

Low-Power Near-Threshold 10T SRAM Bit Cells With Enhanced Data-Independent Read Port Leakage for Array Augmentation in 32-nm CMOS

Shourya Gupta^{ID}, *Student Member, IEEE*, Kirti Gupta, *Senior Member, IEEE*,
Benton H. Calhoun, *Senior Member, IEEE*, and Neeta Pandey^{ID}, *Senior Member, IEEE*

Abstract—The conventional six-transistor static random access memory (SRAM) cell allows high density and fast differential sensing but suffers from half-select and read-disturb issues. Although the conventional eight-transistor SRAM cell solves the read-disturb issue, it still suffers from low array efficiency due to deterioration of read bit-line (RBL) swing and I_{ON}/I_{OFF} ratio with increase in the number of cells per column. Previous approaches to solve these issues have been afflicted by low performance, data-dependent leakage, large area, and high energy per access. Therefore, in this paper, we present three iterations of SRAM bit cells with nMOS-only based read ports aimed to greatly reduce data-dependent read port leakage to enable 1k cells/RBL, improve read performance, and reduce area and power over conventional and 10T cell-based works. We compare the proposed work with other works by recording metrics from the simulation of a 128-kb SRAM constructed with divided-wordline-decoding architecture and a 32-bit word size. Apart from large improvements observed over conventional cells, up to 100-mV improvement in read-access performance, up to 19.8% saving in energy per access, and up to 19.5% saving in the area are also observed over other 10T cells, thereby enlarging the design and application gamut for memory designers in low-power sensors and battery-enabled devices.

Index Terms—10T SRAM, assist techniques, E_{min} , low power, near-threshold, sub-threshold.

I. INTRODUCTION

STATIC Random Access Memory (SRAM) occupies a significant portion of a system-on-a-chip (SoC) and has a notable contribution to the total power consumption and area of the SoC. Since area is an important factor when designing circuits, memory design engineers aim to place as many cells as possible per column to allow sharing of peripheral circuitry. The conventional 6T and 8T cells are greatly limited by their inability to work in longer columns. This is because they

suffer from data dependent leakage and degraded I_{ON}/I_{OFF} ratio and read bit-line swing as more cells are placed on a single column. Therefore, there is a need to design new circuits to address this issue. Previous approaches [1]–[3] have tried to solve this issue by improving the I_{ON}/I_{OFF} ratio to enable up to 1k cells per column. Although these approaches have been successful at this task, these still suffer from large area or varying data-dependent performance. Some also fail to account for the minimum energy point in SRAMs and therefore, consume a lot of energy per access at ultra-low voltages. This work describes three iterations of SRAM bit cells with nMOS-only based read ports aimed to greatly reduce data-dependent read port leakage to enable 1k cells per RBL, improve read performance, and reduce area and power over conventional 6T and 8T cells and other novel read-port based cells. With a unique topology in each of the three cells' read port, we obtain improved read access performance, low energy per access, and low area respectively, thereby enlarging the design and application gamut for memory designers in low power sensors and battery enabled devices.

SRAM's impact has become especially important due to the emergence of battery powered portable devices and low power sensor applications. Most SRAM design effort has been led to facilitate voltage scaling and improving yield. The conventionally implemented six transistor (6T) cell in SRAMs allows high density, bit-interleaving and fast differential sensing but suffers from half-select stability, read-disturb stability, and conflicting read and write sizing. Previous attempts to solve these issues have included the implementation of assist techniques, novel cell design, architectural improvements, or technological developments.

Half-select and read-disturb issues in SRAMs can be mitigated by optimization of word-line voltage level. This includes word-line under-drive assists using process corner tracking [4] or using replica access transistors [5]. Delayed word-line boost [6] to match the internal voltage of half-selected cells to that of the bit-line during a read operation helps to improve their stability but requires fine tuning to establish the sensitive tradeoff between read stability and write ability. Cell supply boost assist can also be used to improve half-select stability by increasing the drive strength of pull down nMOS.

Manuscript received April 27, 2018; revised August 16, 2018 and October 8, 2018; accepted October 16, 2018. This paper was recommended by Associate Editor Y. Pu. (Corresponding author: Shourya Gupta.)

S. Gupta and B. H. Calhoun are with the Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22911 USA (e-mail: shourya.gupta94@gmail.com).

K. Gupta is with the Bharati Vidyapeeth's College of Engineering, New Delhi 110063, India.

N. Pandey is with Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi 110042, India.

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

Digital Object Identifier 10.1109/TCSI.2018.2876785

Negative cell ground [7] implementation to improve read stability is the most effective assist but has high energy cost due to use of multiple *GND* rails [8].

Disturb issues can also be mitigated by partial precharge of bit-lines to decrease the strength of access transistors. Pilo *et al.* [9] make use of regulators to reduce the precharge voltage level of the bit-lines to around 70% of supply voltage to improve the read stability. Alternatively, the bit-lines can be precharged using an nMOS instead of a pMOS to obtain a single V_{TH} drop on the bit-lines [10]. A process variation tolerant selective precharge assist [11] has also been used to decrease bit-line voltage level using charge sharing to improve half-select disturb issues. However, such partial bit-line precharge techniques reduce read ability and become less effective at lower voltages due to reduced V_{DS} of the access transistors.

Multiple supply line assist can also be used to improve read and write half-select stability issues in SRAMs. In [12], a column-based dynamic supply technique was proposed. By implementing different supply voltages for read, write and standby modes, it relieved half-select stability issues and allowed bit-interleaving. However, this resulted in increase in dynamic power, design and routing effort and area due to generation of multiple supply voltages.

Although assist techniques can be beneficial in improving the performance and yield of SRAMs, they can often have a deteriorating complementary effect on write and read operations. They can also incur large area overhead, increase the energy per access, and have a limited and saturating effect on yield. Furthermore, since read and write stability is greatly dependent on temperature variations, an SRAM can either be write-limited at lower temperatures or read-limited at higher temperatures. Therefore, assists often require process and temperature tracking for effective yield improvement.

Apart from assist techniques, improvements on the architectural front have also been made to address half-select and read-disturb stability issues. These include cross-point selection of words using both row and column word-lines to improve half select stability [13], [14]. Shorter bit-lines can also be used to improve read stability. These work by reducing bit-line capacitance, thereby improving dynamic read margin. However, this comes at the expense of large area overhead due to greater number of cell banks. In another work [15], an array architecture with an area overhead of 12% was implemented in order to address the half-select disturb issue by decoupling the large bit-line capacitance from half-selected cells. Read-and-write-back scheme [3] has also been used to alleviate the write-disturb in half-select cells. It allows data retention by writing back the stored data after each read. However, such techniques increase the dynamic power consumption since every column is subjected to full voltage swings. Additionally, the sense amplifier cannot be shared amongst several columns and has to be integrated in each column, thereby incurring a large area overhead.

With the 6T SRAM cell being afflicted by various stability issues, the 8T SRAM cell has been proposed (shown in Fig. 1). It has a decoupled read path comprising of two nMOS transistors. Although it eliminates the read-disturb issue, it is

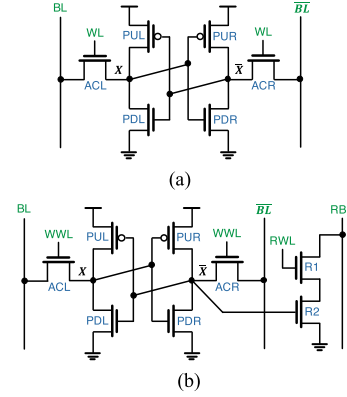


Fig. 1. Schematic of (a) 6T (b) 8T SRAM cell.

still afflicted by a pseudo-read during a write operation in half-selected cells on the same row. As such, the issue of loss of bit-interleaving capability arises. Bit-interleaving is essential to low voltage SRAM operation since it is combined with Error-Correction Code (ECC) to combat soft errors and achieve required yield targets. Soft errors, including Single Bit Upsets (SBUs) and Multiple Cell Upsets (MCUs) are caused by bombardment of alpha-particles, thermal neutrons or high energy cosmic rays [16]. The rate of soft errors increases by 18% for every 10% decrease in supply voltage [17]. This is especially problematic for low voltage SRAMs, since in sub-threshold operation region, the critical charge in nodes is significantly reduced, leading to frequent MCUs [18]. In [19] and [20], MCUs have been mitigated by implementing and combining bit-interleaving structure with ECC. In addition, bit-interleaving capable cell structures such as the column-decoupled 8T cell in [21], disturb-free 9T cell in [13], two-port disturb-free 9T cell in [22], multi-port 9T cell in [23] and the differential 10T cell in [24] have been proposed to enable bit-interleaving and remove half-select disturb issues by using both row and column word-lines. For cell structures without interleaving capability such as the single ended 8T cell, additional parity or ECC bits can be interleaved per word for soft error correction [25].

