CHAPTER 3

HYBRID SRAM CELL DESIGN FOR NORMALLY-OFF APPLICATIONS

3.1 Introduction

With the increasing performance requirements of next-generation applications, the demand for memories is also continuously growing [1]. However, conventional memories such as SRAM, DRAM, and Flash face severe scaling challenges with the shrinking technology nodes. With smaller dimensions, the increasing effect of process variation severely impacts several memory design constraints such as read, write and hold margins. Moreover, the leakage power, which becomes a dominating factor in the total power consumption, also increases with the technology downscaling [2]. This is because SRAM-based memories such as caches, registers, flipflops, and latches always require a constant supply, which adds to the total power consumption even when it is not performing any operation. In recent times, normally-off computing emerges as one of the most promising solutions to address the challenge of increasing static power consumption. In normally-off computing, the system is put into a deep sleep mode by completely gating the power supply to switch off the circuit during the idle mode. As a result, zero power consumption during the standby mode is achieved. However, direct gating of the power supply causes the system to lose its current state. In order to wake up and re-initialize the state, the system requires an energy-intensive boot process. Since the boot process requires thousands of processor cycles, hundreds of microseconds, and hundreds of nanojoules of energy, it becomes limiting in an event-based application where quick wake-up is required.

In addition to large static power consumption, IoT nodes that rely on energy harvesting suffer from the unsteady and uncertain nature of power supply which causes frequent and unpredictable power failures [3]-[16]. In both the 'power off' and 'power failure' cases, the system loses its current state, thus requires an energy-intensive boot process for re-initialization. Also, completing a task may take a longer time since intermediate results could not be saved [4]. These problems have attracted researchers both from the industry and academia to propose some techniques that can be augmented with these systems to ensure continuous progress. The technique of recent state checkpoint ensures continuous computation

by preventing rollbacks [5]. However, in conventional checkpointing techniques, as shown in Figure 3.1, data is backed up from volatile (e.g., SRAM, Flip-Flops) memory to non-volatile memory (e.g., FLASH). However, one-bit transfer to or from an external non-volatile storage element requires 10^7 – 10^8 processor cycle and 1000x more power than cache [6]-[9]. Hence, a slow and energy-intensive conventional checkpoint scheme can still result in reset and rollbacks under an unstable power supply.

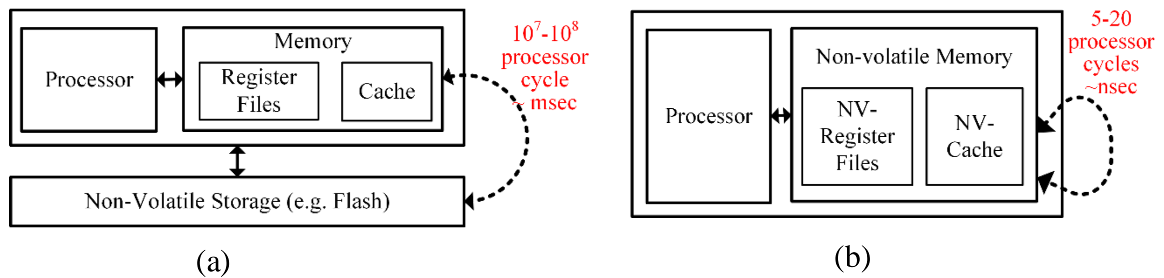

Figure 3.1: Different checkpointing techniques employed in energy harvesting based IoT applications (a) Conventional checkpointing scheme (b) Checkpointing scheme with NVM hierarchy

Whereas the non-volatile memory (NVM) architecture with non-volatile registers and non-volatile memory can maintain the temporary state and quickly resume the computation once the power is restored [17]. Therefore, the NVM checkpoint scheme opens up the possibility to compute continuously with the intermittent power, as shown in Figure 3.2.

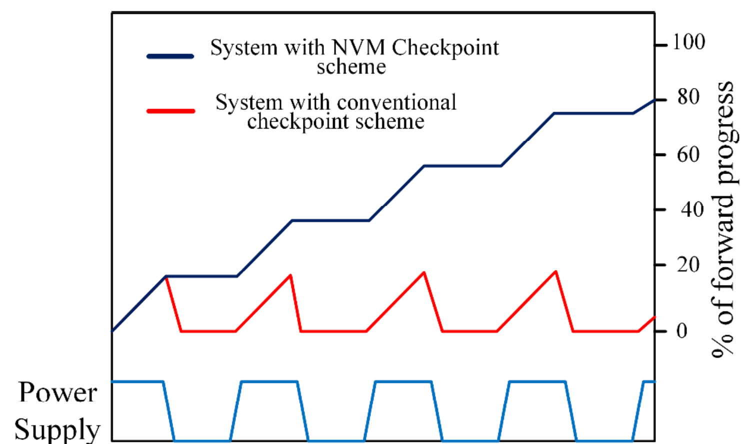

Figure 3.2: Percentage of the forward progress with conventional architecture and NVM checkpointing scheme under unstable power supply [10]–[12]

3.2 Hybrid Memory

Non-Volatile Memory (NVM) architecture can be realized by incorporating emerging non-volatile technologies into the register and cache memory. However, direct replacement of

SRAM and Flip-Flops with pure non-volatile elements degrades the system performance due to their high write energies and large latencies [18]-[24]. Hybrid memory can be a better solution when compared to pure non-volatile cache architecture as it reduces the number of write cycles into the non-volatile element. Hybrid memories have a unique architecture consisting of emerging non-volatile devices and normal cells, offering high speed by isolating the non-volatile element from the conventional cell during normal operation and data storage in local non-volatile elements during power-down periods.

Cache memories occupy more than 50% of the die area and consume 40% of the total power required by the system [25]. As a result, power and performance are critically dependent on its design [26]. Therefore, in this work, we focus on designing a hybrid cache to improve overall performance. Emerging non-volatile elements that can be integrated with CMOS SRAM to achieve ultra-low power operation includes phase change device, ferroelectric capacitor device, memristor device, and magnetoresistive device [27]. Among all non-volatile devices, magnetoresistive device (magnetic tunnel junction) seems more promising because of its low switching currents, smaller switching times, unlimited endurance, and compatibility with the CMOS fabrication process [28]-[34]. Therefore, in this chapter, we propose to design an energy-efficient solution to IoT nodes using a hybrid MTJ/CMOS SRAM cell, which offers high performance of CMOS SRAM with the added advantage of non-volatility.

3.3 Proposed Hybrid SRAM Cells

This section presents proposed non-volatile hybrid SRAM cells for 'Normally OFF - Instant ON' applications. The normally-off and instant-on architecture eliminate the large standby mode power dissipation by turning off the SRAM when not in use and instantly turning it on with full performance when needed. In order to effectively realize the normally-off and instant-on technique, we propose a non-volatile hybrid SRAM, which preserves the state in the locally embedded non-volatile element during the power-off period to provide a faster wake-up time.

3.3.1 Hybrid 7T SRAM Cell

We propose a non-volatile hybrid 7T SRAM cell that retains the state during the power-off period and sudden power failure. The data in the proposed non-volatile SRAM cell is held by both internal storage nodes (q & qc) and the embedded non-volatile magnetoresistive element (spin-transfer torque-magnetic tunnel junction (STT-MTJ)).

3.3.1.1 Structure of proposed hybrid 7T SRAM cell

The proposed hybrid 7T SRAM cell, as shown in Figure 3.3, modifies the conventional 6T SRAM cell using the non-volatile magnetoresistive element (STT-MTJ) to incorporate non-volatility. An MTJ pair is utilized per SRAM cell to preserve the state where one of the MTJ devices (MTJ1a) is connected to bitline BL, and the other MTJ device (MTJ1b) is connected to bitline BL_bar. In addition to 6T SRAM core, an equalization transistor is connected between the internal nodes (q and qc) of the SRAM cell. The data written into the SRAM cell is held by internal nodes as well as the MTJ pair. A logic '0' is represented by the high resistance state (HRS) of the MTJ1a device and the low resistance state (LRS) of the MTJ1b device. Similarly, a logic '1' is represented by the low resistance state (LRS) of the MTJ1a device and the high resistance state (HRS) of the MTJ1b. The main advantage of the proposed hybrid SRAM cell is the retention of intermediate computation status before sudden power failure/off.

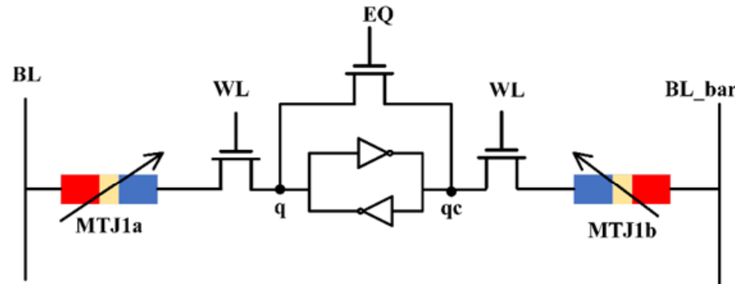

Figure 3.3: Proposed hybrid 7T SRAM Cell

3.3.1.2 Operational modes of proposed hybrid 7T SRAM Cell

The proposed 7T SRAM cell works in three operational modes: Write/Backup, Read and Restore. The write/backup operation of the proposed cell refers to writing into the MTJ device and the cell internal nodes q and qc. Since the write/backup operation store the current data into the MTJ device, there is no need to run a separate backup operation before turning OFF the memory. Like conventional SRAM cell, read operation in the proposed hybrid SRAM cell refers to reading the value stored by the cell nodes q and qc. The restore operation is performed after power restoration to resume the execution from the point of interruption. The details of all the operational mode are discussed in the following sub-sections:

3.3.1.2.1 Write/Backup operation

In this mode of operation, the input data value is simultaneously written into the cell node and the MTJ device through the bitlines BL and BL_bar. We propose two different write/backup operation mechanisms for the proposed hybrid cell: Single-phase and Two-phase operation.

