CHAPTER 2
BACKGROUND AND LITERATURE SURVEY

This chapter briefly explains the principles of operation of different SRAM cells (4T SRAM, 6T SRAM and 6T FGSRAM), their design considerations, leakage current components of an SRAM cell, literature survey and different techniques used to reduce power dissipation in SRAM.

2.1 4T SRAM Cell

Figure 2.1 Resistive Load 4T SRAM Cell

Figure 2.1 shows schematic of 4T SRAM cell that consists of four NMOS transistors and two poly-load resistors [26]. Word line (WL) input is connected to the gate of the two pass transistor M3 and M4 transistors and the cell is tied by the column. The two other NMOS transistors M1 and M2 are called the pull-down network of the inverters. The loads of the inverters are high value poly-silicon resistor. Though the 4T SRAM
cell can be less significant than the 6T SRAM cell, it is at rest regarding four times as big as the cell of a comparable invention DRAM cell.

The difficulty of the 4T SRAM cell is to create a high load resistor (in the range of giga-ohms) to reduce the current. On the other hand, this resistor should not be too high to assure for better performance. The 4T SRAM cells have a number of disadvantages in spite of its advantage of size. Further, the main disadvantage with a 4T SRAM cell is continuous flow of current through one of the pull-up transistor which causes increased static power. These include the reality that each one SRAM cell has a high static current. By the leakage current of the pass transistors in the 4T SRAM cells the retention of the data is ensured. Thus, for low power applications the 4T SRAM cells are not proper aspirant.

As the resistance is much high, and the cell does not have high-speed as the 6T SRAM cell, so the cell is sensitive to soft error and noise [26].

Besides that, the stability of the data in a 6T SRAM cell is free of the leakage current. In addition, the 6T SRAM arrangement exhibits a considerably higher tolerance together to noise which is a significant benefit particularly in the advance technologies where the noise margins have been shrinking. Because of this, 6T SRAM cell has gained more popularity than the 4T SRAM cell configuration in low power SRAM design.

2.2 6T SRAM Cell

A six-transistor SRAM cell (6T SRAM cell) is conventionally used as the memory cell. It has three different states. It operates in three modes, as read, write and hold operation. The SRAM should have "read stability" and "write stability" for read and write operation. Though, the 6T SRAM cell makes a cell of bigger size than that of a DRAM cell, producing in a small memory density. Therefore, to meet the increasing demand of huge capacity memory in mobile applications conventional 6T SRAM cells are not suitable due to complexity. It also exhibits poor noise immunity and stability at low device size and low supply voltage. Because of the voltage division
among the driver and access transistors the stability of SRAM cell decreases during read operation.

Figure 2.2 Basic Structure of SRAM Cell

Figure 2.3 Schematic of 6T SRAM Cell
Figure 2.2 shows the basic structure of SRAM cell. The Figure 2.3 shows the conventional 6T SRAM cell where the two cross coupled inverters are connected back to back (M5 and M6 are the pull-up transistors and the pull-down transistors are M1 and M2). M3 and M4 two access transistors. These transistors are turned on as soon as the horizontal word line (WL) is enabled and a row is selected form the array. It connects the bit line (BL) and bit line bar (BL_bar) to the storage nodes of SRAM cell. For write and read operations they allow contact to the cell [26-28].

2.2.1 Read Operation of 6T SRAM Cell

In Figure 2.4, it is assumed that data '0' and data ‘1’ are stored on the left side of the cell that is at node "1" and on the right side that is at node "2", respectively. Therefore, the driver transistor M2 is off and transistor M1 is on. Initially BL and BL_bar are precharged by a precharge circuit to a voltage just about VDD. The pass transistor M3 and M4 turns on while WL = 1 and current starts to flow through the transistor M3 and M1 to ground which is shown in Figure 2.4. As there is no path to ground through transistor M2, the resulting cell current gradually discharges the bit line capacitance (CC) for the meantime and on the other side of the cell the BL_bar voltage remains high. To produce a valid low output, voltage difference of BL and BL_bar is provided to the sense amplifier. The ratio of the strength of the pull-down
transistor to the pass-gate transistor is called the cell-ratio and it should be suitably large to make sure that the read disorder does not occur [28] i.e. the pass transistor must be weaker than the pull down transistor to guarantee a stable operation. The transistor ratio W1/W3 should be at least 1.3 [26]. To operate the cell reliably, the sizes of the transistors should be properly designed as equation 1 & 2.