Even if the read and write disturb issues are alleviated using the methods described above, an array implemented using the 8T cells has low array efficiency [26]. This is because, its single ended mechanism requires a hierarchical sensing architecture which implements as few as eight cells per local RBL and multiple local RBLs per global RBL. Additionally, unlike the fast differential sensing in the 6T cell, the single ended sensing has a slow full swing operation. As greater number of cells are put on the same local RBL in order to improve array efficiency, both delay and the read bit-line voltage swing are greatly affected. Therefore, this form of hierarchical sensing does not compare to differential sensing in terms of both performance and array efficiency. Although many techniques have been proposed to improve the single ended read sensing performance [27], the area overhead still remains large. In order to improve the array efficiency and read bit-line voltage swing of single-ended-read cells, many modified read ports have been proposed [1]–[3], [28].

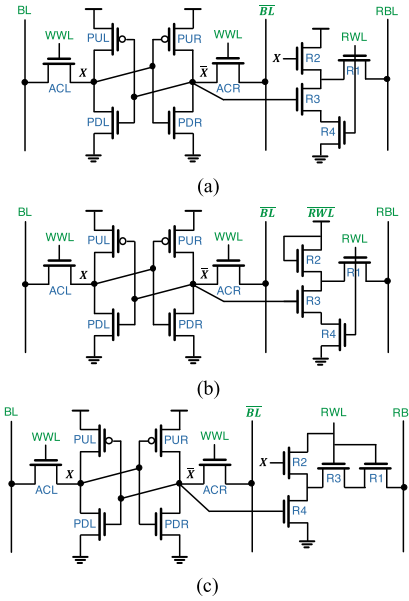


Fig. 2. Schematic of the proposed (a) 10T-P1 (b) 10T-P2 (c) 10T-P3 cells.

These designs aim to put up to 1k cells per bit-line by improving the I_{ON}/I_{OFF} ratio of SRAM read ports. This approach helps to greatly improve the array efficiency as peripheral circuitry can be shared amongst greater number of cells. Although these approaches have been successful at this task, these still suffer from large area, varying data-dependent performance and high energy consumption. In this work, we propose three iterations of SRAM bit cells with nMOS-only based read ports and compare them with conventional 6T and 8T cells and previous 10T cell-based works by measuring metrics from simulation of a 128kb array on the 32nm technology node. We compute minimum energy per access for all cells considering different activity factors for various levels of caches and calculate dynamic failure rate based on operating frequency and process variations.

The paper has been constructed as follows. Section II describes the proposed cells and their working principle and Section III includes the performance evaluation of various bit cells on the basis of read bit-line swing, energy per access, dynamic failure probability and area. Section IV summarizes and concludes the paper.

II. PROPOSED SRAM BIT CELLS

A. Topology of Proposed Bit Cells

The schematic of the proposed 10T SRAM cells is shown in Fig. 2. Each of them comprises of cross coupled inverters (PUL-PDL and PUR-PDR) and two access transistors (ACL and ACR). The read port of each cell consists of four nMOS (R1, R2, R3 and R4). The read port in Fig. 2(a) has improved data-dependent read bit-line leakage and is aimed at high performance. The read ports in Fig. 2(b) and (c) have complete data-independent read bit-line leakage and are aimed at very low power and high density respectively. The working of each port has been explained in the next section.

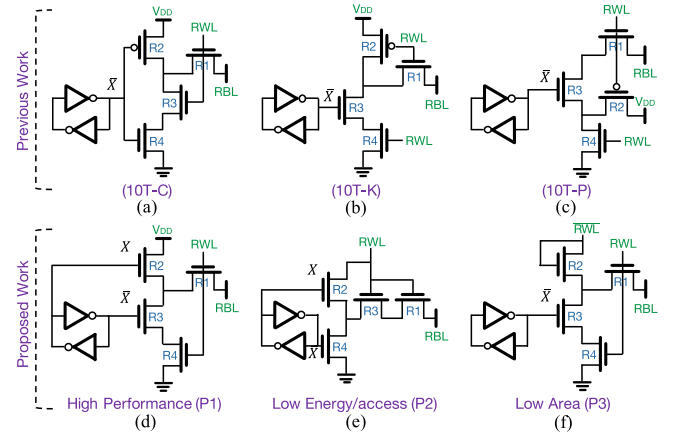


Fig. 3. Schematic of read port of (a) Calhoun and Chandrakasan [1] (b) Kim *et al.* [3] (c) Pasandi and Fakhraie [2] (d) Proposed 10T-P1 (e) 10T-P2 (f) 10T-P3 cell.

From here on, the proposed cells are referred to as 10T-P1, 10T-P2 and 10T-P3.

B. Bit Cell Working Mechanism

When operating in near and sub-threshold region, the I_{ON}/I_{OFF} is severely degraded and it becomes increasingly difficult to implement greater number of cells on a single column. As the number of cells increase, the combined pass-gate leakage becomes comparable to the read current, thereby making it difficult for the sense amplifier to correctly evaluate the read bit-line voltage level. Furthermore, the data stored in the cell also affects the read bit-line leakage, thereby making the off-state read bit-line leakage current to fluctuate highly. This is exacerbated at ultra-low voltages, where the worst-case data pattern can lead to the RBL voltage level of ‘zero’ becoming greater than the RBL voltage level of ‘one’ [29].

In order to improve the I_{ON}/I_{OFF} ratio, the read port shown in Fig. 3(a) was proposed in [1]. When the cell stores ‘one,’ the R2 pMOS charges the intermediate node, thereby greatly reducing the read bit-line leakage through R1 nMOS. However, this also leads to flow of leakage current from intermediate node into the RBL. The combined leakage of all cells on the same column can raise the low logic level of RBL to several hundred millivolts, thereby leading to reduced voltage swing and sensing margin. The conceptual scenario of the effective read bit-line voltage swing for this case has been depicted in Fig. 4(a). On the other hand, when the cell stores ‘zero,’ the RBL leakage is reduced through the stacking effect of nMOS. Therefore, such a topology makes the effective RBL swing largely dependent on the data pattern in the column. In another work [3], the data dependency was removed by creating a data-independent leakage path between the cell’s read port and the RBL. This led to a significant voltage swing on the RBL even at lower voltages. The read port and the corresponding effective RBL swing for the same has been shown in Fig. 3(b) and Fig. 4(b) respectively. A recent work [2], also proposed a modified read port [shown in Fig. 3(c)], to improve the I_{ON}/I_{OFF} ratio. However, it is also

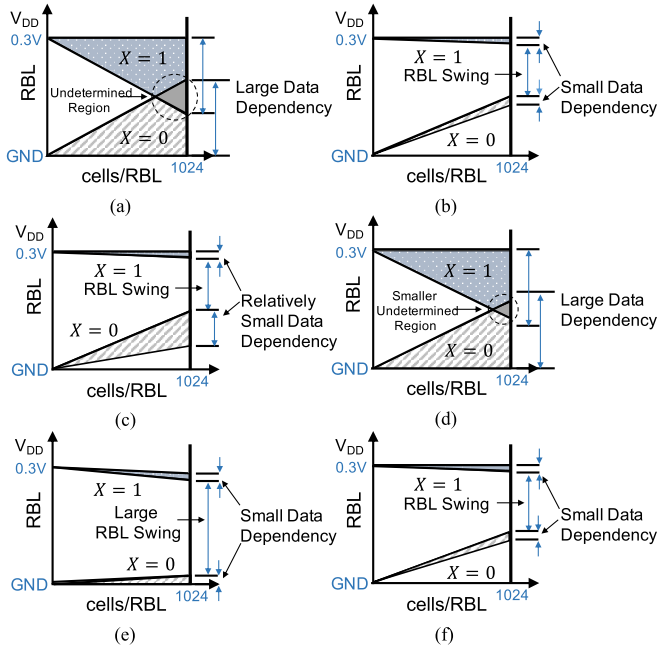


Fig. 4. Conceptual Effective Read Bit-Line Swing of (a) Calhoun and Chandrakasan [1] (b) Kim et al. [3] (c) Pasandi and Fakhraie [2] (d) Proposed 10T-P1 (e) 10T-P2 (f) 10T-P3 cell.

afflicted by the data-dependent leakage path issue. Depending upon the data stored in the cell, the leakage from intermediate node to RBL can change drastically, thereby leading to varying low logic voltage levels of RBL. Despite this issue, it is able to maintain an RBL swing, as shown in Fig. 4(c). From here on, the cells in Fig. 3(a), (b) and (c) will be referred to as the 10T-C, 10T-K and 10T-P cells respectively. Like the proposed cells, these cells also have the same topology for the write port and differ in terms of the read port only.