***Single-phase operation*:** In a single-phase operation, writing into the internal nodes of the cell and the MTJ device is completed in a single step. The status of all the control signals for the write/backup operation is tabulated in Table 3.1. For writing logic '1', bitline BL is charged to Vdd while BL_bar is pulled to the gnd. Now asserting control signal WL and EQ high provides a current path from MTJ1a→q→qc→MTJ1b switching MTJ1a from LRS to HRS and MTJ1b from HRS to LRS as shown in Figure 3.4.

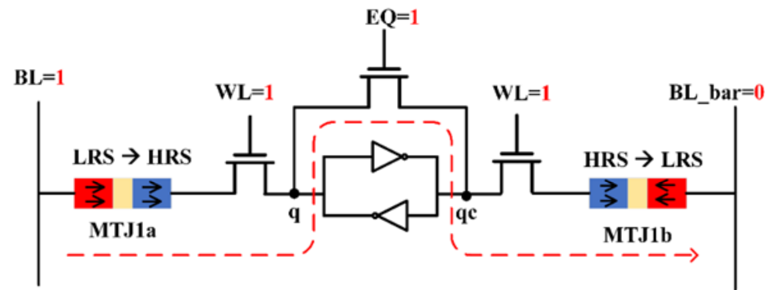

Figure 3.4: Writing logic '1' into the proposed hybrid 7T SRAM cell using single-phase write operation

Two-phase operation: In a two-phase operation, the input data value is first transferred to the MTJ device and then transferred to the cell nodes in two different steps. The status of all the control signals for the write/backup operation is tabulated in Table 3.1. The two-phase write/backup operation is performed in the following two phases:

Phase I: During this phase, the data value is written into the MTJ device, as illustrated in Figure 3.5. Let us assume cell initially stores a logic '0' where node q stores logic '0' and node qc stores logic '1'. For writing logic '1', bitline BL is charged to Vdd while BL_bar is pulled to gnd. Meanwhile, wordline WL is asserted high, allowing current to flow from bitline BL to node q through MTJ1a and from node qc to bitline BL_bar through MTJ1b. The cell is sized such that the opposite value on bitlines will not flip the data held by internal nodes. At the completion of this step, MTJ1a is switched from LRS to HRS, conversely MTJ1b is switched from HRS to LRS.

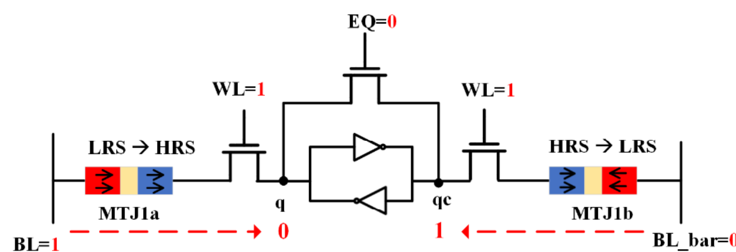

Figure 3.5: Phase I: Writing logic '1' into the MTJ device of the hybrid 7T SRAM cell using two-phase write operation

Phase II: In this phase, the data value is transferred to the cell internal nodes q & qc, as shown in Figure 3.6. The control signal EQ is asserted high to trigger this phase. Meanwhile, both the bitlines are discharged to ground. After attaining equal voltages (Vx) at nodes q & qc, control signal EQ is asserted low. Now, asserting WL high causes current to flow from nodes to bitlines through the MTJ device. The potential drop at the node depends on the resistance state of the MTJ device. The MTJ device with LRS (low resistance state) will cause the corresponding node to fall very quickly, thus creating a differential voltage. This differential potential at the node is pulled to full swing by cross-coupled action.

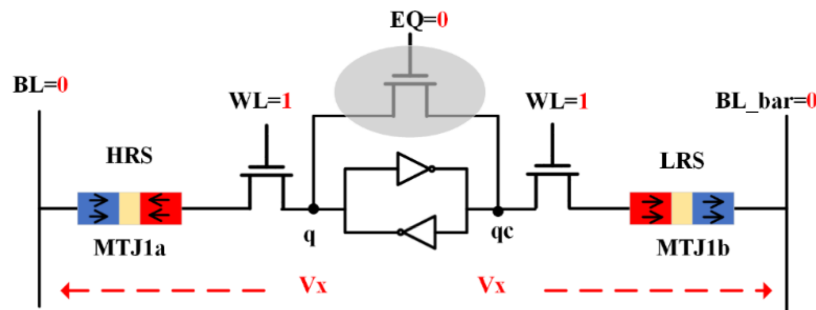

Figure 3.6: Phase II: Writing logic '1' into the internal nodes of the hybrid 7T SRAM cell using two-phase write operation

The two-phase write/backup scheme results in lower power consumption than a single-phase backup operation because of the large potential difference across the MTJ device during phase-I, which causes the MTJ device to change its state quickly. The major advantage of a two-phase write operation is power saving while writing the same bit. During phase I, bitline and node are at the same potential if same bit is to be written. Therefore, no current flows in the circuit for complete phase I duration resulting in lower power consumption.

TABLE 3.1: STATUS OF CONTROL SIGNAL DURING VARIOUS OPERATIONAL MODES OF PROPOSED HYBRID 7T SRAM CELL

		WL	**BL**	**BL_bar**	**EQ**
Single-phase Write/Backup		vdd	vdd/gnd	gnd/vdd	vdd
Two-phase Write/Backup	Phase-I	vdd	vdd/gnd	gnd/vdd	gnd
	Phase-II	vdd	gnd	gnd	Short-pulse
Read		vdd	pre-charged	pre-charged	gnd
Restore		vdd	gnd	gnd	Short-pulse

3.3.1.2.2 Read operation

During the read mode, the data value stored by cell internal nodes q and qc is transferred onto the bitlines BL and BL_bar. The operation starts with pre-charging the bitlines to Vdd. After pre-charging bitlines, wordline WL is asserted high, causing one of the bitline to discharge, as shown in Figure 3.7.

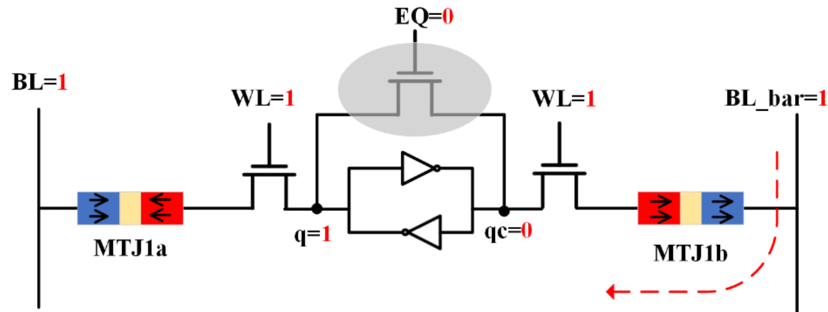

Figure 3.7: Reading Logic '1' in hybrid 7T SRAM cell

3.3.1.2.3 Restore operation

After the power supply is up, a restore operation is performed to retrieve the data stored in the MTJ device into the cell nodes q & qc. During the restore operation, the control signal EQ is asserted high to equalize the nodes q & qc at Vx voltage. Meanwhile, both the bitlines are discharged to ground. After attaining equal voltages at nodes q & qc, control signal EQ is asserted low. Now asserting WL causes current to flow from nodes to bitlines through MTJ, as shown in Figure 3.8. The potential drop at the node depends on the resistance state of the MTJ device. The MTJ device with LRS (low resistance state) will cause the corresponding node to fall very quickly, thus creating a differential voltage. This differential potential at the node is pulled to full swing by cross-coupled action.

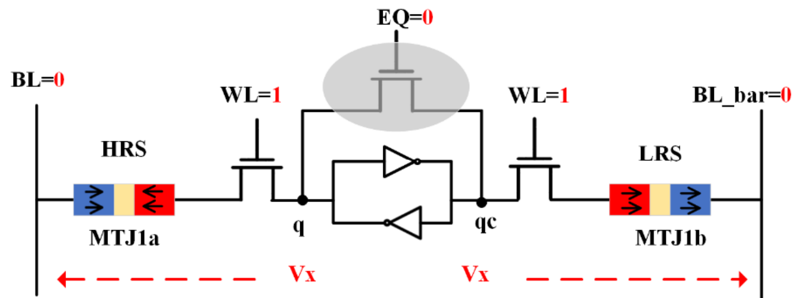

Figure 3.8: Restore operation in hybrid 7T SRAM cell

3.3.1.3 Circuit-level analysis

In this sub-section, we validate the functionality of the proposed hybrid SRAM cell by performing extensive simulations. The simulations are carried out using the SPICE circuit

simulator. Figure 3.9 plots the transient waveform of the proposed hybrid 7T SRAM cell with all the operational modes: write/backup, read and restore operation. The simulation begins with a two-phase write/backup operation, where we write logic '1' into the cell and MTJ device. The initial value of internal nodes q/qc is logic '0'/'1', as depicted in Figure 3.9. For writing logic '1', bitline BL is charged to vdd, and BL_bar is discharged to gnd. During phase I, wordline WL is asserted high, which causes current to flow between bitline and nodes. As a result, MTJ1a is switched from LRS to HRS. Conversely, MTJ1b is switched from HRS to LRS. Next, during phase-II, the control signal EQ is asserted high to equalize both the nodes (q and qc). At the same time, both the bitlines (BL and BL_bar) are discharged to gnd. The wordline WL is then asserted high. Because of the resistance difference between both the MTJs (MTJ1a and MTJ1b), node q stabilizes at logic '1' and node qc stabilizes at logic '0' Next read operation is performed where BL remains at vdd and BL_bar discharges resulting in logic '1' at the output. After the read operation, the power supply is disconnected to represent power-down or power-failure period. Finally, we perform restore operation to retrieve the saved state. The EQ signal is again asserted high to equalize the nodes (q and qc) while bitlines (BL and BL_bar) are discharged to gnd. With EQ signal de-asserted low and wordline WL asserted high, current start flowing from the nodes to the bitlines where its magnitude depends upon MTJ resistance in the restore path.

