably higher than the CMOS designs.

CHAPTER 3

EXISTING MODELS

3.1 GENERAL

This chapter discusses about the existing models (6T SRAM and 10T SRAM cells) with its limitations.

3.2 6T STANDARD SRAM CELL:

The SRAM cell typically consists of a cross-coupled pair of inverters forming a latch as shown in Figure 3.1, controlled by access transistors for read and write operations. This structure ensures stability and enables the storage of a single bit of data.

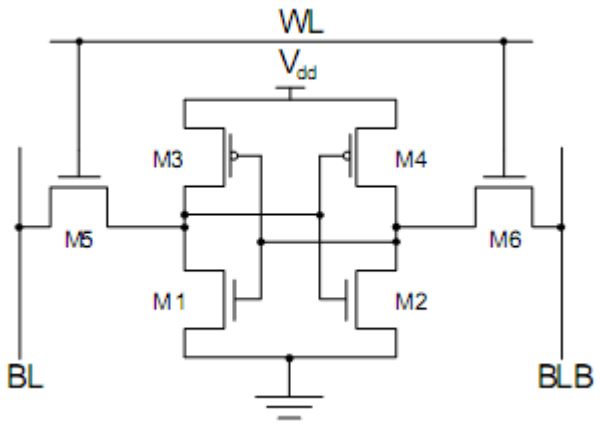


Figure 3.1 6T Standard SRAM

3.2.1 LIMITATIONS

- Its high static power consumption makes it less efficient for battery-powered devices.
- Read and write disturbances can also compromise data integrity, requiring additional circuitry for mitigation.

3.3 10T SRAM CELL

The 10T SRAM cell features additional access transistors compared to conventional 6T cells as shown in Figure 3.2, enabling separate read and write paths. This design enhances stability and reduces the risk of data corruption during read and write operations.

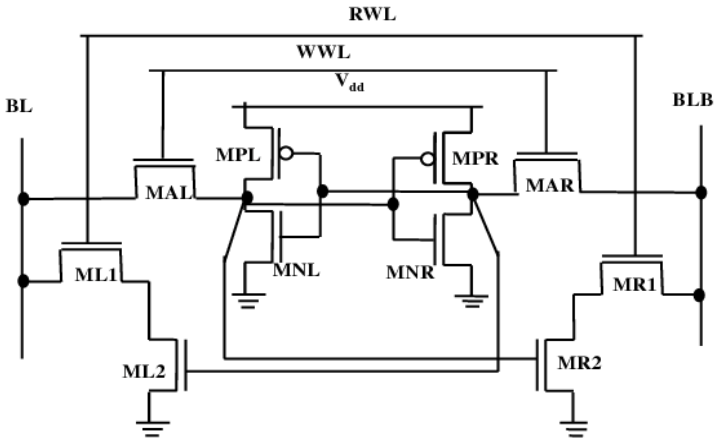


Figure 3.2 10T SRAM Cell.

3.3.1 LIMITATIONS

- 10T SRAM, an extension of the conventional 6T SRAM cell, introduces additional transistors to mitigate read and write disturbances and improve stability.
- The increased number of transistors in the 10T SRAM cell results in larger cell size and higher fabrication complexity compared to the 6T variant, limiting its scalability and integration density.
- Consequently, while 10T SRAM addresses some of the drawbacks of 6T SRAM, its practical implementation may be constrained in applications where space, power efficiency, and cost are critical considerations.

CHAPTER 4

PROPOSED METHOD

4.1 HALF SELECT 11T SRAM CELL

One of the most important considerations that must be considered in the design of an SRAM cell is to avoid unwanted writing in the neighboring cells of the cell of interest in the writing or reading process, known as the half-select issue.

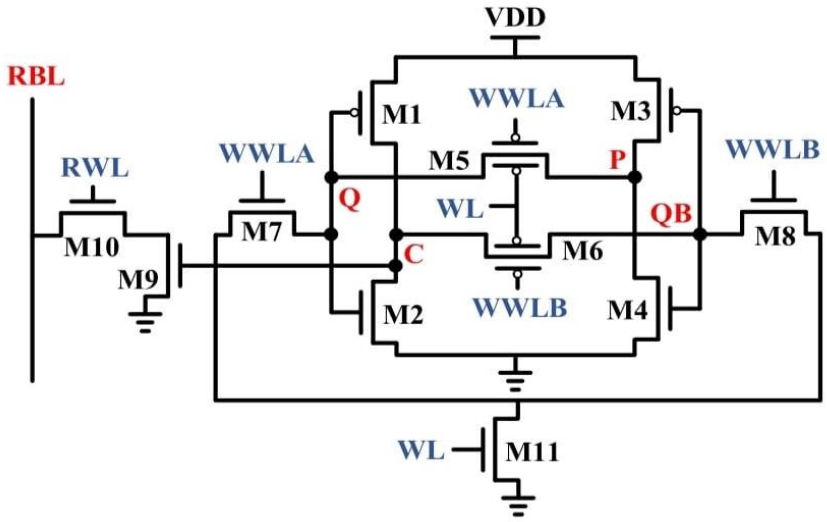


Figure 4.1 Proposed Half Select 11T SRAM

- 11TSRAM cell provides greater read power and write power compared to 6TSRAM cell.
- To write ‘0’ in an arbitrary cell, that cell’s WL and WWLA signals should be activated, as shown in Figure 4.1.
- In any column cell with only the WWLA signal activated, the M5 transistor is still ON, and the feedback paths are active.
- Therefore, the floating problem existing in the cells is solved, and the competition causing unnecessary power dissipation in cells is eliminated

4.2 RECONFIGURABLE HALF SELECT 11T(HF11T) CELL

To obtain the in-memory computing requirements in application areas such as optical character recognition (OCR), the proposed SRAM cell is redesigned to be a sense amplifier for reading data from a nonvolatile memory or a volatile memory element (SRAM mode).