The schematic of the proposed read ports is shown in Fig. 3(d)–(f). The proposed 10T-P2 and 10T-P3 cells are aimed at low power and low area respectively while simultaneously maintaining a data-independent I_{ON}/I_{OFF} ratio. The principle behind their working is depicted in Fig. 5(c) and (d). As seen in Fig. 5, the magnitude of I_{leak} becomes equal in both read ‘zero’ and read ‘one’ case. This helps to maintain the required difference in magnitude between accessed-cell current in both cases. As such, a significant effective RBL swing can be observed, as shown in Fig. 4(e) and (f). This is not possible in the case of conventional 8T cell sensing, because of the large dependence of leakage current on the data pattern.

Although the proposed 10T-P1 cell decreases its data dependency in comparison to the 10T-C cell as seen in Fig. 4(d), it largely remains incapable of performing a read operation at ultra-low voltages. However, in the following subsection, we show that operating at ultra-low voltages increases the energy per access and operating near the threshold point is optimal for lowest energy consumption. As such, the 10T-P1 cell is operated near the sub-threshold region for lowest energy consumption and highest performance. At near-threshold and super-threshold voltages, the read bit-line swing is not an issue for the 10T-P1 cell. A more comprehensive analysis of RBL

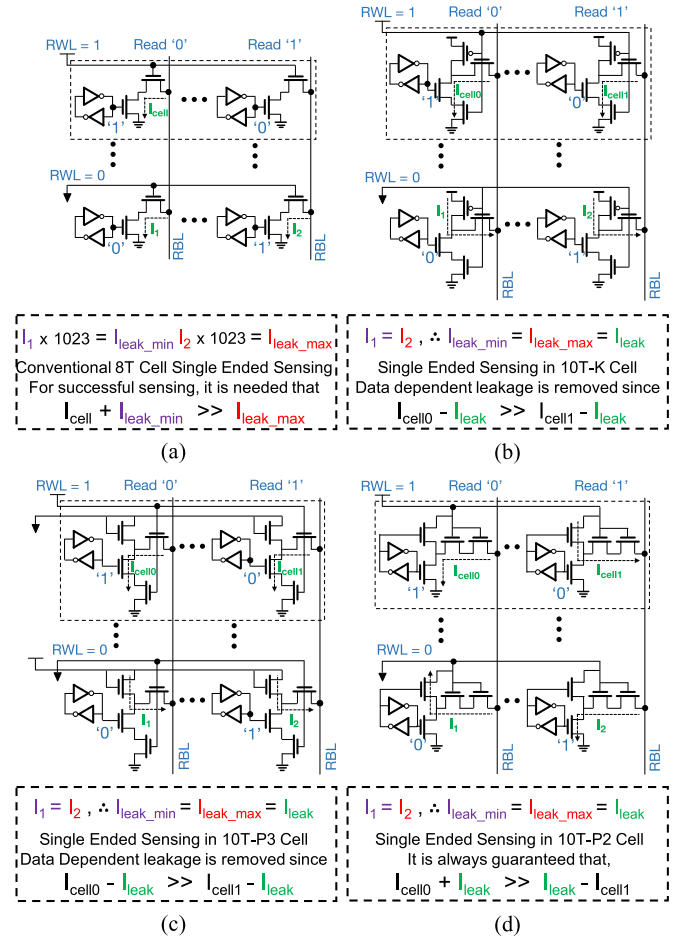


Fig. 5. RBL leakage scenario in (a) Conventional 8T (b) 10T-K (c) 10T-P3 (d) 10T-P2 cells.

swing of each cell with respect to data pattern, supply voltage and temperature is presented in the next section.

III. PERFORMANCE EVALUATION OF SRAM CELLS

A. Array Design

To measure the performance of proposed cells and compare them with previous works, we implemented a 128kb array using each cell. Since all the cells which have been compared are prone to write-disturb issue, the array was constructed in a non-interleaved architecture without column-select circuitry. The array comprises of four 32kb sub-blocks each with 1024 cells per column and a 32-bit word size. High speed Limited Switch Dynamic Logic (LSDL) [30] was used to construct the predecoders and decoders. The circuit implementation for these is shown in Fig. 6. Ten address bits were used as inputs to create sixty-four NOR based predecoders, whose outputs were then used as self-timed pulses to drive the decoder-driver for each row. Hierarchical-Word-Decoding (HWD) results in lower power consumption and faster access time in comparison to Divided-Wordline-Decoding (DWD) [31]. This is because the HWD architecture implements additional levels of word-lines to decrease the overall capacitance per row select path. However, the

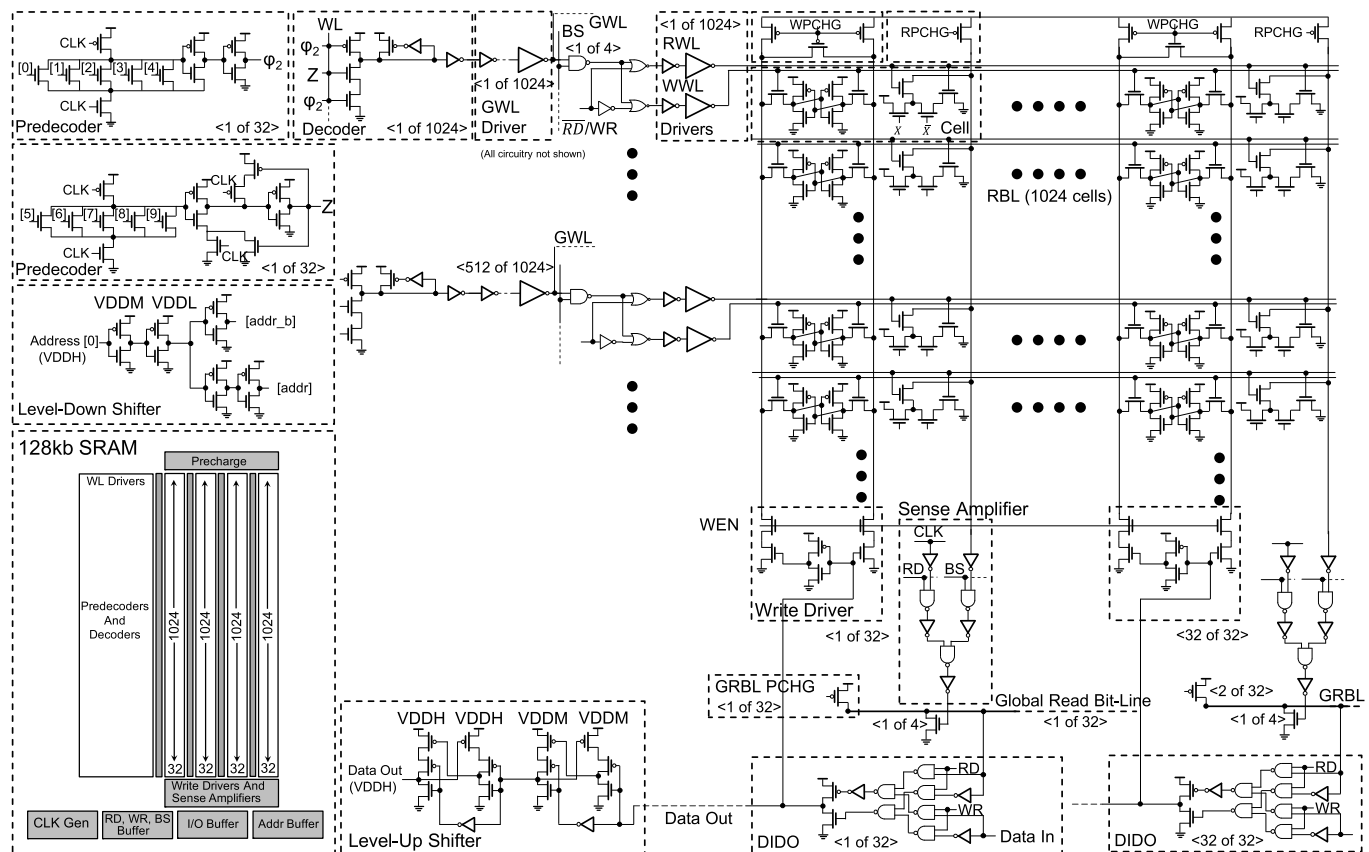


Fig. 6. The critical read path for the proposed 10T-P1 cell and schematic of Limited Switch Dynamic Logic based pre-decoders and decoders.