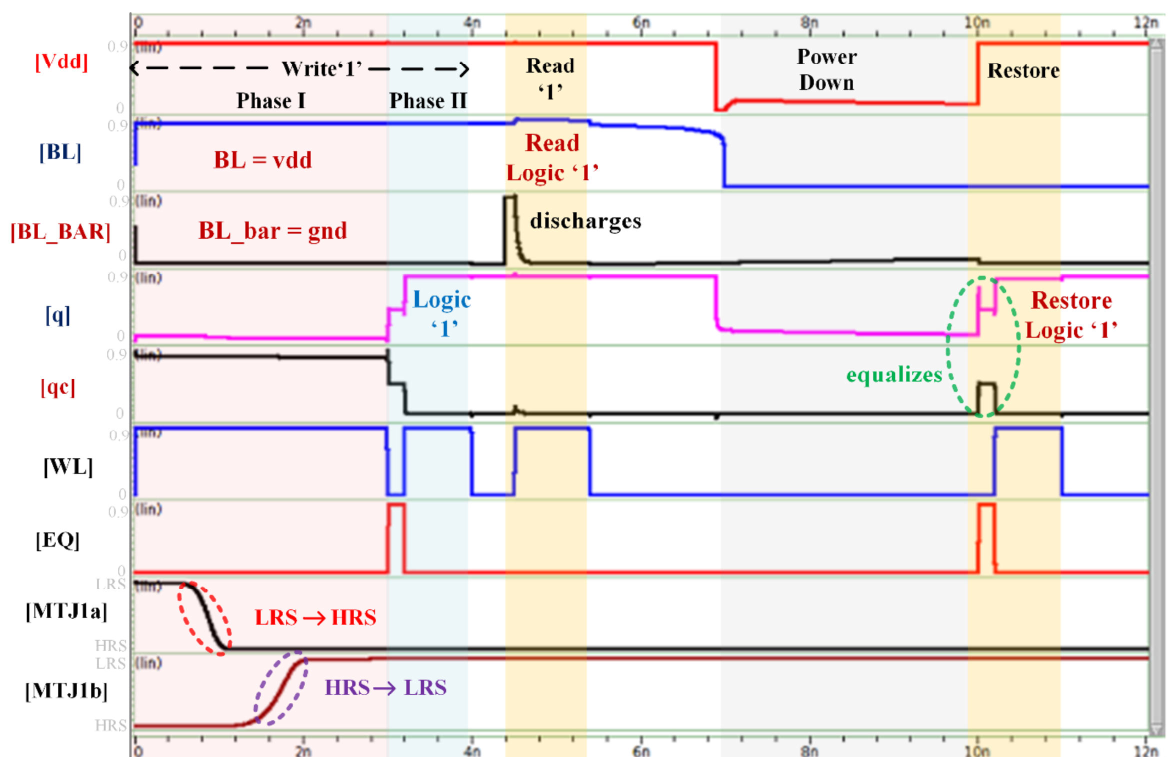

Figure 3.9: Transient waveform of proposed non-volatile hybrid 7T SRAM Cell

Therefore, based on the state stored by the MTJ device, node q and qc take logic ‘1’ and logic ‘0’ respectively at the output hence verifying the successful retention and restoration of the data. The energy consumption and access latency of the proposed hybrid SRAM cell during write/backup, read and restore operational modes are tabulated in Table 3.2. During the write/backup operation, the data value is simultaneously stored in the MTJ and the SRAM storage nodes, resulting in higher energy consumption and latency compared to the read and restore operation. The energy consumption during the write/backup operation is tabulated for two different cases: writing opposite bit and writing same bit. The first case of writing opposite bit refers to overwriting the previously stored data value in the SRAM cell with complementary/opposite input data value, whereas the second case of writing same bit refers to writing the same input data value into the previously stored value in the cell. During single-phase write operation, the equalization transistor is turned ON for the complete write/backup duration, which results in the formation of a short circuit path from bitlines to ground. Therefore, a large amount of current flow increases the energy consumption during single-phase write operation. The same amount of current flows through the cell for both cases; hence, the energy consumption of single-phase write operations remains the same for both cases. On the other hand, if the same bit is to written using the two-phase operation, bitlines and nodes are at same potential. Therefore, no current flows in the circuit for complete phase I duration resulting in lower power consumption. Furthermore, the high restore energy is attributed to the equalization process, where a direct path from vdd to gnd is created for a very short period.

TABLE 3.2: ENERGY CONSUMPTION AND ACCESS LATENCY OF PROPOSED NON-VOLATILE HYBRID 7T SRAM CELL

		Energy Consumption (fJ)		**Latency (ns)**
		While writing opposite bit	**While writing the same bit**	
Write/Backup Operation	**Single -Phase Write Operation**	1720	1720	4n
	Two-Phase Write Operation	478	1.2	4n
Read Operation		7.9		0.148
Restore Operation		109		0.210

Figure 3.10 (a) compares the energy consumption of single-phase and two-phase write operations at different technology nodes. For the lower technology node, the two-phase write operation shows upto 3X improvement in the energy consumption compared to the single-phase write operation.

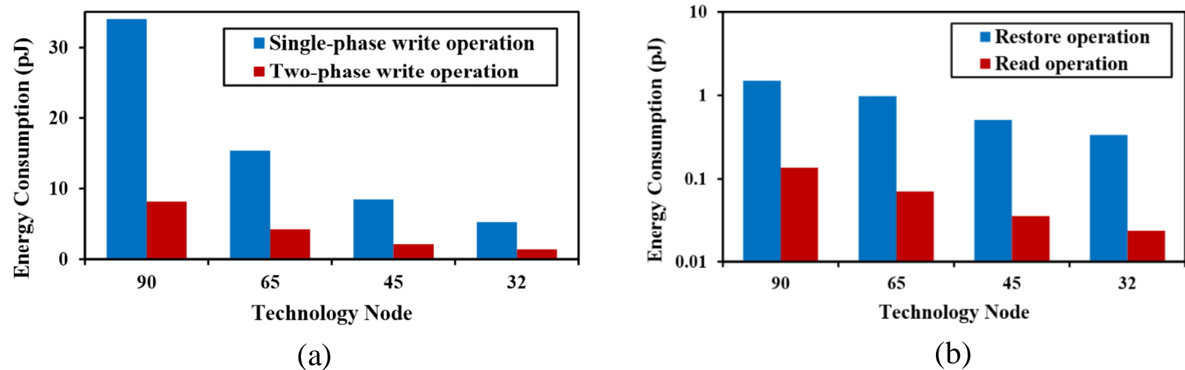

Figure 3.10: Energy consumption during the various operation in proposed hybrid 7T SRAM cell at different technology nodes (a) Energy consumption during the write/backup operation in single phase and two phase write scheme (b) Energy consumption during the read and restore operation.

The energy consumption of the read and restore operation of the proposed 7T hybrid SRAM cell at different technology nodes is compared in Figure 3.10 (b). Due to the equalization process, the energy consumption during restore operation at all the scaled technology nodes is approximately 70 % greater when compared to read operation. Figure 3.11 compares the access latency while performing write/backup operation, restore and read operation in proposed hybrid 7T SRAM cell at different technology nodes. The write/backup latency in different technology nodes is one order of magnitude higher than the restore and read latency due to the large switching time of the MTJ device.

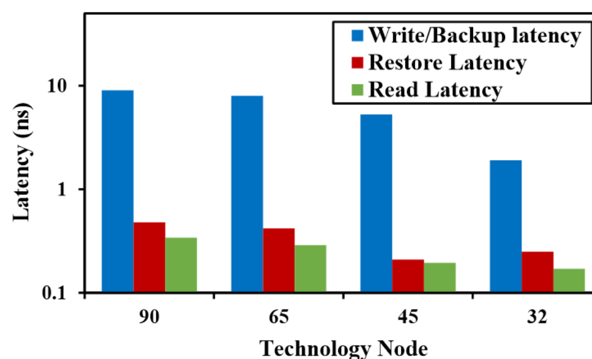

Figure 3.11: Access latency while performing the write/backup, read and restore operation in the proposed hybrid 7T SRAM cell at different technology nodes

3.3.2 Hybrid 8T SRAM Cell

3.3.2.1 Structure of hybrid 8T SRAM cell

Figure 3.12 presents the schematic of the proposed hybrid 8T SRAM cell. It consists of a standard 6T SRAM cell with two MTJ devices (MTJ1a & MTJ1b) and two additional transistors XE1 for isolation & XE2 for equalization.