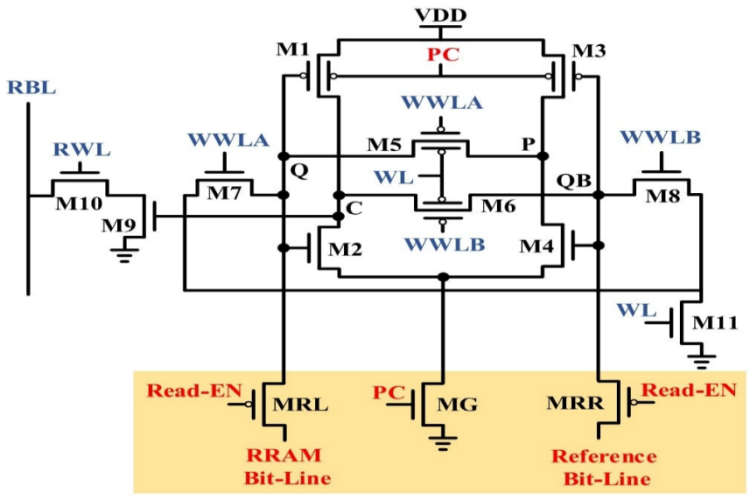


Figure 4.2 Reconfigured Half Select 11T SRAM

- Accordingly, as shown in Figure 4.2, the back gates of the pull-up transistors (M1 and M3) are connected to the PC control signal. An extra n-type transistor (MG) is placed between M2/M4 and ground to float the storage nodes and pre-charge them to prepare the cell for sensing the storage data in nonvolatile memory.
- The proposed cell can be configured as either a fast memory element (SRAM mode) to perform a normal read operation in real-time applications by disabling the Precharge node and Read-EN control signals or a sense amplifier (sense amplifier mode) for in-memory computing by enabling PC followed by Read-EN.

CHAPTER 5

IN-MEMORY ARCHITECTURE

A new hybrid memory architecture, including an RRAM-based main memory array (MMA), a reconfigurable SRAM array (SA-SRAM), and polymorphic logic units (PMU) is Proposed.

The MMA row (MMRD) and source line (SLD) decoders are considered to select cells for either reading, writing or computing operations. The SA-SRAM unit can be configured in SRAM mode and used as a sense amplifier to read the MMA data and then switch back to SRAM mode to maintain the reading data.

It is worth mentioning that the SA-SRAM cells can perform in-memory computing besides the storage mode. Switching between the SRAM and sense amplifier modes is done by the Mode Decoder (MD) unit, which provides the PC signals.

In the meantime, the Read-Enable Decoder (RED) contributes to the communication between the MMA and SA-SRAM units. In the SRAM mode, the Source Column Decoders (SCD) and Source Row Decoders (SRD) decoders are utilized for reading from and writing to the SA-SRAM cells.

The SA-SRAM output can be maintained in the latch unit presented in the Sense Amplifier Unit (SAU) to provide proper inputs for PMUs. PMUs act as security-processing blocks to increase the computation capability in this architecture, designed to handle real-time and non-real-time applications.

5.1 MMA MEMORY MODE

In MMA, the data stored in a memory cell is represented by the resistance of RRAM. To write a bit into a memory cell, the corresponding Source Line (SL) and bit-line (BL) are connected to the respective voltages to modify the stored data.

For example, to write ‘0’ into a cell, the corresponding SL should be grounded, and BL should be connected to the write voltage (2V). Then, the RRAM cell will be reset by enabling the corresponding DWL signal.

On the other hand, for writing ‘1’ into a cell, BL should be grounded, and SL should be connected to the write voltage (2V). Then, the RRAM cell will be set by asserting the DWL signal.

The read process is accomplished by connecting SL to the ground and pre-charging BL and the reference bit line (ReBL) to 0.7V.ReBL is connected to the reference resistance through a transistor connected to the L1 signal.

For instance, when the word line of the desired cell and L1are asserted, Bl and ReBL are discharged at different rates depending on the resistance of the corresponding cell. Since an SA-SRAM is configured in the sense amplifier mode, it acts as a comparator.

To this end, the discharge voltage in BL is compared with the discharge voltage in the ReBL for a certain period, and the desired data is detected and stored.

5.2 MMA COMPUTING MODE

To enable IMC in MMA, SLs are connected to the ground, the BLs are pre-charged to 0.7V, and the word lines of the two desired rows are activated to start the computing process.

With the activation of the two desired rows, two RRAM cells become parallel in each column, and according to the data of each cell, the discharge voltage of BL can be stable at three voltage levels.

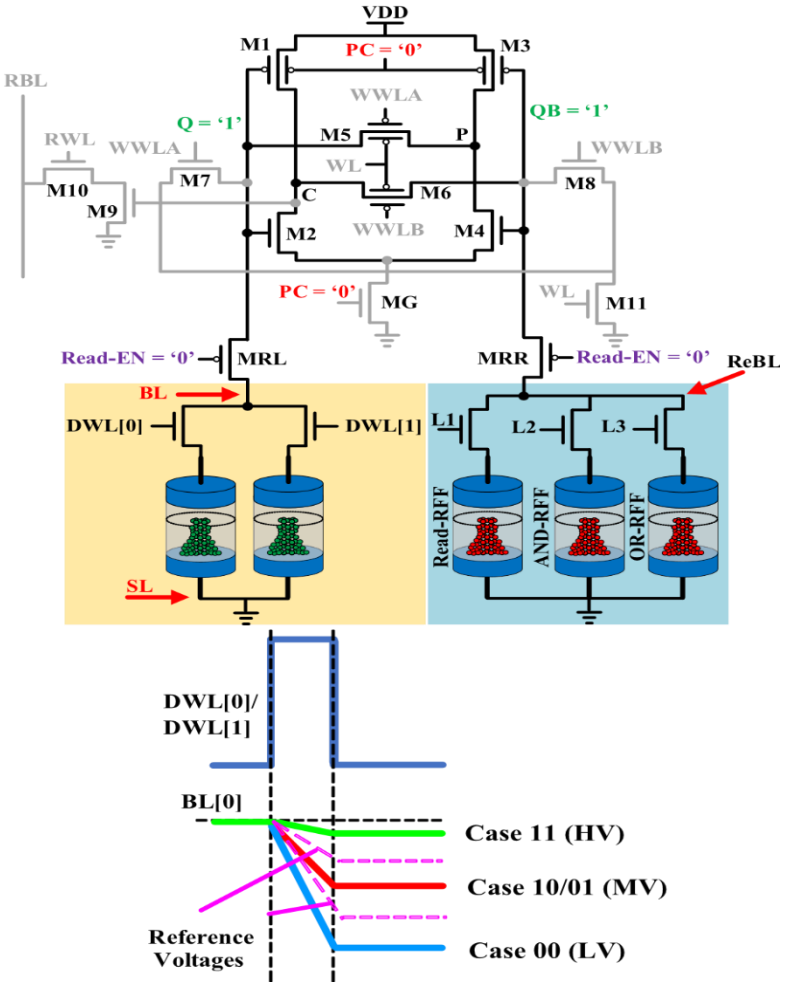


Figure 5.1 In-Memory Computing structure in MMA

As shown in Figure 5.1, the data on both cells are ‘‘00’’ (low-low), ‘‘01’’ or ‘‘10’’ (low-high), or ‘‘11’’ (high- high), the voltage on BL will be low (LV), medium (MV), or high (HV), respectively.