TABLE I
TRANSISTOR WIDTH SIZING (nm)

S. No.	Bit Cell Topology	PUL/PUR	PDL/PDR	ACL/ACR	R1/R4	R3	R2
1	6T Cell	72	144	72	-	-	-
2	8T Cell	72	72	72	108	-	-
3	10T-C Cell	72	72	72	108	108	72
4	10T-K Cell	72	72	72	108	108	72
5	10T-P Cell	72	72	72	108	108	72
6	10T-P1 Cell	72	72	72	108	180	72
7	10T-P2 Cell	72	72	72	108	108	72
8	10T-P3 Cell	72	72	72	108	108	72

Length of all transistors: 36nm

advantages are negligible for smaller arrays ($< 256\text{kb}$). Since the array capacity is only 128kb , the DWD scheme was used. Selective precharge using *BS* (Block-Select) was used to enable the precharge of bit-lines of only the accessed block to reduce active power consumption.

Four metal layers were used to route V_{DD} , GND , bit-lines and the local and global word-lines. The transistor sizing for all cells compared in this work is shown in Table I.

The proposed 10T-P1 cell, with its unique topology and layout, was able to increase the sizing of R3 nMOS, thereby leading to large improvements in read performance. While this change did not lead to an increase in the area of cell, it did result in slightly increased standby power.

While local RWLs and WWLs were implemented in each row for all cells, the RWL for the 10T-P3 cell was shared amongst two adjacent rows. Such an implementation was made due to the limited vertical pitch of each row and the wiring requirement of several horizontal local and global lines in each row. Whenever a read operation is performed, the RWL goes to low level for two rows. Although the leakage current from the half-selected row increases, the overall increase in I_{OFF} for each RBL is extremely small and does not impact read performance. Only the dynamic and standby power increase, which has been addressed in the following subsection.

The timing diagram for the memory operation is shown in Fig. 7. The WL (word-line), RD (read) and WR (write)

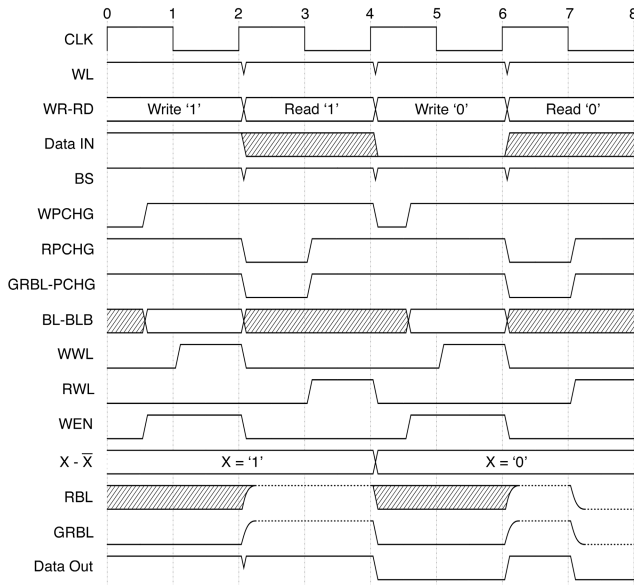


Fig. 7. Timing diagram for subsequent write-read operation in the SRAM.

signals are positive-edge synchronized with the clock signal. Both the local RWL and WWL are enabled with the negative edge of clock signal. The WL is combined with the output of each predecoder to enable new address decoding, only at the beginning of each operation. The bit-lines are precharged at the beginning of each write operation, after which the data is loaded onto them before the enabling of WWL. Similarly, all the local bit-lines of the accessed block and global RBLs are precharged during the first half of the read clock cycle. The RWL is enabled during the second half of the clock to allow the RBL to develop conditionally. The voltage level on the RBL is detected by the sense amplifier, which is then used to evaluate the global RBL. The DIDO (Data-In-Data-Out) produces logical output according to global RBL level.

B. Read Bit-Line Swing

When performing a read operation, a sufficient read margin is required for correct evaluation by the sense amplifier. This margin is ensured by a large RBL swing, which is ideally the difference between V_{DD} and GND . However, the combined leakage from all cells into the RBL or from RBL into the cells, can severely degrade the low and high logic levels respectively, thereby leading to diminished margin for correct sensing at lower voltages. The deterioration in RBL swing is also exacerbated at higher temperatures due to increase in leakage current. In some cases, the RBL swing is also affected by the data-dependent leakage path in the read ports of SRAM cells.

The effective RBL swing, as a percentage of V_{DD} , and with respect to varying voltage, temperature and data pattern is shown for all cells compared in this work in Fig. 8. The following three cases have been considered when measuring the RBL swing -

- 1) All cells in the column store 'zero.'

- 2) All cells in the column store 'one.'

- 3) 'One' and 'zero' are distributed equally in the column.

As seen in Fig. 8, the 10T-C cell's effective RBL swing is lowest amongst all cells compared and varies greatly according to the data pattern. The 10T-K cell has a data-independent leakage path in its read port, which leads to a data-independent RBL swing. The 10T-P cell also has a data-dependent RBL swing, albeit with a lower level of variation. The effective RBL swing in the proposed 10T-P1 cell has a much lower dependence on the data pattern in comparison to the 10T-C cell. Although it is data-dependent, it is aimed at high performance and near/super-threshold operation, where RBL swing is not an issue. Both the 10T-P2 and 10T-P3 cells have a data-independent RBL swing, with the 10T-P2 exhibiting the highest RBL swing amongst all the cells compared.

C. Standby Leakage Power

At any given voltage and temperature, the absolute power of the cell should be as low as possible. Therefore, a complete comparison of average standby leakage power per cell with respect to voltage, temperature and data pattern variations has been shown in Fig. 9. The following three cases have been considered when measuring standby power-

- 1) All cells in the array store 'zero.'
- 2) All cells in the array store 'one.'
- 3) 'One' and 'zero' are distributed equally.

It can be observed from Fig. 9, that the proposed cells consume lowest leakage power in comparison to previous work at most voltages and temperatures. It can also be observed that an equal distribution of 'zero' and 'one' data leads to the worst-case leakage power in case of 10T-C, 10T-P, 10T-P1 and 10T-P2 cells. The 10T-K and 10T-P3 mostly have the highest leakage power when they store 'zero.' However, this trend does not maintain across all supply voltages and temperatures. The 10T-P2 cell, which consumes lowest power in near and sub-threshold region, has a sudden increase in power consumption at higher voltages due to increase in flow of gate tunneling leakage current. At 27°C, the 10T-P3 cell goes from consuming lower power than the 10T-C cell in super-threshold region to higher power consumption in sub-threshold region. Although this trend holds true at lower temperatures (-10°C) as well, at high temperatures (80°C), the 10T-P3 cell always consumes higher power than the 10T-C cell.

The total power is also affected by the frequency of operation, switching activity, layout dependent interconnect parasitics and supply voltage. Since all cells compared in this work have significant differences in interconnect wiring due to varying aspect ratios, the total active power will also vary greatly. Therefore, in the next subsection we evaluate the total energy per access for all cells.

D. Total Energy per Access and E_{min}

In super-threshold region, the active power dominates over standby power. But as we approach near-threshold and sub-threshold region, active power is reduced greatly due to its quadratic dependence on current. This makes the