\[ V_2 < V_{T,1} \]

\[ \frac{k_{n,3}}{k_{n,1}} = \frac{W_3}{L_3} \frac{W_1}{L_1} < \frac{2(V_{DD} - 1.5V_{T,n})V_{T,n}}{(V_{DD} - 2V_{T,n})^2} \]

DC analysis of the SRAM cell transistors operation is conventionally accepted to make sure the cell stability throughout the read operation [22]. Since, it was mentioned earlier, a little voltage difference between BL and BL_bar \( \Delta V \) makes sure that the output of the inverter constructed from M6 and M2 transistor stays stable at node 2. M1 and M3 form a voltage divider and raise node "1" voltage by \( \Delta V \). The voltage level \( \Delta V \) is managed by the resistive ratio of transistor M3 and M1 to make sure a non-critical read operation.

### 2.2.2 Cell-Ratio (CR)

Cell-ratio is the ratio among sizes of the driver transistor and pass transistor for the period of the read operation [12]. Cell-ratio (CR) for Figure 2.4 can be represented as:

\[ \text{CR} = \frac{W_1}{L_1} / \frac{W_3}{L_3} \]

where, W and L are the width and length of the corresponding MOS transistors respectively. A higher cell-ratio leads to a lower \( \Delta V \) and results in a more stable read operation.

### 2.2.3 Read Access Time (TRA)

When BL or BL_bar is released by 50-mV since its preliminary high level, word line is activated and read access time (TRA) or read delay is obtained [29]. Keep away from misread the 50-mV difference between BL and BL_bar is sufficient to be detected by a Sense Amplifier [26]. For read operation, the bit lines are precharged and transistor M1 is turned on. Node “2” stores a data ‘1’ and node “1” stores a data ‘0’. When WL is activated, BL falls through transistor M3 or transistor M1.
The delay between the application of the WL signal and the response time of the sense amplifier is known as the read delay. It can also be represented by the delay occupied in allocating the bit lines (BL and BL_bar) to release by about 10% of the peak rate.

### 2.3 Write Operation of 6T SRAM Cell

![Figure 2.5 Circuit of 6T SRAM Cell for Write Operation](Image)

Assume that, '1' is stored in the cell as shown in Figure 2.5. The transistor M5 and M2 will operate in the active mode and transistor M6 and M1 will be turned off. Thus, at the beginning the internal node voltages are V1 (voltage at node "1") = VDD and V2 (voltage at node "2") = 0 V. Now to perform a write '0' operation, cell access transistor M3 and M4 are turned on and BL is set low and BL_bar is set high. The node "2" voltage level should not be large enough to turn on NMOS M1, otherwise the voltage at node "1" will be discharged through M1. So, the following two conditions (3) and (4) must be satisfied for write operation [26].

\[ V_1 < V_{T,2} \] .................................................................(3)

\[
\frac{k_{p,5}}{k_{n,3}} = \frac{W_5}{L_5} \frac{W_3}{L_3} < \frac{\mu_n}{\mu_p} \frac{2(V_{DD} - 1.5V_{T,n})V_{T,n}}{(V_{DD} + 2V_{T,p})^2} \] ...........................................(4)
The discharging force of the access transistor should rise above the restoring force of the pull-up transistor. The ratio of the strength of the access transistor to the pull-up transistor should be adequately large to guarantee that write failure is not to be occurred [22]. When the write operation is completed, both WL and write-enable signals are turned off. Then the next precharge circuit is activated to prepare the bit lines for the next access operation. Both inverter 1 and inverter 2 have the same switching voltage because they have the same transistor sizing. The cell ratio should be less than 1.5 for a stable write operation [22]. A successful write operation can be guaranteed by choosing a proper pull-up ratio (PR). In a smaller ΔV, a smaller pull-up ratio results and a smaller ΔV connected with superior drive on the input of inverter constructed from transistor M5-M1. To achieve a low PR, a wider access transistor is desirable. However, increasing the width of the access transistor degrades the stability of the cell during the read operation by asserting (cell-ratio) CR.