5.2.1 AND/NAND LOGIC

To obtain the AND/NAND logic, the reference voltage must be placed between the MV and HV resistance regions. At the same time, as the word lines of the desired cells are activated, and BL and ReBL are pre-charged, the L2 signal is activated (reference AND/NAND resistance), and SA-SRAM detects the desired data by comparing the voltages on BL and ReBL. Finally, the appropriate output result is stored in the memory.

5.2.2 OR/NOR LOGIC

The OR/NOR logic, like AND/NAND, is obtained by setting the appropriate reference resistance that places the ReBL voltage between LV and MV during IMC. The reference resistance in this operation is connected by activating the L3 signal. In this operation, the desired output is generated based on the different voltages on BL and ReBL according to the resistance of their paths.

5.3 SA-SRAM MEMORY MODE

SA-SARM operates in the SRAM mode by disabling the active-low PC signal (PC= ‘1’). In summary, the write operation is accomplished by selecting the required signals (WWLA, WWLB, and WL) for the desired cell, activated by the SRD and SCD decoders. Based on the data stored in the cell, RBL discharges or remains at its pre-charged value in the read operation.

5.4 SA-SRAM COMPUTING MODE

The SA-SRAM cell provides in-memory computations in the SRAM mode, benefiting from the isolated read path due to the decoupled read ports. The isolated read mechanism makes it possible to perform the OR/NOR AND/NAND, and XOR/XNOR functions within the SASRAM array. In the following, how each of these operations is performed is described.

5.4.1 AND/NAND OPERATION

In the AND/NAND operation, the output is '1'/'0' only if both inputs are '1'. By activating two RWLs of the selected cells connected to RBL, RBL remains at its high value only if the Q(QB) nodes of the two activated cells in the same column contain '1' ('0') values.

Otherwise, the RBL is discharged to the ground, indicating the '0' output. As shown in Figure 5.2, a high-skewed inverter (INV1) acts as a sense amplifier for each column gated by the corresponding RBL for fast detection of the output at a certain period. Placing the high-skewed inverter causes the NAND operation to be executed, so a subsequent unskewed inverter (INV2) is needed to perform the AND operation.

5.4.2 OR/NOR OPERATION

As mentioned in the AND/NAND operation, when the RWLs of the cells of interest are activated simultaneously, based on the data stored in these cells, the selected RBL can be expected to discharge to the ground or not.

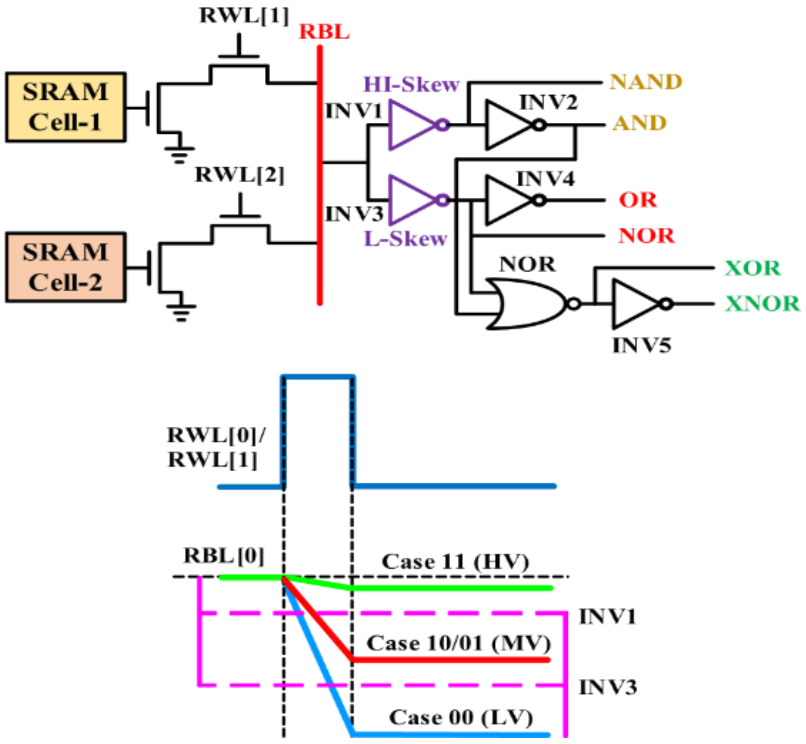


Figure 5.2 SRAM-mode IMC structure

If two desired cells contain ‘‘00’’, ‘‘01’’, or ‘‘10’’, RBL starts to discharge. However, the discharge rate in the ‘‘00’’ state is much faster than in the ‘‘01’’/‘‘10’’ states due to different discharge path resistances.

As shown in Figure 5.2, placing a low-skewed inverter (INV3) with a switching threshold between the LV and MV regions implements the NOR function. In addition, a cascaded inverter (INV4) is needed to realize the OR function.

5.4.3 XOR/XNOR OPERATION

The XOR operation can be performed straightforwardly by NORing the AND (INV2) and NOR (INV3) outputs.

5.5 POLYMORPHIC UNIT (PMU)

A polymorphic gate is a reconfigurable component that can perform various logic functions non-conventionally, boosting processing capacity.

Moving toward polymorphic structures is due to their flexibility in implementing multiple processes and preventing security risks such as penetration and reverse engineering.

For this purpose, some control signals are also considered as inputs to the polymorphic unit and the primary input data. Different logic configurations can be selected for input processing by changing these control signals.