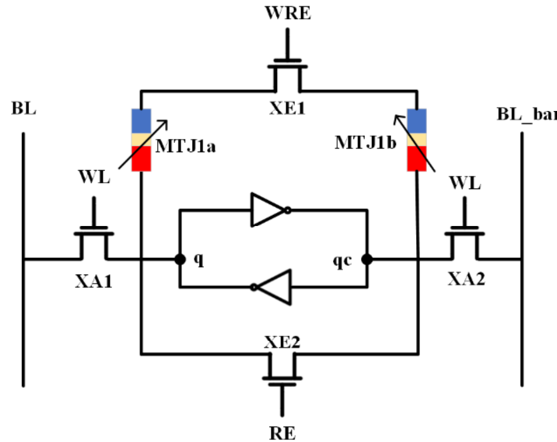

Figure 3.12: Proposed hybrid 8T SRAM cell

The data bit is stored in cross-coupled inverters, which is accessed similar to a conventional SRAM cell. Whereas, the data is stored in the MTJ before power off. A logic '0' is represented by high resistance state (HRS) of MTJ1a and low resistance state (LRS) of MTJ1b. Similarly, a logic '1' is represented by the low resistance state (LRS) of MTJ1a and the high resistance state (HRS) of MTJ1b.

3.3.2.2 Operational modes of hybrid 8T SRAM cell

The proposed hybrid memory has four different operations: Write, Read, Backup, and Restore operation. During the read and write operations, the transistors XE1 and XE2 are turned OFF to disconnect the MTJ device from the circuit. As a result, the proposed 8T SRAM cell act as a conventional 6T SRAM cell during these operational modes. The additional operations, backup, and restore operations are performed before and after the power-off period to effectively implement 'Normally-OFF & Instant-ON' computing. In the following sub-section, the details of each operation mode are presented.

3.3.2.2.1 Write operation

During the write operation, the data value is written into the proposed hybrid 8T SRAM cell. The status of various control signals during the write operation is tabulated in Table 3.3. The voltage at the bitlines is set based on the input data value to be written. For example, the bitline

BL is set to vdd/gnd, and BL_bar is set to gnd/vdd for writing logic '1'/logic '0' into the cell. The wordline WL is then enabled to connect bitlines and transfer the data value to the cell internal nodes q and qc.

3.3.2.2.2 Read operation

During the read operation, the stored data value is read from the proposed hybrid 8T SRAM cell. The status of various control signals during the read operation is tabulated in Table 3.3. The bitlines are pre-charged to vdd for performing the read operation. The wordline WL is then asserted high to connect bitlines to the cell internal nodes q and qc. Based on the data value stored in the internal nodes, one of the bitline discharges creating a differential voltage at the bitlines. This differential voltage is applied to the sense amplifier to generate the valid logic output.

3.3.2.2.3 Backup operation

During the backup operation, the hybrid 8T SRAM cell is first isolated from the bitlines by turning OFF the access transistors XA1 and XA2, and then the state of the memory cell is written into the MTJ device, as shown in Figure 3.13. The control signal WRE is enabled to connect both the MTJ device. It remains logic high for the complete duration of the backup operation. Since writing into the MTJ is an asymmetric operation, i.e., the transition from LRS (low resistance state) → HRS (high resistance state) takes more time than HRS→LRS, therefore pulse width of the control signal WRE depends on the switching time of LRS→HRS of MTJ.

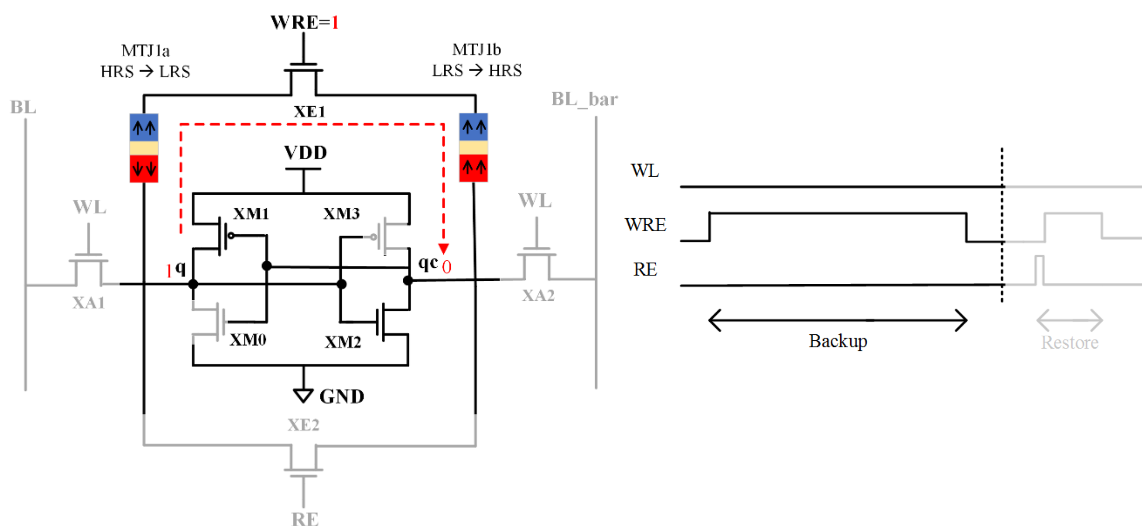

Figure 3.13: Direction of current flow and status of control signal during backup operation in hybrid 8T SRAM cell

Writing into the MTJ device depends on the direction of current flow, which is dependent on the value stored in the nodes q and qc. For example, if node q stores logic ‘1’ and node qc stores logic ‘0’, then the current flows from q→MTJ1a→MTJ1b→qc storing LRS to MTJ1a and HRS to MTJ1b. The status of control signals during backup operation is also shown in Figure 3.13.

TABLE 3.3: STATUS OF CONTROL SIGNALS DURING VARIOUS OPERATION IN HYBRID 8T SRAM CELL

	WL	**BL**	**BL_bar**	**WRE**	**RE**
Write	vdd	vdd/gnd	gnd/vdd	gnd	gnd
Read	vdd	pre-charged	pre-charged	gnd	gnd
Backup	gnd	vdd	vdd	vdd	gnd
Restore	gnd	vdd	vdd	vdd	Short-pulse

3.3.2.2.4 Restore operation

During the restore operation, the value stored in MTJ device is reloaded into the cell nodes q and qc. The restore operation starts with equalizing the node voltages by turning ‘ON’ the equalization transistor XE2. The control signal RE is pulled to vdd to enable access transistor XE2. Since equalized nodes form a short circuit path between vdd and gnd, the duration of control signal RE is kept small, as shown in Figure 3.14. Subsequently, asserting the control signal WRE high while nodes q and qc are at the same voltage will restore the value at the nodes based on the resistance difference of the MTJ device.

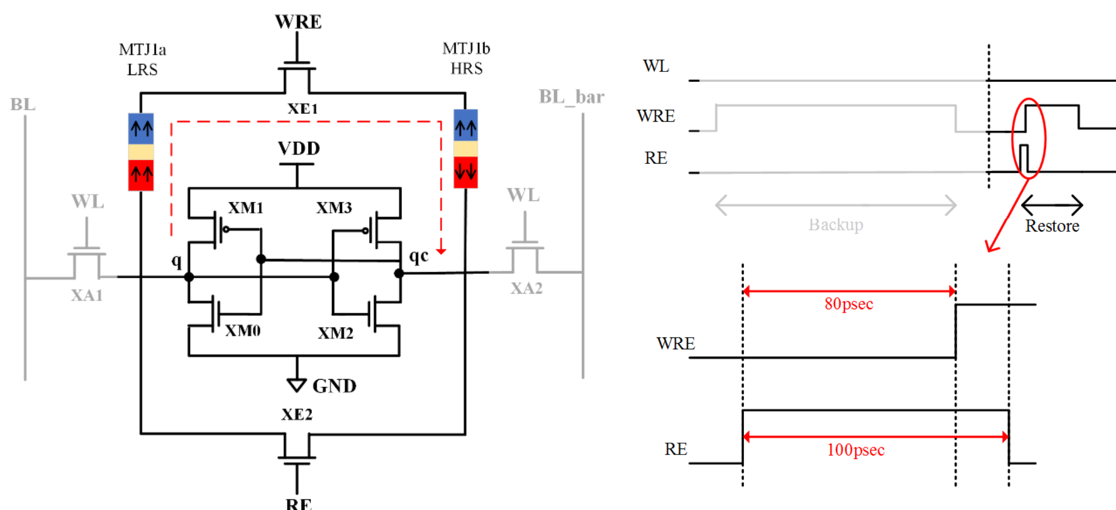

Figure 3.14: Restore operation in hybrid 8T SRAM cell

However, if the control signal WRE is asserted high after RE goes low, it will result in node voltages q and qc to change irrespective of the MTJ device state. Therefore, the control signal WRE is asserted high slightly before RE goes low to provide a current path [q→ MTJ1a→XE1→MTJ1b→qc] which creates differential voltage across nodes. This differential voltage is then translated to full swing by the cross-coupled inverter action. Figure 3.14 shows the restoration of logic '1' where MTJ1a is in LRS, and MTJ1b is in HRS, resulting in logic '1' being restored at node q and logic '0' at node qc.

3.3.2.3 Circuit-level analysis

To verify the correct functionality of the hybrid 8T SRAM cell at the circuit level, we simulate the proposed SRAM cell using the SPICE simulator. Figure 3.15 plots the transient waveform of the proposed 8T SRAM cell. The simulation starts from the write operation in which input data is written into the storage nodes q and qc, followed by a read operation that reads the newly written value. Then, a backup operation is performed before power down to store the value of the cell into the MTJ device. Finally, to check the non-volatility, the restore operation is performed after a power-failure, which shows that the restored value is the same as the backed-up value.