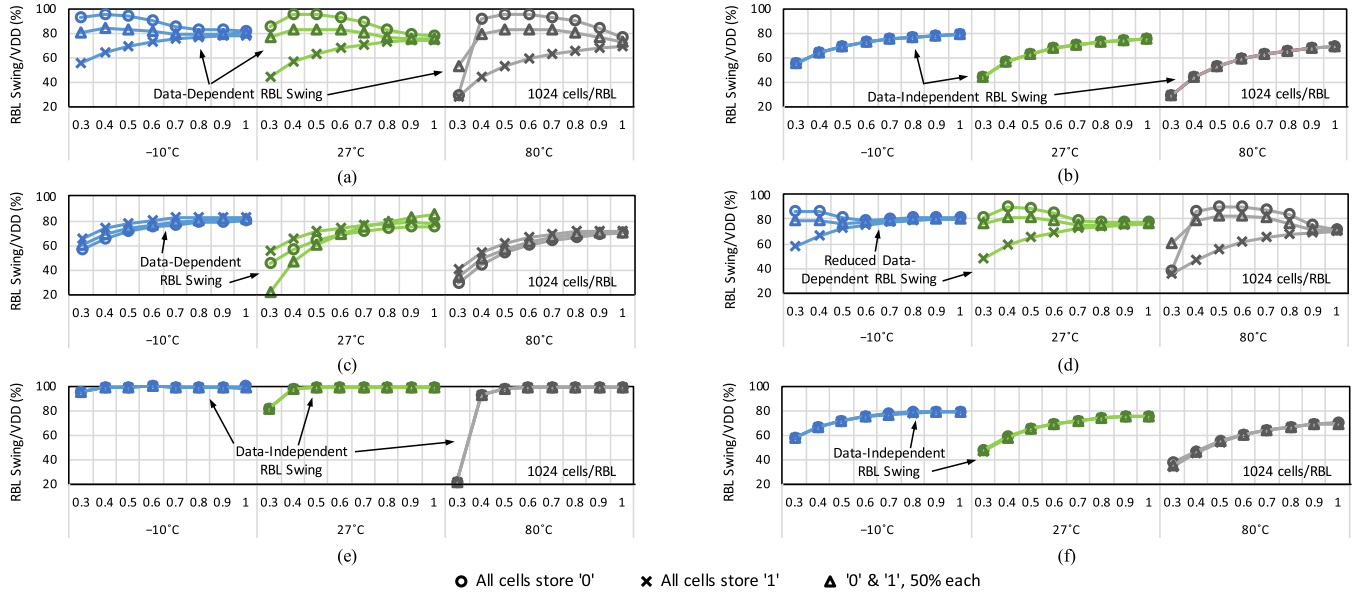


Fig. 8. Comparison of Effective Read Bit-Line swing in (a) 10T-C (b) 10T-K (c) 10T-P (d) 10T-P1 (e) 10T-P2 (f) 10T-P3 cells with varying supply voltage, temperature and data pattern.

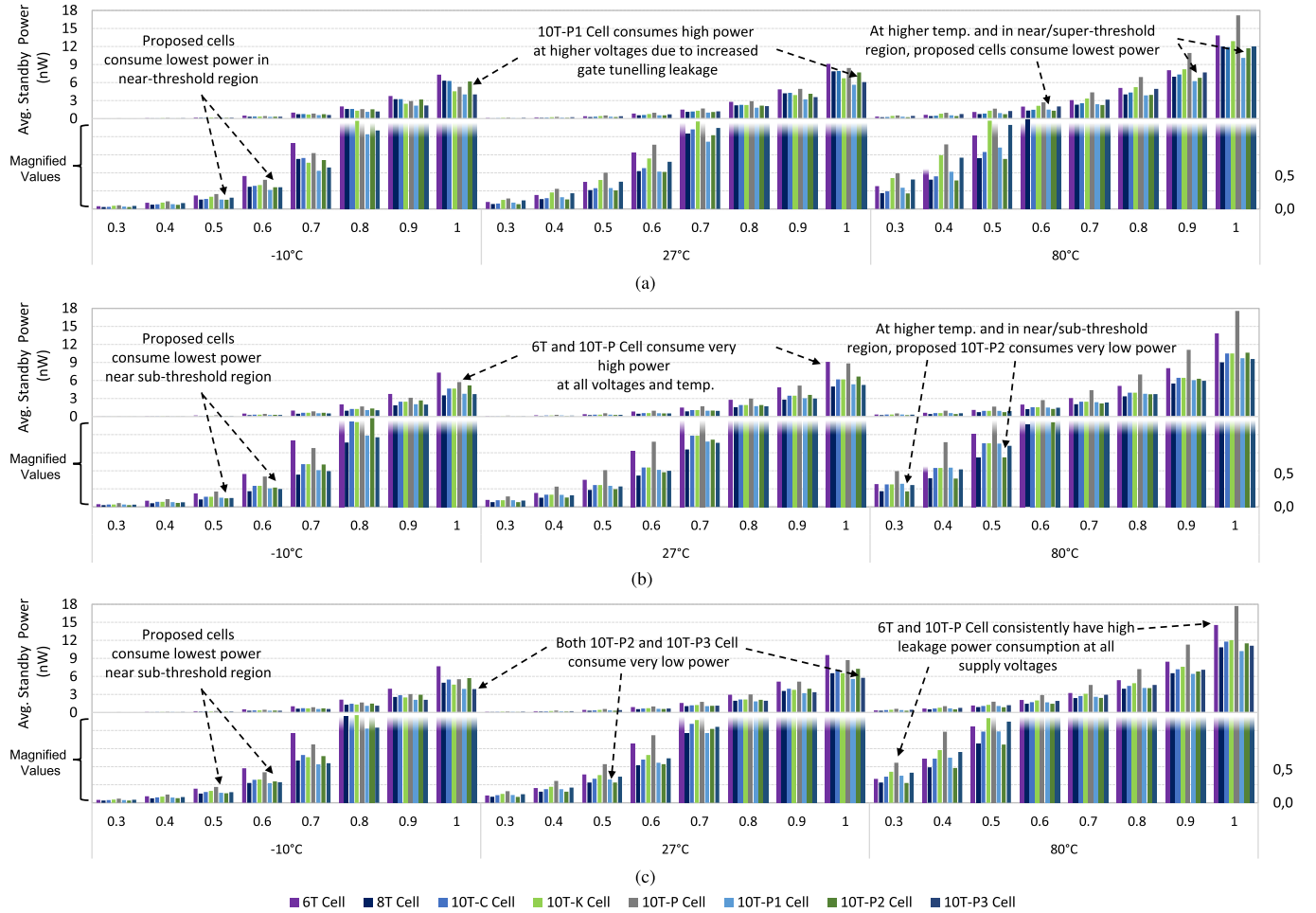


Fig. 9. Comparison of average standby leakage power per cell with varying supply voltage and temperature when (a) All cells store 'zero' (b) All cells store 'one' (c) 'zero' and 'one' are equally distributed.

standby power the dominant component in these operating regions [32]. However, when operating in these regions, a significant loss in performance is also observed due to the

deteriorated $V_{DD}-V_{TH}$, thereby leading to several orders of magnitude of diminution in frequency. Although an absolute reduction in power can be observed, the exponential rise in

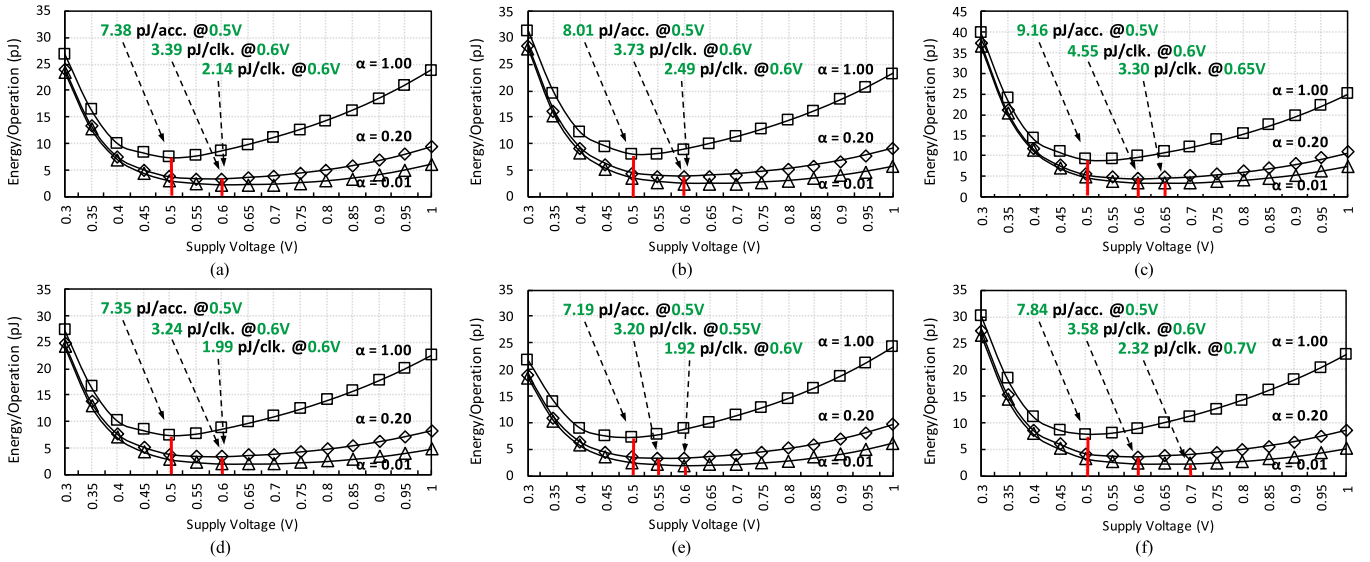


Fig. 10. Comparison of total energy per access in (a) 10T-C (b) 10T-K (c) 10T-P (d) 10T-P1 (e) 10T-P2 (f) 10T-P3 cells with varying supply voltage and activity factor.

delay in sub-threshold region increases the total energy per access [33]. As a consequence of the contention among the varying trends of power and delay, a point where energy per access reaches a minimum value is observed. This point is referred to as E_{min} and is located near the threshold point. However, the E_{min} point can vary according to the activity factor of the cache memory. L1 caches, which are small in size (\sim KBs) and dedicated for each core in a processor, are accessed more often, leading to higher switching activity. Whereas, shared L3 caches are much larger in size (\sim MBs) and have low access rates and therefore low switching activity [33]. Based on activity factor, their total power consumption per cycle can vary greatly. For L1 caches with higher activity factor, the E_{min} point is much closer to sub-threshold region. In such a case, voltage scaling is beneficial for lowering energy consumption. On the other hand, the leakage power dominates in L3 caches. As such, the E_{min} point lies closer to nominal voltages, thereby making ultra-low voltage operation contrary to the intention of energy saving.