A new PMU design based on the IG-FINFET technology is proposed, as shown in Figure 5.3. This structure can be configured as a dual-purpose design that can perform different simple logic functions (logic mode) and be considered as a full-adder (full-adder mode).

This structure is inherently a full adder with four outputs, where signals A and B are intended as input data signals, and C can be either data or a control signal.

5.5.1 LOGIC MODE

Signal C acts as a control signal in this mode. If we set C to ‘0’ in the proposed PMU, the majority part (C_{OUT}/C_{OUT} outputs) acts as a two-input AND/NAND, and the SUM part (SUM/SUM outputs) realizes the two-input XOR/XNOR operations.

On the other hand, if the C signal is ‘1’, the majority part performs OR/NOR operations, and the SUM unit performs XNOR/XOR.

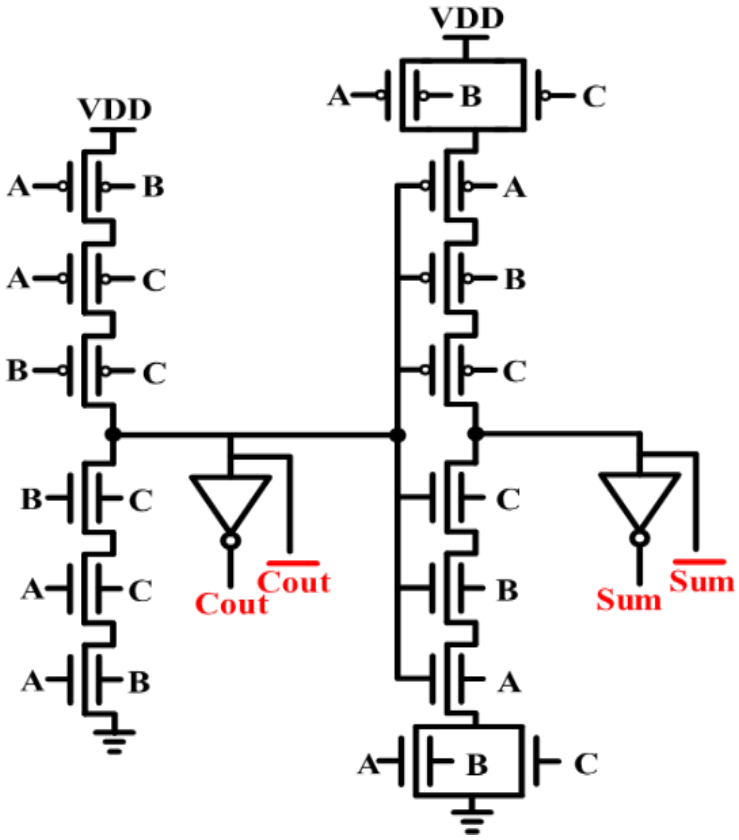


Figure 5.3 Polymorphic logic design

One of the data signals (A or B) should also be considered a control signal to implement the INV logic. For example, if C and A are set to ‘‘01’’ or ‘‘10’’,the output of the SUM part gives the inverted result.

Overall, seven different Boolean logic can be implemented with the proper setting of the C signal.

5.5.2 FULL-ADDER MODE

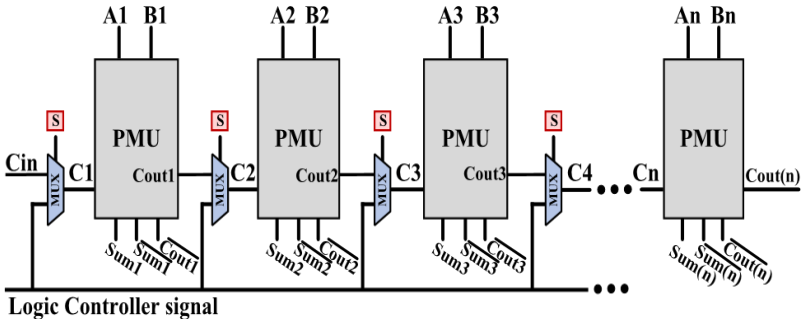


Figure 5.4 RCA structure based on PMU logic

As the proposed structure is inherently a full adder, the full adder outputs are generated in one cycle if all three inputs are considered data. Suppose the majority output of each polymorphic unit is used as the input C of the next polymorphic unit.

In that case, these blocks will form a ripple carry adder (RCA) for more complex calculations.

As demonstrated in Figure 5.4, the C input of each block must be signaled in different ways to change the configuration mode.

Therefore, it is necessary to embed a multiplexer on the input C of each block. Suppose the selector of this multiplexer switches to ‘1’.

The logic controller signal is connected to input C of each cell to perform the desired operations.

On the other hand, if the selector switches to ‘0’, The majority of the output of the previous block is connected to the current block, which makes the RCA available for processing.

CHAPTER 6
SOFTWARE DESCRIPTION

Software used to analyze the performance of the SRAM is **HSPICE** version 2019 and **Avanwaves (**commonly known as Awaves**)** used to plot the waveforms.

HSPICE is a widely used commercial electronic circuit simulation software developed by Synopsys. It is an advanced simulation tool used primarily for simulating and analyzing analog, digital, and mixed-signal circuits at the transistor level. HSPICE is renowned for its accuracy, speed, and extensive device modeling capabilities, making it a staple in the semiconductor industry for circuit design and verification.

6.1 KEY FEATURES OF HSPICE

1. Transistor-Level Simulation: HSPICE allows designers to simulate circuits at the transistor level, providing detailed insights into circuit behavior, performance, and characteristics.

2. Device Modeling: The software supports a wide range of device models, including MOSFETs, BJTs, diodes, resistors, capacitors, and inductors. These models accurately capture device physics and behavior for precise circuit analysis.

3. Mixed-Signal Simulation: HSPICE can handle mixed-signal designs, combining analog and digital components in the same simulation environment. This capability is crucial for verifying complex integrated circuits (ICs) and system-on-chip (SoC) designs.

4. Accuracy and Convergence: HSPICE is known for its accurate simulation results and robust convergence algorithms. It can handle challenging circuit configurations and non-linear behaviors, ensuring reliable simulation outcomes.