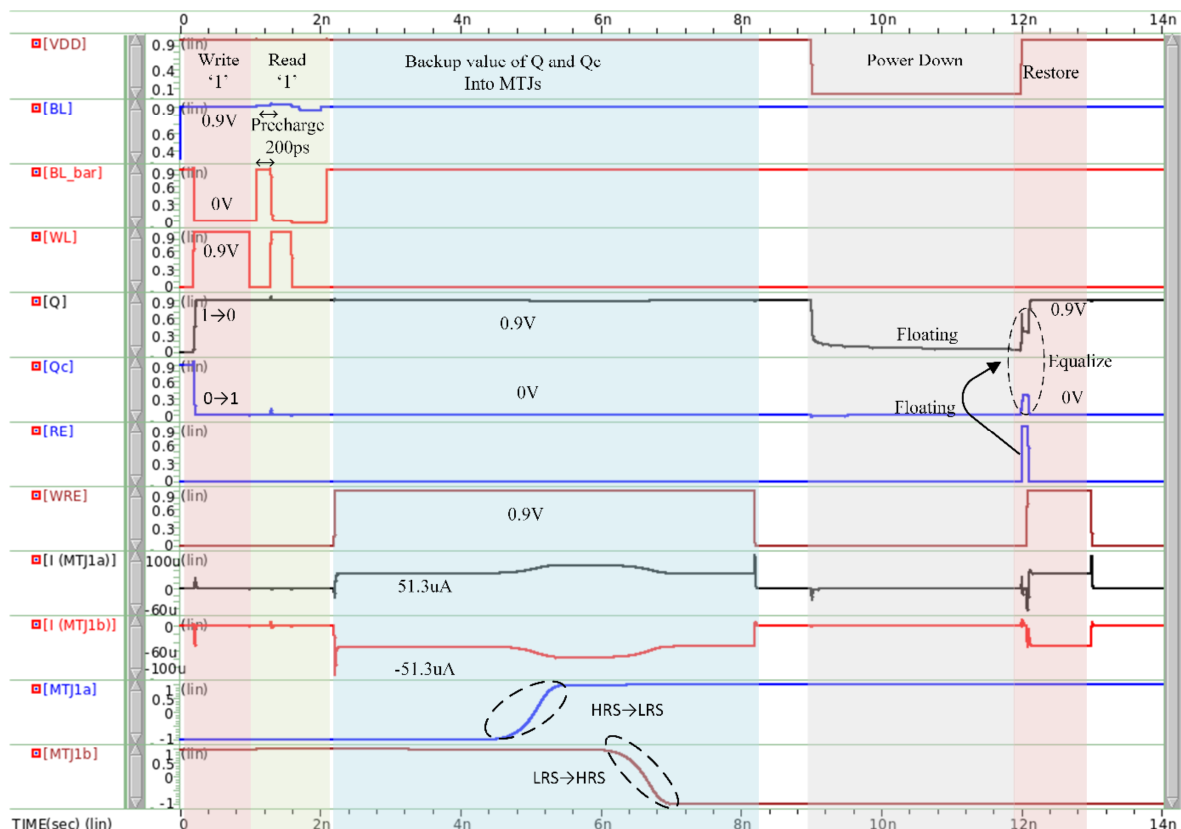

Figure 3.15: Simulation waveform of non-volatile hybrid 8T SRAM cell

In Figure 3.15, the first two highlighted regions plot the write & read operations of the hybrid cell, which is performed similar to a conventional SRAM cell. Let us assume cell stores a logic ‘0’ where node q stores a logic ‘0’ and node qc stores a logic ‘1’ before starting the write operation. For writing logic ‘1’ into the cell, BL is charged to vdd while BL_bar is pulled to gnd. Now asserting WL high causes node q to transition from ‘0’ to ‘1’ and alternately node qc to transition from ‘1’ to ‘0’. And then read operation is performed in which both BL and BL_bar are pre-charged. After pre-charging the bitlines, WL is asserted high, causing BL_bar to discharge through pass transistor and pull-down transistor. The region in blue color plots the backup operation where logic ‘1’ stored in internal node q and logic ‘0’ in node qc is to be written into the MTJ device. Before the backup operation, MTJ1a is in HRS, and MTJ1b is in LRS, which corresponds to logic '0’. Now for storing logic ‘1’ data stored in MTJ device has to be overwritten. The backup operation begins with asserting control signal WRE to logic high. With WRE high, it is observed in Figure 3.15 that the current I (MTJ1a) and I (MTJ1b) start flowing from q → MTJ1a → MTJ1b → qc switching MTJ1a from HRS → LRS and MTJ1b from LRS → HRS. After regaining power, restore operation is initiated by equalizing the internal nodes of the cell through activation of RE control signal. Before control signal RE is asserted low, control signal WRE is asserted high. With WRE turned ON, different value of resistance in each branch (low resistance of MTJ1a and high resistance of MTJ1b) pull node q to logic ‘1’ and node qc to logic ‘0’. Table 3.4 quantifies the energy consumptions during various operation performed in hybrid 8T SRAM cell.

TABLE 3.4: ENERGY CONSUMPTION AND ACCESS LATENCY OF THE PROPOSED HYBRID 8T SRAM CELL DURING VARIOUS OPERATION

	Energy Consumption	**Latency**
Write operation	15 fJ	22ps
Read operation	0.9 fJ	30ps
Backup operation	288.6 fJ	6ns
Restore operation	30 fJ	140ps

Figure 3.16 compares the energy consumption and latency of backup, restore, write and read operation performed in hybrid 8T SRAM cell using different technology nodes. It is observed that the energy consumption and latency during the backup operation is high compared to other operations at every technology node due to large switching energy and switching time of magnetic tunnel junction device. Whereas, energy consumption and latency of memory

read and write operations is similar to conventional memory cell because of the isolation of magnetic tunnel junction device from the cell during memory operation by disabling the isolation transistor (XE1 shown in Figure 3.12).

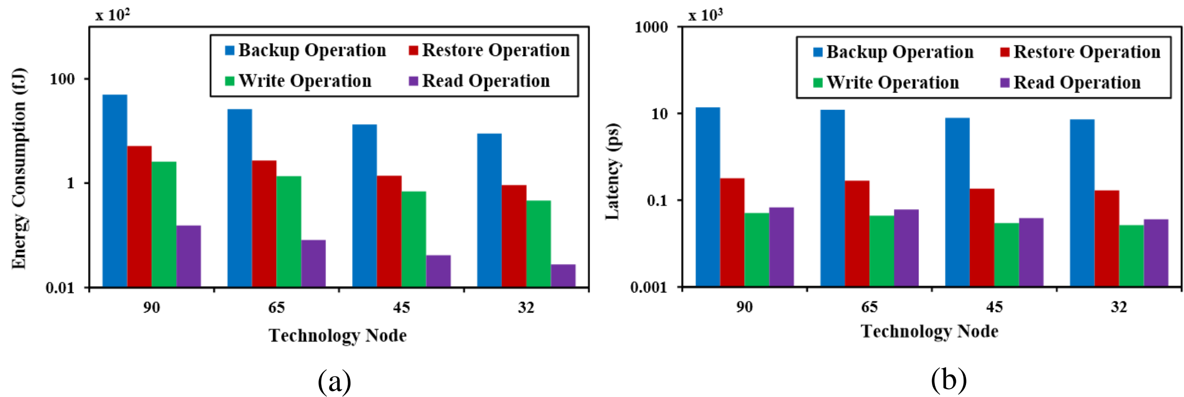

Figure 3.16: Energy consumption and latency of various operation of hybrid 8T SRAM cell at different technology nodes

3.3.3 Hybrid 11T SRAM Cell

3.3.3.1 Structure of hybrid 11T SRAM cell

The schematic of the proposed hybrid 11T SRAM cell is shown in Figure 3.17. The proposed cell uses only one MTJ device (MTJ1) and two additional NMOS transistors (N7 and N8) to incorporate non-volatility in transmission gate-based feedback cutting 9T SRAM cell (TFC-9T) [41].

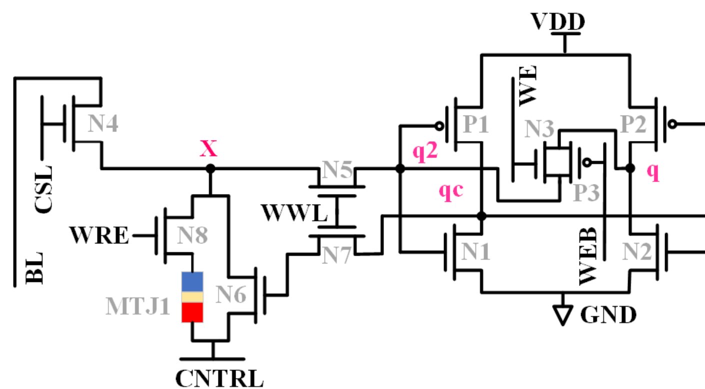

Figure 3.17: Proposed hybrid 11T SRAM Cell.

The MTJ1 is connected to the bitline BL through the transistor N4 and N8. A logic '0' is represented by the low resistance state of the MTJ1 device. In contrast, a logic '1' is represented by the high resistance state of the MTJ1 device. The control signal WRE connected to the gate of N8 transistor is used to provide access to the MTJ1. During the normal

memory operations (read/write), WRE is disabled to isolate the MTJ1 from the main SRAM cell, thereby eliminating the formation of a direct path between node X and CNTRL. As a result, significant power is saved, which reduces the impact of the MTJ device addition to the TFC-9T SRAM cell. The transmission gate (N3-P3) controlled by WE and WEB is connected in between the feedback path of two inverters.