In this work, we evaluate the energy per access for $\alpha = 1.0, 0.2$ and 0.01 . The E_{min} point is evaluated when energy per access reaches minimum value. When the difference between two adjacent E_{min} values is negligible, the point with higher corresponding supply voltage is chosen to achieve higher performance and yield. The comparison for all cells is shown in Fig. 10. For a fair comparison, the operating frequency is assumed to be the same for all cells and is shown in Fig. 11(a).

As seen in Fig. 10, the E_{min} point for high activity factor lies at the threshold point ($V_{THN} \approx 0.5V$) and subsequently shifts towards nominal voltage ($0.8V$) as activity factor decreases. The proposed 10T-P2 cell, which is aimed at low-power consumption, observes at least 2.6% and up to 21.5% energy savings in comparison to previous 10T cell-based works. The 10T-P1 cell also observes at least 0.4% and up to 19.8% energy

savings in comparison to these previous 10T cell-based works. The 10T-P3 cell implements an additional rail RWL , thereby leading to increased standby and active power. Although this leads to the 10T-P3 cell having a 6.2% energy overhead over the 10T-C cell, it offers a 19.5% area saving over it, thereby making it a good trade-off. Additionally, the 10T-P3 cell observes 2.1% and 14.4% in energy savings over the 10T-K and 10T-P cell respectively.

E. Access Failure Probability

An access failure occurs when the read bit-line fails to develop conditionally within the duration of the pulse width of the clock. For differentially sensing cells, the access time is calculated till the development of differential voltage between bit-lines. For single ended cells, the access time is calculated as the total time elapsed from global word-line enable to global read bit-line evaluation. The full swing sensing of single ended cells is much slower than differential sensing, thereby increasing the total access time and failure probability.

The probability of access failure has been calculated for a $3\sigma/\mu = 30\%$ variation in V_{TH} using the method provided in [34]. The access V_{min} is determined at the 6σ failure probability (i.e., when $P_{FAIL} = 10^{-9}$) and is shown in Fig. 11. Four cases of operating frequency corresponding to the delay of 549-FO4, 599-FO4, 649-FO4 and 699-FO4 inverters have been considered. As seen in Fig. 11, all proposed read ports have lower access failure probability than previous work. The proposed 10T-P1 cell, which is aimed at high performance, has an access V_{min} of 483mV in the best case, which is up to about 100mV lower than the previous 10T cell-based works. In other words, the 10T-P1 cell can be run at higher frequency than other cells at any given voltage.

F. Cell Area Comparison

The thin-form layout [35] for all SRAM cells compared in this work is shown in Fig. 12. The cells have been drawn on

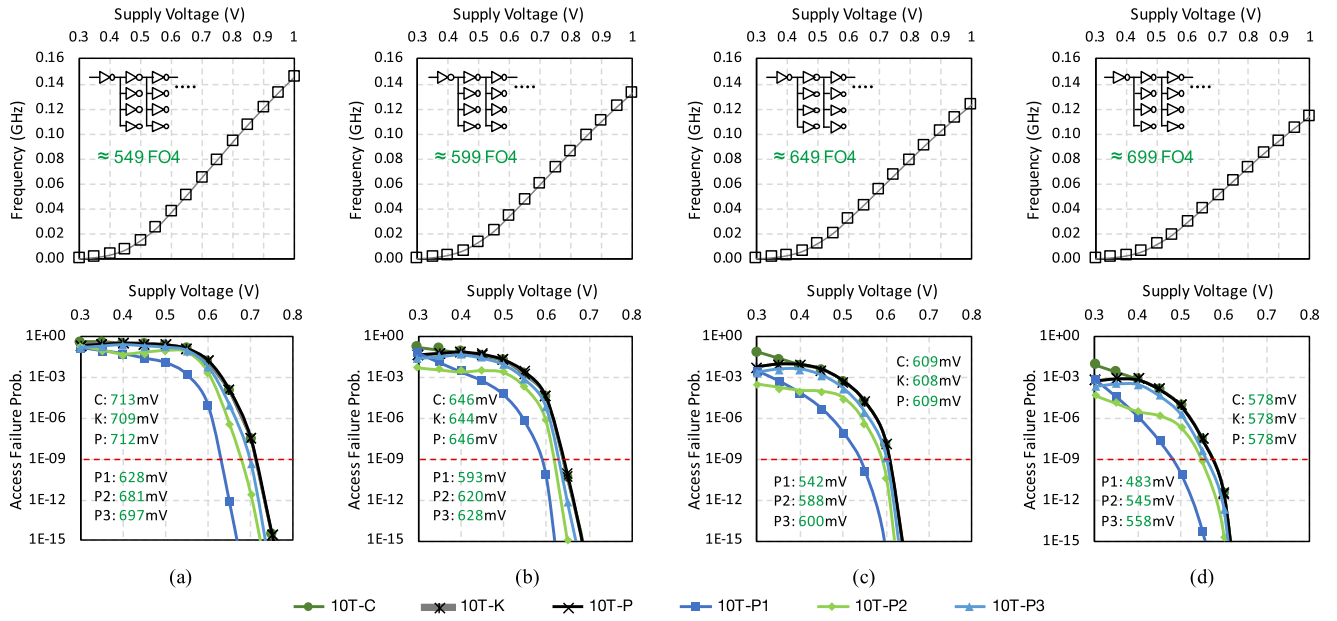


Fig. 11. Access failure probability at operating frequency corresponding to the delay of (a) 549-FO4 (b) 599-FO4 (c) 649-FO4 (d) 699-FO4 inverters.