5. SPICE Compatibility: HSPICE follows the SPICE (Simulation Program with Integrated Circuit Emphasis) standard, ensuring compatibility with SPICE netlists and models. This allows users to leverage existing SPICE-based designs and libraries seamlessly.

6. Analysis Capabilities: The software offers a range of analysis options, including transient analysis, AC analysis, DC analysis, parametric sweeps, Monte Carlo simulations, sensitivity analysis, and more. These analyses help designers understand circuit performance under different operating conditions and variations.

7. Graphical Visualization: HSPICE provides graphical visualization tools for analyzing simulation results. Users can plot waveforms, gain insights into circuit behavior, and identify potential design issues or optimizations.

8. Integration with Design Flows: HSPICE integrates with popular electronic design automation (EDA) tools and design flows, allowing seamless transfer of designs, netlists, and simulation setups between different software platforms.

Overall, HSPICE is a powerful and versatile simulation tool used by engineers and designers across the semiconductor industry for verifying and optimizing electronic circuits, ensuring high performance, reliability, and functionality in integrated circuit designs.

CHAPTER 7

RESULTS AND DISCUSSIONS

7.1 TIMING DIAGRAM

A Timing diagram for an Half select 11T SRAM cell, Reconfigurable Half select SRAM cell and Polymorphic logic design, HSPICE simulations are performed as shown in Figure 7.1,7.2,7.3, which illustrates the various signals and their timing relationships during read and write operations.