3.3.3.2 Operational modes of hybrid 11T SRAM cell

The proposed hybrid 11T SRAM cell has four operational modes: Read, Write, Backup, and Restore. The memory read and write operations are performed similar to conventional TFC-9T SRAM cell. However, two additional operations are performed before power gating the memory array. Before turning OFF the memory block, the content of the SRAM cell is stored in the MTJ device using the backup operation. After turning ON the block, the value stored in the MTJ device is transferred back to the SRAM cell using the restore operation. The detail description of all four operational modes are provided in the following sub-sections.

3.3.3.2.1 Write operation

The write operation in the proposed hybrid 11T SRAM cell is performed using a single bitline (BL). The status of various control signals during the write operation is tabulated in Table 3.5. The control signal CSL and WWL are enabled to connect bitline BL directly to the internal nodes of cell. Since feedback is cut-off during the write operation, the value of bitline is first transferred to node qc, followed by node q. To write logic '1' into the SRAM cell, BL is set to gnd. In contrast, to write logic '0', BL is set to vdd.

TABLE 3.5: STATUS OF CONTROL SIGNAL DURING VARIOUS OPERATIONAL MODES OF PROPOSED HYBRID 11T SRAM CELL

	CSL	WWL	BL	CNTRL	WE	WRE
Read	vdd	vdd	Pre-Charged	gnd	vdd	gnd
Write	vdd	vdd	gnd/vdd	gnd	gnd	gnd
Backup	vdd	vdd	gnd/vdd	vdd/gnd	gnd	vdd
Restore	gnd	vdd	-	vdd	gnd	vdd

3.3.3.2.2 Read operation

The read operation in the proposed hybrid 11T SRAM is also performed through a single bitline BL. The bitline is pre-charged to vdd. In order to read the value stored in the internal

nodes (q and qc) of the SRAM cell, column shared CSL is asserted high while CNTRL is asserted low as tabulated in Table 3.5. If the cell stores logic '0'(q=0/qc=1), transistor N6 is turned ON, forming a discharge path for the BL through the transistors N4 and N6. Conversely, if the cell stores logic '1' (q=1/qc=0), the transistor N6 is turned OFF, and BL remains pre-charged.

3.3.3.2.3 Backup operation

The backup operation of the proposed hybrid 11T SRAM cell simultaneously writes the data value into the SRAM cell nodes (q and qc) and the MTJ device. The value of control signals required to perform the backup operation in the proposed SRAM cell is tabulated in Table 3.5. The control signal CSL and WWL are enabled to connect bitline BL to the nodes. Whereas the control signal WRE is enabled to connect the bitline BL to the MTJ device. The bi-directional current required for MTJ device switching is provided through BL and CNTRL. To store logic '1' into the MTJ device, BL is set to vdd, and CNTRL is set to gnd. Conversely, to store logic '0' into the MTJ device, BL is set to gnd, and CNTRL is set to vdd.

3.3.3.2.4 Restore operation

The restore operation writes back the value stored in the MTJ device to the SRAM cell. During the restore operation, CNTRL is set to vdd, and BL is set to gnd. With control signals WRE and CSL enabled, current flow through the CNTRL-BL path whose magnitude depends upon MTJ device resistance value. The two different current values corresponding to two resistive states of the MTJ device result in different voltage drop across node X. This difference in voltage across node X is interpreted as logic '0' and logic '1' by the SRAM cell. The LRS of the MTJ device results in a lower voltage at node X, which writes logic '0' into the cell. On the contrary, the HRS of MTJ device results in higher voltage at node X, which writes logic '1' into the cell.

3.3.3.3 Circuit-level analysis

In this sub-section, we validate the functionality of the proposed hybrid 11T SRAM cell using the SPICE tool. Figure 3.18 plots the transient waveform of the proposed hybrid cell. The simulation starts with a backup operation where we store logic '1' into the MTJ device. Let us assume the cell initially stores logic '1' (q=1/qc=0) and the MTJ1 stores logic '0' (low resistance state), as depicted in Figure 3.18. To store logic '1' into the MTJ1 device, bitline BL is pulled to vdd, and CNTRL is set to gnd. The control signals WWL, WRE, and CSL are asserted high. As a result, the current start flowing through the MTJ1 device, which causes

MTJ1 device to change its state from low resistance state to high resistance state, representing logic ‘1’. In the next step, the power supply is gated to represent a power-down or power-failure period. We then perform a restore operation to retrieve the previously saved state. The CNTRL is set to vdd while bitline BL is set to gnd. The control signals WRE and WWL are enabled to allow current to flow from CNTRL to BL through the MTJ1 device. As the MTJ1 device is in high resistance state, the voltage drop at node X is large, which is translated to logic ‘1’ at the output node q by the SRAM cell. The output logic ‘1’ at node q demonstrates the successful restoration of data after the power-off period. Further, we perform basic memory operations: write and read. We first perform a memory write operation to overwrite the logic ‘1’ stored by the cell to logic ‘0’. Finally, we perform read operation where BL discharges resulting in logic ‘0’ at the output.

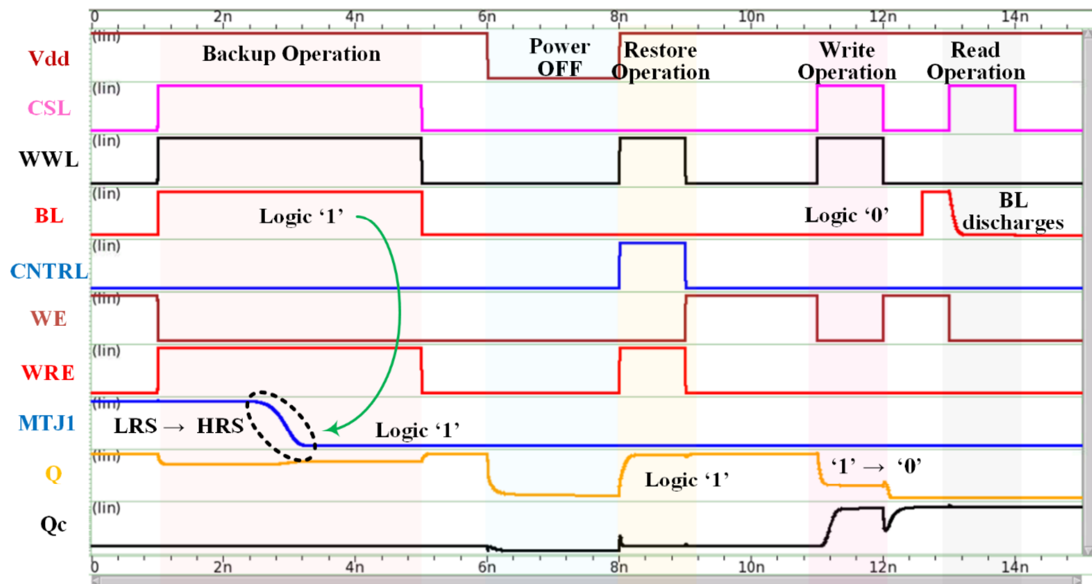

Figure 3.18: Simulation waveform of proposed hybrid 11T SRAM cell

Table 3.6 summarizes the energy consumption and access latency of the proposed hybrid 11T SRAM cell.

TABLE 3.6: ENERGY CONSUMPTION AND DELAY OF PROPOSED HYBRID 11T SRAM CELL

	Energy Consumption	**Latency**
Write operation	22.2 fJ	30.4ps
Read operation	2.03 fJ	35ps
Backup operation	456 fJ	4ns
Restore operation	100 fJ	133ps

During the backup operation, a large current is required for a long duration to switch the state of the MTJ device, hence, the energy consumption and latency during the backup operation are highest. However, during the memory read and write operation, the MTJ device is isolated from the cell, thus have a little impact on the performance of these operations.

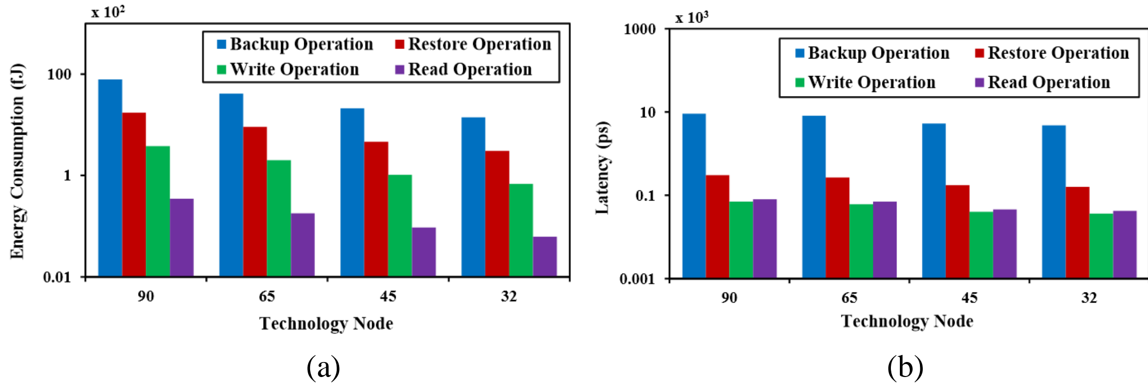

Figure 3.19: Energy consumption and latency of various operation of hybrid 11T SRAM cell at different technology nodes

Figure 3.19 compares the energy consumption and latency of backup, restore, write, and read operation performed in hybrid 11T SRAM cell using different technology nodes. It is observed that the energy consumption and latency during the backup operation are high compared to other operations at every technology node due to the large switching energy and switching time of the magnetic tunnel junction device. Whereas, energy consumption and latency of memory read and write operations is similar to conventional memory cell because of the isolation of magnetic tunnel junction device from the cell during memory operation.