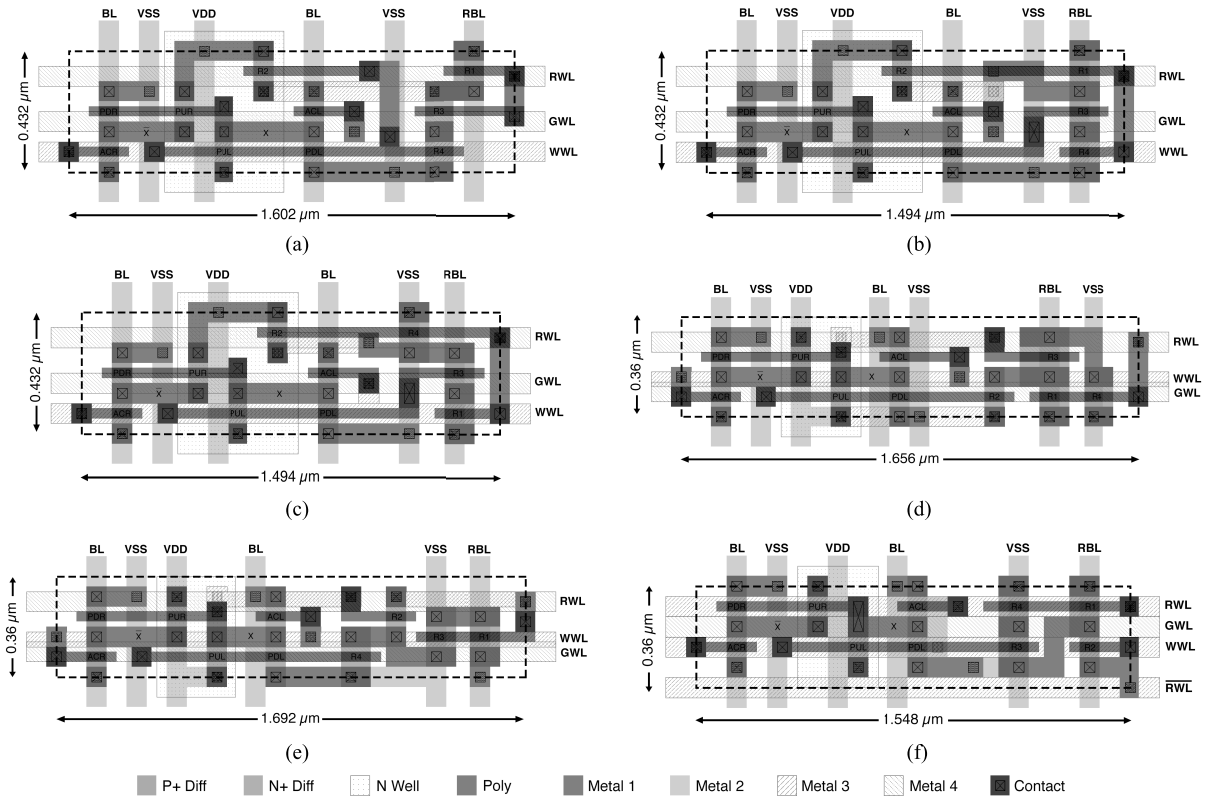


Fig. 12. Layout of (a) 10T-C (b) 10T-K (c) 10T-P (d) 10T-P1 (e) 10T-P2 (f) 10T-P3 SRAM cells.

the 32nm technology node ($\lambda = 0.018\mu\text{m}$). The absolute and normalized area comparison is shown in Table II. As observed in Fig. 12, the cells from previous works have additional implementations of pMOS besides the inverter pair. This increases the n-well size, which leads to longer vertical sizing and ultimately longer bit-lines. Subsequently, the dynamic read

performance and power consumption are also deteriorated. On the other hand, all proposed cells have smaller vertical sizing, which leads to smaller bit-line capacitance, thereby decreasing power consumption and increasing performance. As seen in Fig. 12, the 10T-C cell has the largest area, whereas the proposed 10T-P3 cell has the lowest area. When comparing

TABLE II
PERFORMANCE COMPARISON OF VARIOUS BIT CELL TOPOLOGIES

S. No.	Bit Cell Topology	Read V_{\min} (mV)	Access V_{\min} (mV)	Area (μm^2)	Normalized Area	Worst Mean RBL Swing (%)	E_{\min} (pJ/acc) $\alpha = 1$	E_{\min} (pJ/clock) $\alpha = 0.2$	E_{\min} (pJ/clock) $\alpha = 0.01$	E_{\min} at/above V_{\min} (pJ/clock)
1	6T	985	663	0.23846	0.34	-	9.27	4.44	2.78	29.85
2	8T (1024 cells/RBL) \dagger	465	Fail	0.38880	0.56	0.0	-	-	-	-
3	8T (16 cells/RBL)	465	614	0.38880	0.56	96.23	6.91	3.67	2.46	8.73
3	10T-C	465	578	0.69206	1	65.38	7.38	3.39	2.14	8.39
4	10T-K	465	578	0.64541	0.93	65.38	8.01	3.73	2.49	8.77
5	10T-P	465	578	0.64541	0.93	65.09	9.16	4.55	3.30	9.66
6	10T-P1	465	483	0.59616	0.86	67.19	7.35	3.24	1.99	7.35
7	10T-P2	465	545	0.60912	0.88	97.34	7.19	3.20	1.92	7.87
8	10T-P3	465	558	0.55728	0.81	67.12	7.84	3.58	2.32	8.43
	Best Cell	-	10T-P1*	6T**	6T	10T-P2	8T***	8T	8T	10T-P1

* Highest Performance and Yield

** 10T-P3 has lowest area in comparison to other 10T cells

*** 10T-P2 has lowest Energy per Access for 1k cells/RBL

 \dagger Array-area efficiency for the 8T cell for 1024 cells/RBL is 60-70% but falls to 30-40% for 16 cells/RBL.

proposed cells with other 10T cells, up to 19.5% and 5.6% area savings can be observed in the best and worst case respectively. These savings come at the expense of ultra-low voltage operation capability. However, we have shown that such capability is not required since the minimum energy point lies near and above the threshold point.

IV. CONCLUSION

In this work, we presented three niche-specific read ports with enhanced data-independent read port leakage for SRAM cells aimed at high performance, low power and low area respectively. All three proposed read ports did not implement any pMOS, thereby leading to smaller n-well size, which in turn led to smaller vertical sizing and shorter bit-lines in the thin-form layouts. This reduced the area per cell and energy per access. Each of the SRAM cells with the proposed read ports improved the effective read bit-line voltage swing and enabled 1k cells per read bit-line, allowing great potential for area saving in terms of sharing peripheral circuitry. With a unique topology in each of the three cells' read port, we obtain a best-case access V_{\min} of 483mV for the 10T-P1 cell, an E_{\min} of 7.19pJ/acc for the 10T-P2 cell, and a low area of $0.55728\mu\text{m}^2$ for the 10T-P3 cell. In comparison to conventional cells, this translates to up to 180mV improvement in read access performance and up to four times reduction in energy per access at their respective V_{\min} . When compared to previous 10T cell-based works, about 100mV improvement in read access performance, up to 19.8% saving in energy per access, and up to 19.5% saving in area can be observed, thereby enlarging the design and application gamut for memory designers in low power sensors and battery enabled devices.

REFERENCES

- [1] B. H. Calhoun and A. P. Chandrakasan, "A 256-kb 65-nm sub-threshold SRAM design for ultra-low-voltage operation," *IEEE J. Solid-State Circuits*, vol. 42, no. 3, pp. 680-688, Mar. 2007.
- [2] G. Pasandi and S. M. Fakhraie, "A 256-kb 9T near-threshold SRAM with 1 k cells per bitline and enhanced write and read operations," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 11, pp. 2438-2446, Nov. 2015.
- [3] T.-H. Kim, J. Liu, J. Keane, and C. H. Kim, "A 0.2 V, 480 kb subthreshold SRAM with 1 k cells per bitline for ultra-low-voltage computing," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 518-529, Feb. 2008.
- [4] V. P.-H. Hu, M.-L. Fan, P. Su, and C.-T. Chuang, "Analysis of GeOI FinFET 6T SRAM cells with variation-tolerant WLUD read-assist and TVC write-assist," *IEEE Trans. Electron Devices*, vol. 62, no. 6, pp. 1710-1715, Jun. 2015.
- [5] S. Ohbayashi *et al.*, "A 65-nm SoC embedded 6T-SRAM designed for manufacturability with read and write operation stabilizing circuits," *IEEE J. Solid-State Circuits*, vol. 42, no. 4, pp. 820-829, Apr. 2007.
- [6] M. E. Sinangil, H. Mair, and A. P. Chandrakasan, "A 28 nm high-density 6T SRAM with optimized peripheral-assist circuits for operation down to 0.6 V," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2011, pp. 260-262.
- [7] M. Yabuuchi, K. Nii, Y. Tsukamoto, S. Ohbayashi, Y. Nakase, and H. Shinohara, "A 45 nm 0.6 V cross-point 8T SRAM with negative biased read/write assist," in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2009, pp. 158-159.
- [8] B. Zimmer *et al.*, "SRAM assist techniques for operation in a wide voltage range in 28-nm CMOS," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 12, pp. 853-857, Dec. 2012.
- [9] H. Pilo *et al.*, "A 64 Mb SRAM in 32 nm high- k metal-gate SOI technology with 0.7 V operation enabled by stability, write-ability and read-ability enhancements," *IEEE J. Solid-State Circuits*, vol. 47, no. 1, pp. 97-106, Jan. 2012.
- [10] A. J. Bhavnagarwala *et al.*, "A sub-600-mV, fluctuation tolerant 65-nm CMOS SRAM array with dynamic cell biasing," *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 946-955, Apr. 2008.
- [11] M. H. Abu-Rahma, M. Anis, and S. S. Yoon, "A robust single supply voltage SRAM read assist technique using selective precharge," in *Proc. 34th Eur. Solid-State Circuits Conf.*, 2008, pp. 234-237.
- [12] D. Anh-Tuan, J. Y. S. Low, J. Y. L. Low, Z.-H. Kong, X. Tan, and K.-S. Yeo, "An 8T differential SRAM with improved noise margin for bit-interleaving in 65 nm CMOS," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 6, pp. 1252-1263, Jun. 2011.
- [13] K. Shin, W. Choi, and J. Park, "Half-select free and bit-line sharing 9T SRAM for reliable supply voltage scaling," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 8, pp. 2036-2048, Aug. 2017.
- [14] M.-H. Tu *et al.*, "A single-ended disturb-free 9T subthreshold SRAM with cross-point data-aware write word-line structure, negative bit-line, and adaptive read operation timing tracing," *IEEE J. Solid-State Circuits*, vol. 47, no. 6, pp. 1469-1482, Jun. 2012.