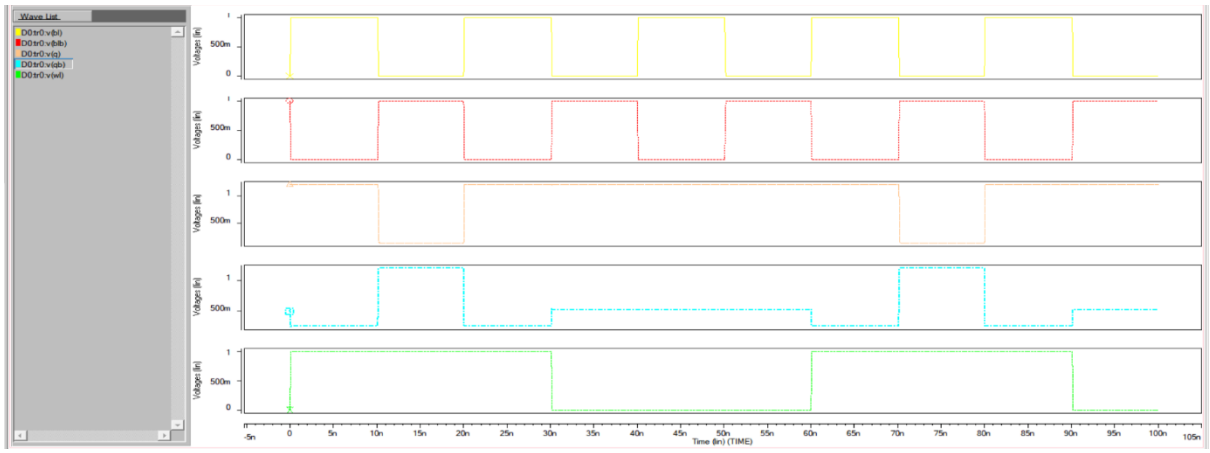


Figure 7.1 Transient waveform of Half Select 11T SRAM Cell

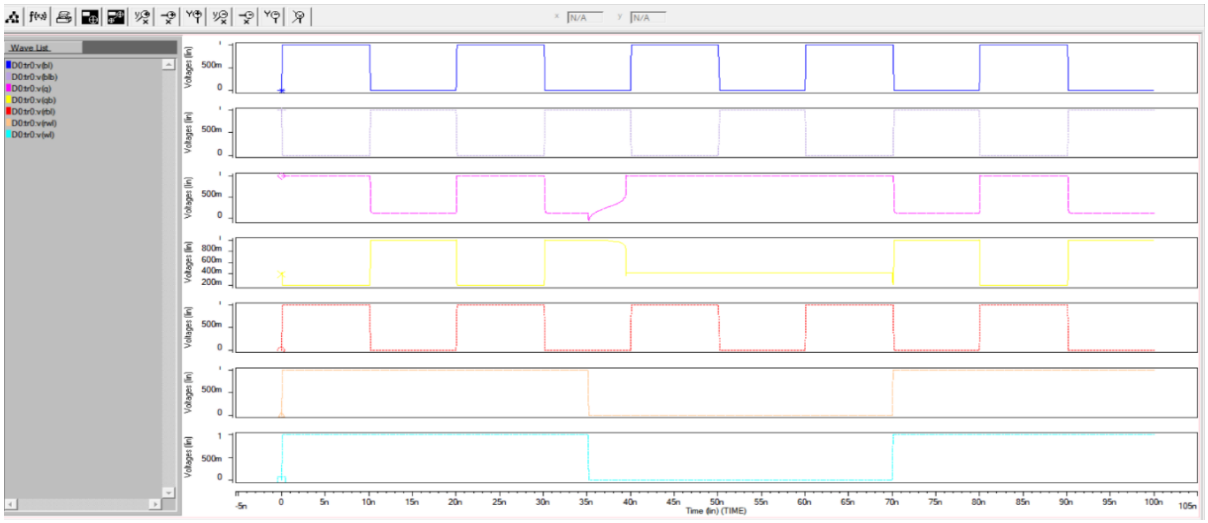


Figure 7.2 Transient waveform of Reconfigurable Half Select 11T SRAM Cell

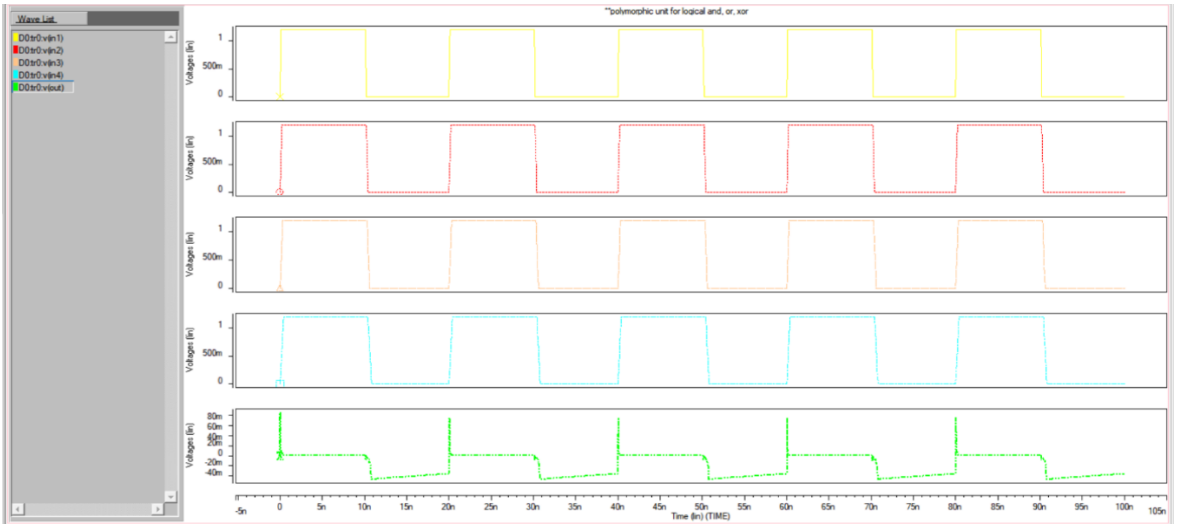


Figure 7.3 Transient waveform of Polymorphic logic design

7.2 STATIC NOISE MARGIN (SNM)

The static noise margin (SNM) of an SRAM (Static Random-Access Memory) cell refers to its tolerance to noise in maintaining stored data reliably.

The read margin determines how much voltage can be dropped on the bit-lines before a read operation fails.

The write margin refers to the voltage difference between the values of the bit-lines that are necessary to successfully write data into the cell. It's essentially the difference in voltage between the bit-line and the sensing voltage necessary for correct read operation.

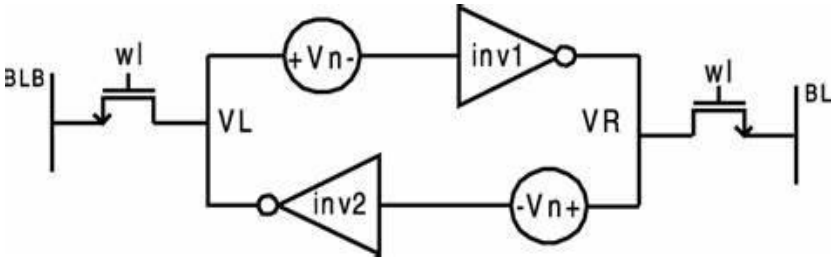


Figure 7.4 Measurement setup for Static Noise Margin

The two DC noise voltage sources V are placed in series with the cross-coupled inverters as shown in Figure 7.4, with worst-case polarity at the internal nodes VR and VL of the SRAM cell.

To evaluate the performance of proposed SRAM cell, HSPICE simulations are performed in terms of read/hold/write static noise margin, write/read power, and leakage power in 7nm FINFET Technology. Several SRAM cells are selected and implemented for comparisons, including the conventional 6T, 10T and 11T SRAM cells as shown in Table 7.1.

Table 7.1 Comparison of various cells in the 7nm FINFET Technology

PARAMETERS	6T SRAM CELL	10T SRAM CELL	PROPOSED 11T SRAM
RSNM (mv)	0.498	0.558	0.616
HSNM (mv)	0.082	0.091	0.112
WSNM (mv)	0.498	0.521	0.545
READ POWER (uw)	695.9	558.5	426.2
WRITE POWER (uw)	0.0183	0.521	16.70
HOLD POWER (mw)	0.303	0.294	0.277
WRITE PDP(m)	0.481	0.733	0.881
READ PDP(m)	0.859	0.975	1.328

This results indicates that the leakage power consumption in 11T SRAM cell is decreased by 38.7% and 23.6% compared to 6T, 10T SRAM.

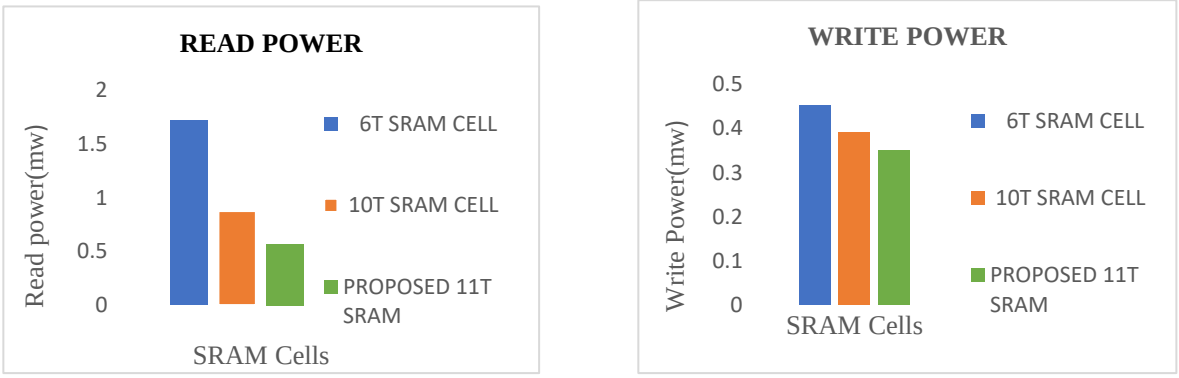
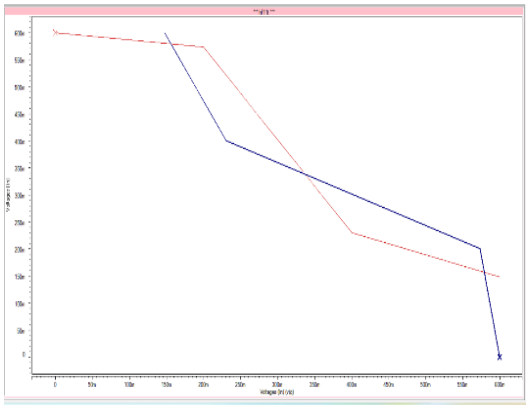
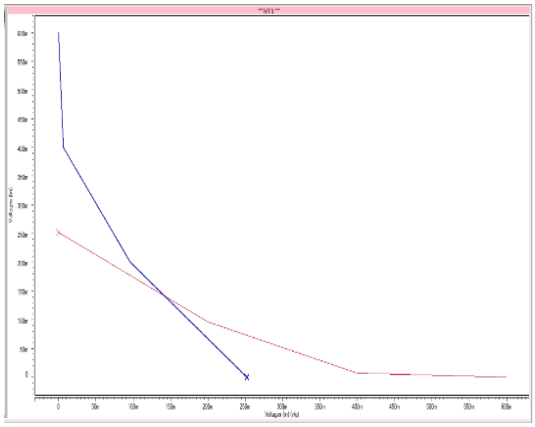