2.3.1 Pull-Up Ratio (PR)

In the write operation, the ratio among the sizes of the load transistor and the pass transistor is known as pull-up ratio [12]. In the Figure 2.5, pull-up ratio can be expressed as $PR = \frac{W6/L6}{W4/L4}$. For high density and reliable read/write operations, both the pull-up PMOS and the pass NMOS transistors have the minimum width, while the pull down NMOS transistors have double the minimum width.

2.3.2 Write Access Time (TWA)

Write access time (TWA) or write delay for writing zero (’0’) to storage node "1" is the time when word line (WL) is turned on to the time when storage node "2" falls to 10% of its preliminary high level (i.e. it’s 90% swing). Similarly, write access time for writing one (’1’) to storage node "2" is the time when WL is turned on to the time when storage node "2" increase to 90% of its starting low level. Write delay is the delay between the application of the WL signal and the time at which the data is written. The write delay is proportional to the resistance of the driver and the resistance of the pass transistor in addition to bit line capacitance.
2.4 Hold or Standby Operation

The pass transistors M3 and M4 disconnect the SRAM cell from the bit lines (BL and BL_bar) when the WL is not connected. The two cross coupled inverters formed by M1, M5, M2 and M6 will go on to reinforce each other on condition that they are associated to the power supply [27].

2.5 Data Retention Voltage (DRV)

Data retention voltage (DRV) is the lowest amount power supply voltage essential for holding the data in the standby mode [30]. An SRAM cell has two nodes indicated as "1" and "2" for holding the data either zero ('0') or one ('1'). The supply voltage is reduced until there is flip in the state of the SRAM cell or content of the SRAM cell stay as invariable. When supply voltage scales down to data retention voltage, the voltage transfer curve (VTC) of the inner inverters disgrace to such a level that static noise margin of the SRAM cell decreases to zero [12].

The sizing of the transistors is the most important factor for appropriate working of SRAM cell. The thumb rule is that the ratio of the width of transistors should be \( \frac{w_5}{w_3} = \frac{w_3}{w_1} = 1.5 \) and \( \frac{w_6}{w_4} = \frac{w_4}{w_2} = 1.5 \). The size formation of transistor gives the correct driving voltage to transistors for on and off state [31].

2.6 Static Noise Margin

The most significant metric for the SRAM is the read stability of the circuit as per the study in the year 1987, which is in other word called the static noise margin (SNM). SNM can be described by means of the input voltage to output VTC curve. SNM is also defined as the utmost dc voltage that the cell can stand prior to changing state in read mode [32]. Good SNM is required for stability of the SRAM cell.
Figure 2.6 "Butterfly Curve" For Measuring SNM

The SNM is a magnitude of the window between the two inverter characteristics [33]. In the maximum square system, the VTC of the feed familiar inverter and the mirrored VTC of the feedback inverter is described to form a butterfly shape shown in Figure 2.6. It is a measure of the binary cell memory stability and is achieved mirroring the inverter characteristics and drawing the maximum square between them [34],[35]. The read VTC for the SRAM Cell [12], [34],[36] can be defined as the side of the largest square which can be fitted into the "eyes" of the butterfly curves. The SNM is affected by the threshold voltages of the NMOS and PMOS devices in the SRAM cells. To increase the SNM, the threshold voltage should be increased.

SNM of the SRAM cell depends on the cell-ratio, pull-up ratio and supply voltage. 70 percent value of the SNM [30], [34] depends upon the driver transistor of the SRAM cell. In Figure 2.6 voltage 1 is the voltage at node "1" of SRAM cell and voltage 2 is the voltage at node “2” of SRAM cell.