3.4 Simulation Result and Comparative Analysis

Out of three proposed hybrid 7T, 8T, and 11T SRAM cells, the hybrid 8T SRAM cell is more promising due to its low store/restore power and simple control signaling. Therefore, further simulations are performed using the hybrid 8T SRAM cell. The proposed 8T SRAM cell is compared with the existing MTJ device-based hybrid SRAM cells. One such memory cell is a hybrid 4T SRAM cell based on a loadless 4T SRAM cell [30]. It does not have connection to a power supply (Vdd). Therefore, data in hold state discharges very quickly, resulting in the lowest SNM value among all other non-volatile hybrid SRAM cells. One way to retain the value in the hybrid 4T SRAM cell is to perform multiple write operations. Another non-volatile memory cell is the hybrid 6T SRAM cell, which is based on the conventional 6T SRAM cell with two MTJ devices having a direct bit to bit connection with the internal storage nodes of the cell [31]. The main disadvantage of hybrid 6T SRAM cell is the formation of a

direct path between the nodes and gnd during the idle period resulting in high leakage powers. To overcome this problem, the hybrid 8T SRAM cell [34] uses two extra transistors that eliminate the direct path between nodes and gnd. The two extra transistors, also known as isolation transistors, isolates the cell internal storage nodes during idle mode to significantly reduce the leakage power. Similarly, the proposed hybrid 8T SRAM cell isolates the nodes from the MTJ deice during the idle period, but it uses only one transistor. Table 3.7 summarizes the simulation results obtained for the various design of MTJ device-based hybrid SRAM cells. In order to have fair comparative analysis, transistor sizing of conventional 6T cell structure used in the designs of 6T, Hybrid 6T, Hybrid 8T, and proposed hybrid 8T SRAM cells are kept similar. The design criterion used for sizing 6T cell is to obtain similar read and write margins. The sizes of extra two isolation transistors used in the design of the hybrid 8T cell are kept equal. Whereas the sizes of extra two transistors used in the design of proposed 8T SRAM cell are different. Isolation transistor XE1 is made approximately 2x wider than equalization transistor XE2 to allow large current flow during backup operation.

3.4.1 Energy estimation

Table 3.7 summarizes the energy consumption by various hybrid SRAM cells during backup operation. Compared with the above-mentioned MTJ based hybrid SRAM cell, the proposed 8T SRAM cell shows a significant reduction in backup energy. The hybrid 4T SRAM cell [30] has high backup energy due to additional writing circuitry for the MTJ device with a power supply requirement greater than vdd. A separate writing interface improves switching time at the cost of increased area overhead and large energy consumption. The other hybrid SRAM cells, 6T SRAM cell [31] and 8T SRAM cell [34], have a similar backup process in which two distinct operations are performed to complete the backup operation. The first operation is a) Reset operation in which both the MTJ devices are reset to a high resistance state by raising the control voltage higher than vdd, and second operation is b) Store operation in which current is supplied by the cell internal storage nodes (q and qc) to store the corresponding value into the MTJ device. Due to these two operations, energy consumption during backup operation in hybrid 6T cell and hybrid 8T cell increases. Also, the control signal used in hybrid 6T [31] and hybrid 8T [34] cells have complex timing and amplitude requirements such that their generation results in increased system complexity. In contrast, the proposed hybrid 8T SRAM cell significantly reduces the backup energy by eliminating the need of a separate writing interface, dual-phase backup, and multiple control signals. It is observed that backup energy in the proposed hybrid 8T SRAM cell reduces by 51.9%, 78%,

and 65% compared to 4T, 6T, and 8T hybrid SRAM cells. Also, all the above-mentioned hybrid SRAM have similar restore operation during which a higher supply voltage is applied either through bitlines (in case of hybrid 4T SRAM cell [15]) or control signal (in case of hybrid 6T SRAM cell [31], and hybrid 8T SRAM cell [34]) which forces current to flow from MTJ device to nodes. Consequently, voltage drop across the cell nodes q and qc depends entirely on MTJ device resistances which is then translated to full swing by cell feedback action. The aforementioned hybrid SRAM cells are for 'normally-off and instantly-on' application. So, the node voltages are expected to be discharged completely. Any residual voltage can result in a false restore operation. However, our proposed 8T SRAM cell overcomes this dependency of restore operation on node voltage by introducing an equalization transistor.

TABLE 3.7: COMPARISON RESULTS OF DIFFERENT MTJ DEVICE BASED HYBRID SRAM CELL

	6T SRAM Cell	**Hybrid 4T SRAM cell [15]**	**Hybrid 6T SRAM Cell [31]**	**Hybrid 8T SRAM Cell [34]**	**Hybrid 8T SRAM Cell (proposed work)**
Write Energies (fJ)	14	26	15	16	15
Read energies (fJ)	0.5	18	9.3	0.98	0.9
Backup Energies (fJ)	-	600	1331	769	288.6
Restore Energies (fJ)	-	52	100	38	30
Latency (ps)	23	348	24.5	25.5	24
Leakage Power	0.9nW	250nW	92uW	0.9nW	0.9nW
SNM (mV)	407	98	325	407	407

3.4.2 Area estimation

The integration of the MTJ device with CMOS technology in the hybrid SRAM cells does not increase the die area, as it is stacked over the chip surface. However, transistor sizing in hybrid cells is kept 2-3 times bigger than conventional SRAM cell to allow a large switching current to flow. The hybrid 4T SRAM cell has a separate current generator circuit that generates the spin-polarized current required to switch the state of the MTJ device during the backup operation. Due to a separate writing interface, the area overhead of the hybrid 4T SRAM cell is very large. The hybrid 8T SRAM cell has an area of 0.355um^{2}, calculated from

the layout, as shown in Figure 3.20, results in an area overhead of 28% compared to hybrid 6T SRAM cell due to the two extra isolation transistors used to reduce the leakages. The total area of the proposed hybrid 8T SRAM cell is $0.405um^2$, also calculated from the layout, as shown in Figure 3.20. The larger XE1 and XE2 transistors in the proposed 8T SRAM cell lead to an area overhead of 12% and 37% compared to hybrid 8T and 6T SRAM cells, respectively. Although the area overhead of the proposed hybrid SRAM cell is slightly greater than hybrid 8T and 6T SRAM cells, its energy consumption during backup and restore operation is significantly lower when compared to hybrid 8T and 6T SRAM cells. Therefore, in energy-limited systems such as IoT, the proposed 8T SRAM cell is more efficient in meeting the demand for ultra-low-power operation.

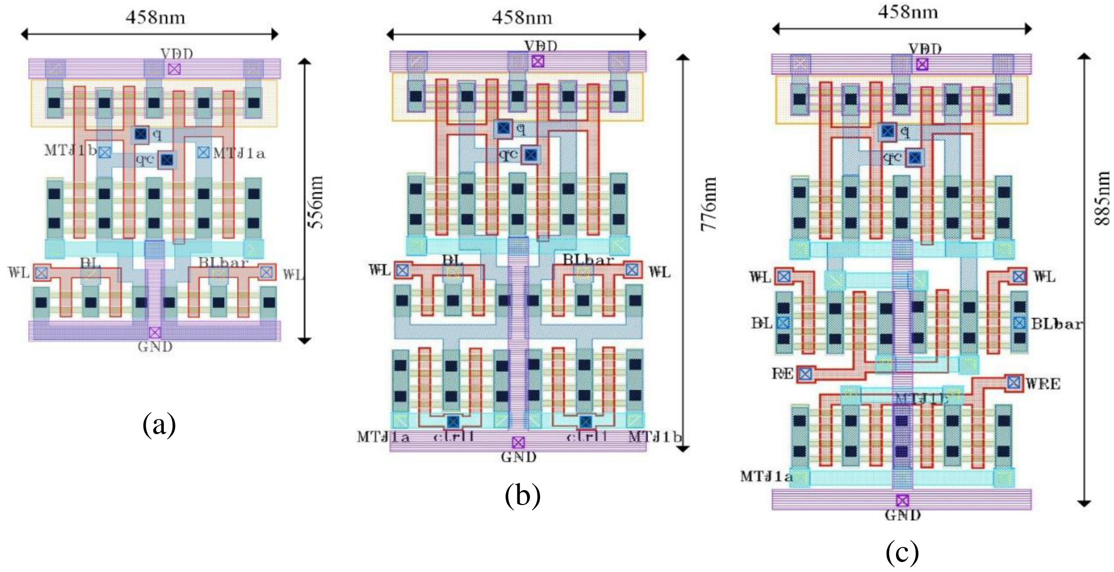

Figure 3.20: Layout design of (a) Hybrid 6T SRAM cell (b) Hybrid 8T SRAM cell (c) Proposed hybrid 8T SRAM cell

3.5 Conclusion

In this work, we address the challenges of data loss incurred during direct implementation of normally-off computing technique and sudden power outage in the batteryless system. We present three hybrid SRAM cells that can store the state within the locally embedded non-volatile magnetic tunnel junction (MTJ) device using a simple backup operation. The proposed hybrid 7T and 8T SRAM cells utilize an MTJ device pair, whereas the hybrid 11T SRAM cell uses a single MTJ device to store the content of SRAM during the power OFF state. The addition of the MTJ device to the conventional SRAM cells has little impact on their read and write performance. The different operational modes of all these proposed cells are explained in detail. We also perform exhaustive circuit analysis to validate the design and

evaluate multiple key performance parameters, including read/write energies, backup/restore energies, access times. Furthermore, the proposed hybrid 8T SRAM cell shows the best all-around performance when compared with the existing MTJ device-based hybrid SRAM cell. It is worth noting that the proposed 8T SRAM cell shows 51-78% reduction in backup energy due to elimination of a separate MTJ writing interface and extra control signals. Also, a reliable restore operation is ensured by the introduction of an extra equalization transistor. Thus, the low-power backup operation and reliable restore operation of the proposed design effectively tackle the problem of erratic power supply in energy-autonomous IoT systems.