Figure 7.5 (a) Read power (b) Write power

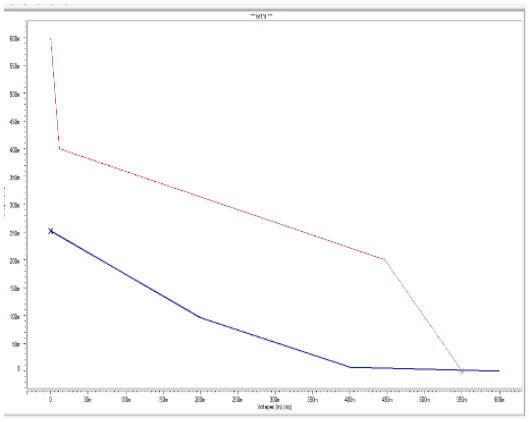
Hence, the power efficiency of 11T SRAM cell is greater than 6T and 10T SRAM cell.



(a)

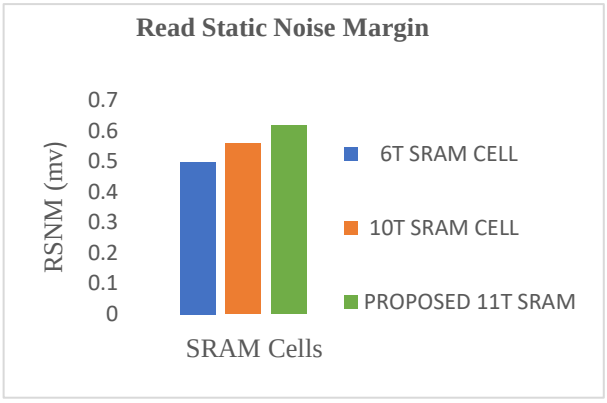


(b)

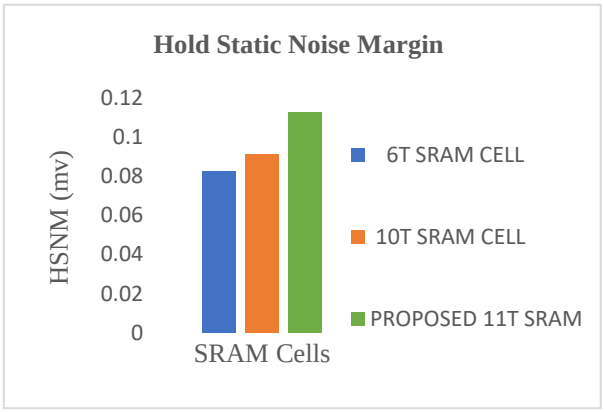


(c)

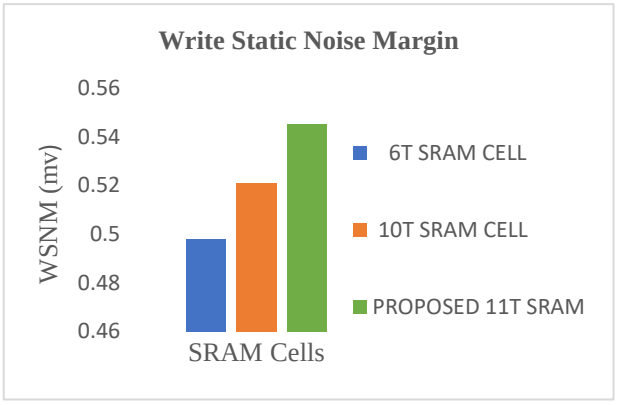
Figure 7.6 Butterfly diagrams of (a)RSNM (b) HSNM (c) WSNM



(a)



(b)



(c)

Figure 7.7 (a) RSNM (b)HSNM(c)WSNM

The SNM for SRAM cells is displayed in Fig 7.7, the Proposed 11T achieves 19.2% and 40.2% improvement in RSNM and WSNM respectively.

CHAPTER 8
CONCLUSION & SCOPE FOR FUTURE WORK

In-memory processing is a promising paradigm that improves the throughput and energy consumption, especially in data intensive applications. The proposed reconfigurable SRAM array can be configured as a sense amplifier to read data from RRAM memory and also can be configured as an SRAM cell for in-memory computation besides data storage ability. Moreover, the proposed reconfigurable SRAM cell has a high static noise margin and a free half-select issue which is essential for memory design. The simulation results indicate that the Proposed 11T achieves 19.2% and 40.2% improvement in RSNM and WSNM respectively, at a voltage of 1.2 V when compared to a 6T SRAM cell. Its Read static power consumption is also decreased by 38.7% and 23.6% compared to 6T, 10T SRAM Cell.In the proposed hybrid in-memory architecture, AND/NAND, OR/NOR operations can be performed in the RRAM main memory.

SCOPE FOR FUTURE WORK

The future scope for integrating a hybrid SRAM/RRAM in-memory computing architecture is vast and promising and some potential applications are,

- This technology could significantly enhance AI algorithms by providing faster processing speeds and improved memory density. It can accelerate tasks such as deep neural networks and executing complex AI models.
- It could enable researchers to process large amounts of biological data more efficiently, leading to advancements in healthcare and biotechnology.
- It could enable real-time analysis of large datasets and improve the accuracy of image recognition algorithms.
- Improved computational efficiency and memory density can enhance the High Performance Systems.

REFERENCE

[1] 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, doi: 10.1109/TCSI.2018.2848999.applications,’’ *IEEE Access*, vol. 10, pp. 48229–48241, 2022, doi: 10.1109/ACCESS.2022.3171799.