- [15] M. E. Sinangil, N. Verma, and A. P. Chandrakasan, "A 45 nm 0.5 V 8T column-interleaved SRAM with on-chip reference selection loop for sense-amplifier," in *Proc. IEEE Asian Solid-State Circuits Conf.*, Nov. 2009, pp. 225–228.
- [16] Y.-P. Fang and A. S. Oates, "Characterization of single bit and multiple cell soft error events in planar and FinFET SRAMs," *IEEE Trans. Device Mater. Rel.*, vol. 16, no. 2, pp. 132–137, Jun. 2016.
- [17] P. Hazucha *et al.*, "Neutron soft error rate measurements in a 90-nm CMOS process and scaling trends in SRAM from 0.25- μ m to 90-nm generation," in *IEDM Tech. Dig.*, Dec. 2003, pp. 21.5.1–21.5.4.
- [18] H. Fuketa, M. Hashimoto, Y. Mitsuyama, and T. Onoye, "Alpha-particle-induced soft errors and multiple cell upsets in 65-nm 10T subthreshold SRAM," in *Proc. IEEE Int. Rel. Phys. Symp.*, May 2010, pp. 213–217.
- [19] S. M. Jahinuzzaman, J. S. Shah, D. J. Rennie, and M. Sachdev, "Design and analysis of a 5.3-pJ 64-kb gated ground SRAM with multiword ECC," *IEEE J. Solid-State Circuits*, vol. 44, no. 9, pp. 2543–2553, Sep. 2009.
- [20] A. Kumar, H. Qin, P. Ishwar, J. Rabaey, and K. Ramchandran, "Fundamental data retention limits in SRAM standby experimental results," in *Proc. 9th Int. Symp. Qual. Electron. Design*, Mar. 2008, pp. 92–97.
- [21] R. V. Joshi, R. Kanj, and V. Ramadurai, "A novel column-decoupled 8T cell for low-power differential and domino-based SRAM design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 5, pp. 869–882, May 2011.
- [22] C.-Y. Lu *et al.*, "A 0.325 V, 600-kHz, 40-nm 72-kb 9T subthreshold SRAM with aligned boosted write wordline and negative write bitline write-assist," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 5, pp. 958–962, May 2015.
- [23] D.-P. Wang, H.-J. Lin, C.-T. Chuang, and W. Hwang, "Low-power multiport SRAM with cross-point write word-lines, shared write bitlines, and shared write row-access transistors," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 61, no. 3, pp. 188–192, Mar. 2014.
- [24] I. J. Chang, J.-J. Kim, S. P. Park, and K. Roy, "A 32 kb 10T sub-threshold SRAM array with bit-interleaving and differential read scheme in 90 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 44, no. 2, pp. 650–658, Feb. 2009.
- [25] L. Chang *et al.*, "A 5.3 GHz 8T-SRAM with operation down to 0.41 V in 65 nm CMOS," in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2007, pp. 252–253.
- [26] J. P. Kulkarni and K. Roy, "Ultralow-voltage process-variation-tolerant Schmitt-trigger-based SRAM design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 2, pp. 319–332, Feb. 2012.
- [27] H. Jeong, T. Kim, T. Song, G. Kim, and S.-O. Jung, "Trip-point bit-line precharge sensing scheme for single-ended SRAM," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 7, pp. 1370–1374, Jul. 2015.
- [28] B. Wang, T. Q. Nguyen, A. T. Do, J. Zhou, M. Je, and T. T. H. Kim, "Design of an ultra-low voltage 9T SRAM with equalized bitline leakage and CAM-assisted energy efficiency improvement," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 2, pp. 441–448, Feb. 2015.
- [29] T.-H. Kim, J. Liu, and C. H. Kim, "A voltage scalable 0.26 V, 64 kb 8T SRAM with V_{min} lowering techniques and deep sleep mode," *IEEE J. Solid-State Circuits*, vol. 44, no. 6, pp. 1785–1795, Jun. 2009.
- [30] W. Belluomini *et al.*, "Limited switch dynamic logic circuits for high-speed low-power circuit design," *IBM J. Res. Develop.*, vol. 50, nos. 2–3, pp. 277–286, 2006.
- [31] T. Hirose *et al.*, "A 20-ns 4-Mb CMOS SRAM with hierarchical word decoding architecture," *IEEE J. Solid-State Circuits*, vol. 25, no. 5, pp. 1068–1074, Oct. 1990.
- [32] E. Morifuji, T. Yoshida, M. Kanda, S. Matsuda, S. Yamada, and F. Matsuoka, "Supply and threshold-voltage trends for scaled logic and SRAM MOSFETs," *IEEE Trans. Electron Devices*, vol. 53, no. 6, pp. 1427–1432, Jun. 2006.
- [33] G. Chen *et al.*, "Yield-driven near-threshold SRAM design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 18, no. 11, pp. 1590–1598, Nov. 2010.
- [34] S. Gupta, K. Gupta, and N. Pandey, "Pentavariate V_{min} analysis of a subthreshold 10T SRAM bit cell with variation tolerant write and divided bit-line read," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 10, pp. 3326–3337, Oct. 2018.
- [35] K. Osada *et al.*, "Universal- V_{dd} 0.65–2.0-V 32-kB cache using a voltage-adapted timing-generation scheme and a lithographically symmetrical cell," *IEEE J. Solid-State Circuits*, vol. 36, no. 11, pp. 1738–1744, Nov. 2001.



Shourya Gupta (S'17) was born in New Delhi, India, in 1994. He received the B.Tech. degree in electronics and communication engineering from Guru Gobind Singh Indraprastha University, New Delhi, in 2017. He is currently pursuing the Ph.D. degree in electrical engineering with the University of Virginia, Charlottesville, VA, USA.

His current research interests include the design of low-power logic and memory circuits in emerging and exploratory technologies.



Kirti Gupta (M'15–SM'18) received the B.Tech. degree in electronics and communication engineering from the Indira Gandhi Institute of Technology, New Delhi, India, in 2002, the M.Tech. degree in information technology from the School of Information Technology, New Delhi, in 2006, and the Ph.D. degree in electronics and communication engineering from Delhi Technological University, New Delhi, in 2016.

From 2002 to 2007, she was an Assistant Professor, and from 2008 to 2016, she was an Associate Professor with the Bharati Vidyapeeth's College of Engineering, New Delhi, where she is currently a Professor. She has authored over 70 technical papers in various international conferences and journals. Her current research interests include digital integrated circuit design.

Dr. Gupta is a Life Member of ISTE.



Benton H. Calhoun (SM'12) received the B.S. degree from the University of Virginia in 2000 and the M.S. degree and the Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology in 2002 and 2006, respectively. He is currently a Professor with the Electrical and Computer Engineering Department, University of Virginia. He is the Campus Director and the Technical Thrust Leader with the NSF Nanosystems Engineering Research Center for Advanced Self-Powered Systems of Integrated Sensors and Technologies.

He has co-authored *Sub-Threshold Design for Ultra Low-Power Systems* (Springer, 2006) and has authored *Design Principles for Digital CMOS Integrated Circuit Design* (NTS Press, 2012). His research interests include body area sensor nodes, wireless sensor networks, low-power digital circuit design, sub-threshold digital circuits, sub-threshold field-programmable gate arrays, static random access memory design for end-of-the-roadmap silicon, power delivery circuits and architectures, variation tolerant circuit design methodologies, and low-energy electronics for medical applications.



Neeta Pandey (M'04–SM'14) received the M.E. degree in microelectronics from the Birla Institute of Technology and Sciences, Pilani, India, and the Ph.D. degree from Guru Gobind Singh Indraprastha University, New Delhi, India.

She was with the Central Electronics Engineering Research Institute, Pilani; IIT Delhi, New Delhi; the Priyadarshini College of Computer Science, Noida; and the Bharati Vidyapeeth's College of Engineering, New Delhi, in various capacities. She is currently a Professor with the Electronics and Communication Engineering Department, Delhi Technological University, New Delhi. She has authored over 200 technical papers in reputed national and international conferences and journals. Her current research interests include analog and digital very large-scale integration design.

Dr. Pandey is a member of the WIE Affinity Group and a Life Member of ISTE.