REFERENCES

[1] J. Chang et al., "Embedded memories for mobile, IoT, automotive and high-performance computing," *2017 Symposium on VLSI Technology,* Kyoto, Japan, 2017, pp. T26-T27, doi: 10.23919/VLSIT.2017.7998187.

[2] S. H. Zadeh, T. Ytterdal and S. Aunet, "An ultra low voltage subthreshold standard cell based memories for IoT applications," *2020 28th Iranian Conference on Electrical Engineering (ICEE),* Tabriz, Iran, 2020, pp. 1-5, doi: 10.1109/ICEE50131.2020.9260950.

[3] C. Zhu, V. C. M. Leung, L. Shu, and E. C. H. Ngai, “Green Internet of Things for Smart World,” IEEE Access, vol. 3, pp. 2151–2162, 2015.

[4] S. Sudevalayam and P. Kulkarni, “Energy harvesting sensor nodes: Survey and implications,” IEEE Commun. Surv. Tutorials, vol. 13, no. 3, pp. 443–461, 2011.

[5] V. Raghunathan and P. H. Chou, “Design and Power Management of Energy Harvesting Embedded Systems,*” Proc. 2006 Int. Symp. Low Power Electron. Des., ISLPED’06,* pp. 369–374, 2006.

[6] Y. Liu et al., “Ambient energy harvesting nonvolatile processors: From circuit to system,” *52nd ACM/EDAC/IEEE Design Automation Conference*, San Francisco, CA, USA, June 8-12, 2015.

[7] K. Ma et al., “Architecture exploration for ambient energy harvesting nonvolatile processors,” *2015 IEEE 21st International Symposium on High Performance Computer Architecture, HPCA 2015*, Burlingame, CA, USA, Feb 7-11, pp. 526–537.

[8] R. Torah, P. Glynne-Jones, M. Tudor, T. O’Donnell, S. Roy, and S. Beeby, “Self-powered autonomous wireless sensor node using vibration energy harvesting,” Meas. Sci. Technol., vol. 19, no. 12, p. 125202, 2008.

[9] C. Moser, L. Thiele, D. Brunelli, and L. Benini, “Adaptive power management in energy harvesting systems,” *Design, automation and test in Europe Conference and Exhibition (DATE ’07),* Nice, France, April 16-20, 2007, pp. 1–6.

[10] J. A. Paradiso and M. Feldmeier, “A Compact, Wireless, Self-Powered Pushbutton Controller,” *Proceedings of the 3rd International Conference on Ubiquitous Computing*, Atlanta, Georgia, USA, pp. 299–304, 2001.

[11] M. Gorlatova, J. Sarik, G. Grebla, M. Cong, I. Kymissis, and G. Zussman, "Movers and Shakers: Kinetic Energy Harvesting for the Internet of Things," IEEE J. Sel. Areas Commun., vol. 33, no. 8, pp. 1624–1639, 2015.

[12] M. Raju and M. Grazier, "Energy harvesting--ULP meets energy harvesting: a game-changing combination for design engineers, White Paper," Texas Instruments. p. 8, 2010.

[13] P. Kamalinejad, C. Mahapatra, Z. Sheng, S. Mirabbasi, V. C. M. Leung, and Y. L. Guan, "Wireless energy harvesting for the Internet of Things," IEEE Commun. Mag., vol. 53, no. 6, pp. 102–108, 2015.

[14] H. Jayakumar, K. Lee, W. S. Lee, A. Raha, Y. Kim, and V. Raghunathan, "Powering the internet of things," *Proc. 2014 Int. Symp. Low power Electron. Des. - ISLPED '14*, pp. 375–380, 2014.

[15] N. Onizawa, A. Mochizuki, A. Tamakoshi, and T. Hanyu, "Sudden Power-Outage Resilient In-Processor Checkpointing for Energy-Harvesting Nonvolatile Processors," IEEE Trans. Emerg. Top. Comput., vol. 5, no. 2, pp. 151–163, 2017.

[16] M. Xie et al., "Checkpoint Aware Hybrid Cache Architecture for NV Processor in Energy Harvesting Powered Systems," *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, 2016.

[17] T. Miwa et al., "NV-SRAM: A nonvolatile SRAM with backup ferroelectric capacitors," IEEE J. Solid-State Circuits, vol. 36, no. 3, pp. 522–527, 2001.

[18] L. Baldi, R. Bez, and G. Sandhu, "Emerging memories," Solid. State. Electron., vol. 102, pp. 2–11, 2014.

[19] K. Ma et al., "Nonvolatile Processor Architecture Exploration for Energy- Harvesting Applications," IEEE Micro, vol. 35, no. 5, pp. 32–40, Sep. 2015.

[20] R. Aitken, V. Chandra, J. Myers, B. Sandhu, L. Shifren, and G. Yeric, "Device and technology implications of the Internet of Things," *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, pp. 1–4, 2014.

[21] F. Su, K. Ma, X. Li, T. Wu, Y. Liu, and V. Narayanan, "Nonvolatile Processors : Why is it Trending ?," *Design, Automation & Test in Europe Conference & Exhibition (DATE),* vol. 2, pp. 966–971, 2017.

[22] K. Ma et al., "Nonvolatile Processor Architectures: Efficient, Reliable Progress with Unstable Power," IEEE Micro, vol. 36, no. 3, pp. 72–83, 2016.

[23] I. Arsovski, Q. Li, M. Kuemerle, R. Tu, and H. Pilo, "Memory design considerations for high-performance networking SoCs," *ISOCC 2013 - 2013 International SoC Design Conference*, pp. 286–289, 2013.

[24] B. Jovanović, R. M. Brum, and L. Torres, "Evaluation of hybrid MRAM/CMOS cells for 'normally-off and instant-on' computing," Analog Integr. Circuits Signal Process., vol. 81, no. 3, pp. 607–621, 2014.

[25] S. Senni, L. Torres, G. Sassatelli, and A. Gamatie, "Non-Volatile Processor Based on MRAM for Ultra-Low-Power IoT Devices," ACM J. Emerg. Technol. Comput. Syst., vol. 13, no. 2, pp. 1–23, 2016.

[26] H. Cai, Y. Wang, L. A. De Barros Naviner, J. Yang, and W. Zhao, “Exploring Hybrid STT-MTJ/CMOS Energy Solution in Near-/Sub-Threshold Regime for IoT Applications,” IEEE Trans. Magn., vol. 54, no. 2, pp. 1–9, 2018.

[27] A. Chen, “A review of emerging non-volatile memory (NVM) technologies and applications,” Solid. State. Electron., vol. 125, pp. 25–38, 2016.

[28] L. Cargnini, L. Torres, R. Brum, S. Senni, and G. Sassatelli, “Embedded Memory Hierarchy Exploration Based on Magnetic Random Access Memory,” J. Low Power Electron. Appl., vol. 4, no. 3, pp. 214–230, 2014.

[29] K. Chen, J. Han, and F. Lombard, “On the Restore Operation in MTJ-Based Nonvolatile SRAM Cells,” IEEE Trans. Very Large Scale Integr. Syst., vol. 23, no. 11, pp. 1–5, 2014.

[30] B. Jovanović, R. M. Brum, and L. Torres, “MTJ-Based Hybrid Storage Cells for ‘NORMALLY-OFF and INSTANT-ON’ Computing,” Electron. Energ., vol. Vol. 28, no. 3, pp. 465–476, 2015.

[31] T. Ohsawa et al., “A 1.5nsec/2.1nsec random read/write cycle 1Mb STT-RAM using 6T2MTJ cell with background write for nonvolatile e-memories,” VLSI Circuits (VLSIC), 2013 Symp., vol. 46, no. 1, pp. C110–C111, 2013.

[32] K. Chen, J. Han, and F. Lombardi, “On the nonvolatile performance of flip-flop/SRAM cells with a single MTJ,” IEEE Trans. Very Large Scale Integr. Syst., vol. 23, no. 6, pp. 1160–1164, 2015.

[33] S. Yamamoto and S. Sugahara, “Nonvolatile static random-access memory using magnetic tunnel junctions with current-induced magnetization switching architecture,” Jpn. J. Appl. Phys., vol. 48, no. 4, 2009.

[34] Y. Gong, N. Gong, L. Hou, and J. Wang, “MTJ based data restoration in non-volatile SRAM,” *2016 13th IEEE International Conference on Solid-State and Integrated Circuit Technology, ICSICT 2016 - Proceedings*, vol. 1, pp. 1011–1013, 2017.

[35] G. Panagopoulos, C. Augustine, and K. Roy, “A framework for simulating hybrid MTJ/CMOS circuits: Atoms to system approach,” *Des. Autom. Test Eur. Conf. Exhib. (DATE), 2012*, pp. 1443–1446, 2011.

[36] Y. Wang et al., “Compact Model of Dielectric Breakdown in Spin-Transfer Torque Magnetic Tunnel Junction,” IEEE Trans. Electron Devices, vol. 63, no. 4, pp. 1762–1767, 2016.

[37] Y. Zhang et al., “Compact model of subvolume MTJ and its design application at nanoscale technology nodes,” IEEE Trans. Electron Devices, vol. 62, no. 6, pp. 2048–2055, 2015.

[38] S. Pal, S. Bose, W. Ki and A. Islam, "Characterization of Half- Select Free Write Assist 9T SRAM Cell," in IEEE Transactions on Electron Devices, vol. 66, no. 11, pp. 4745- 4752, Nov. 2019.