[2] F. Razi, M. H. Moaiyeri, and S. Mohammadi, ‘‘Toward efficient logicin-memory computing with magnetic reconfigurable logic circuits,’’ *IEEE Magn. Lett.*, vol. 13, pp. 1–5, 2022, doi: 10.1109/LMAG.2022.3146060.

[3] H. Kim, J. Mu, C. Yu, T. T. Kim, and B. Kim, ‘‘A 1-16b reconfigurable 80Kb 7T SRAM-based digital near-memory computing macro for processing neural networks,’’ *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 70, no. 4, pp. 1580–1590, Apr. 2023, doi: 10.1109/tcsi.2022.3232648.

[4] H. Zhang, W. Kang, K. Cao, B. Wu, Y. Zhang, and W. Zhao, ‘‘Spintronic processing unit in spin transfer torque magnetic random access memory,’’ *IEEE Trans. Electron Devices*, vol. 66, no. 4, pp. 2017–2022, Apr. 2019, doi: 10.1109/TED.2019.2898391.

[5] J. Chen, W. Zhao, Y. Wang, Y. Shu, W. Jiang, and Y. Ha, ‘‘A reliable 8T SRAM for high-speed searching and logic-in-memory operations,’’ *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 30, no. 6, pp. 769–780, Jun. 2022, doi: 10.1109/TVLSI.2022.3164756.

[6] J. Wang, X. Wang, C. Eckert, A. Subramaniyan, R. Das, D. Blaauw, and D. Sylvester, ‘‘A 28-nm compute SRAM with bit-serial logic/arithmetic operations for programmable in-memory vector computing,’’ *IEEE J. Solid-State Circuits*, vol. 55, no. 1, pp. 76–86, Jan. 2020, doi:10.1109/JSSC.2019.2939682.

[7] M. Bettayeb, F. Zayer, H. Abunahla, G. Gianini, and B. Mohammad, ‘‘An efficient in-memory computing architecture for image enhancement in AI

[8] M. F. Ali, A. Jaiswal, and K. Roy, ‘‘In-memory low-cost bit-serial addition using commodity DRAM technology,’’ *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 1, pp. 155–165, Jan. 2020, doi: 10.1109/TCSI.2019.2945617.

[9] M. K. Q. Jooq, M. H. Moaiyeri, and K. Tamersit, ‘‘A new design paradigm for auto-nonvolatile ternary SRAMs using ferroelectric CNTFETs: From device to array architecture,’’*IEEE Trans. Electron Devices*,vol.69,no.11, pp. 6113–6120, Nov. 2022, doi: 10.1109/TED.2022.3207703.

[10] M. T. Nasab, A. Amirany, M. H. Moaiyeri, and K. Jafari, ‘‘Highperformance and robust spintronic/CNTFET-based binarized neural network hardware accelerator,’’ *IEEE Trans. Emerg. Topics Comput.*, vol. 11, no. 2, pp. 527–533, Apr./Jun. 2022, doi: 10.1109/TETC.2022. 3202113.

[11] Q. Dong, S. Jeloka, M. Saligane, Y. Kim, M. Kawaminami, A. Harada, S. Miyoshi, M. Yasuda, D. Blaauw, and D. Sylvester, ‘‘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.

[12] S. Angizi, Z. He, A. Chen, and D. Fan, ‘‘Hybrid spin-CMOS polymorphic logic gate with application in in-memory computing,’’ *IEEE Trans. Magn.*, vol. 56, no. 2, pp. 1–15, Feb. 2020, doi: 10.1109/TMAG.2019.2955626.

[13] 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.

[14] S. Yin, Y. Kim, X. Han, H. Barnaby, S. Yu, Y. Luo, W. He, X. Sun, J. Kim, and J. Seo, ‘‘Monolithically integrated RRAM- and CMOS-based in-memory computing optimizations for efficient deep learning,’’ *IEEE Micro*, vol. 39, no. 6, pp. 54–63, Nov. 2019, doi: 10.1109/MM.2019.2943047.

[15] V. Jamshidi, A. Patooghy, and M. Fazeli, ‘‘MagCiM: A flexible and non-volatile computing-in-memory processor for energy-efficient logic computation,’’ *IEEE Access*, vol. 10, pp. 35445–35459, 2022, doi: 10.1109/ACCESS.2022.3159967.

[16] Y. Kong, X. Chen, X. Si, and J. Yang, ‘‘Evaluation platform of time-domain computing-in-memory circuits,’’ *IEEE Trans. Circuits Syst.II, Exp. Briefs*, vol. 70, no. 3, pp. 1174–1178, Mar. 2023, doi:10.1109/TCSII.2022.3229437.

[17] Y. Park, Z. Wang, S. Yoo, and W. D. Lu, ‘‘RM-NTT: An RRAM-based compute-in-memory number theoretic transform accelerator,’’ *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 8, no. 2, pp. 93–101, Dec. 2022, doi: 10.1109/JXCDC.2022.3202517.

[18] Z. Lin, H. Zhan, X. Li, C. Peng, W. Lu, X. Wu, and J. Chen, ‘‘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.

[19] Z. Yang, K. He, Z. Zhang, Y. Lu, Z. Li, Y. Wang, Z. Wang, and W. Zhao, ‘‘A novel computing-in-memory platform based on hybrid spintronic/CMOS memory,’’ *IEEE Trans. Electron Devices*, vol. 69, no. 4, pp. 1698–1705, Apr. 2022, doi: 10.1109/TED.2021.3137761.

SIMATS ENGINEERING

Approved By AICTE | IET-UK Accreditation

Certificate of Presentation

This is to certify that

Dr / Mr / Miss / Mrs Shruthi. B. K

Thanthai Periyar Government Institute of Technology.

Have successfully presented the paper entitled

Energy Efficient In-memory Computing Architecture with Enhanced polymorphic Circuit.

at the ***International Conference on Recent Trends in Computing and Communication Engineering (ICRTCCE-2024)*** organized by Department of Wireless Communication SIMATS School of Engineering. Chennai.Tamil Nadu. India held on 23rd March 2024.

HoD
Dr. A. Mary Joy Kinol

Principal
Dr. B. Ramesh


SIMATS ENGINEERING
ICRTCCE 2K24
23.03.2024
BEST
PRESENTER
SAVEETHA