2.7 Leakage Mechanism in CMOS Transistor

The static power of a CMOS circuit is determined by leakage current through each off transistor. The leakage current is the highest contributor to the standby power...
consumption of the Intel Pentium processors and there is an on-going effort to restrain this current through device enhancement and circuit techniques [37].

**Figure 2.7. Static CMOS Leakage Sources**

Figure 2.7 shows the sources of leakage for the CMOS circuits.

### 2.7.1 PN Junction Reverse Bias Leakage Current ($I_1$)

Drain and source to well junctions are usually reverse biased and this causes the PN junction reverse biased leakage current [38], [39]. It is a function of junction area and doping concentration. There are two reasons for PN junction reverse bias leakage current. First due to minority carrier diffusion close to the boundary of the depletion region and the second is caused by electron hole pair creation of the reverse biased junction's depletion region. Additional leakage took place among the drain and well junction for an MOS transistor by overlapping of the gate to the drain-well PN junctions. If together N and P regions are heavily doped, band-to-band tunnelling (BTBT) controls the PN junction leakage [40]. PN Junction leakage current highly depends on the temperature.
2.7.2 Band-To-Band Tunnelling Current ($I_2$)

High electric field across reverse-biased p-n junction causes significant current known as band-to-band tunnelling (BTBT) current [40], which dominates the p-n junction leakage current. In scaled devices, high doping concentrations and abrupt doping profiles cause significant band-to-band tunnelling current through the drain-well junction.

2.7.3 Subthreshold Leakage Current ($I_3$)

When the threshold voltage is greater than the voltage across gate-source terminal ($V_{gs}$), the current flowing is known as subthreshold leakage current [39], [41]. Subthreshold or weak inversion conduction current among source and drain in a MOS transistor occurs when gate voltage is less than threshold voltage. Due to the low threshold voltage ($V_{th}$), weak inversion usually dominates recent device off-state leakage. The source and drain terminals are divided far enough that their depletion regions have no effect on the potential for the most part of the device in long-channel devices. Therefore, the threshold voltage is virtually self-determining of the drain bias and length of the channel in such devices. Inside a short channel device, though, the drain and source depletion thickness in the perpendicular path and the drain-source potential have a strong cause on the band bending above an important part of the device. As a result, the threshold voltage and then the short channel sub threshold current of devices differ through the drain bias. This outcome is known as drain-induced barrier lowering (DIBL) [40]. The flow of subthreshold leakage is shown in Figure 2.8. In deep sub-micron technology due to small threshold voltage of the device the subthreshold leakage current is a major contributor of total leakage power [42]. While transistors are switched off subthreshold leakage current of large quantity are produced.

2.7.4 Gate Oxide Tunneling Current($I_4$)

The electric field across the oxide increases due to the decrease in gate oxide. The large electric field together with small oxide thickness marks in electrons tunnelling from substrate to gate and furthermore from gate to substrate through the gate oxide.
This results in the gate oxide tunnelling current. The method of tunnelling among gate poly-silicon and substrate is mainly divided into two parts, that is, (I) Fowler-Nordheim (FN) tunnelling through oxide band and (II) direct tunnelling through gate. For the regular device operations Fowler-Nordheim tunnelling is small, but while the oxide thickness is smaller than 2-3 nm direct tunnelling is much vital [40].

2.7.5 Hot Carrier Injection (I₅)

In a short-channel transistor, owing to high electric field close to the Si to SiO₂ edge, electrons or holes can increase large amount of energy from the electric field to cross the edge potential barrier and go through the oxide layer. This is called as hot-carrier injection. The injection from Si to SiO₂ is added liable for electrons than holes, while electrons have a lower effective mass than holes. With this, barrier height for holes (4.5 eV) is higher than electrons (3.1eV) [40].

2.7.6 Punch Through (I₆)

The depletion areas at the source-substrate and drain-substrate junctions expand into the channel in short channel devices. As the channel length is decreased, if the doping is maintained stable, the partition among the depletion region boundaries reduces. An increase in the reverse bias across the connections furthermore pushes the links closer to each other. While the grouping of reverse bias and channel length directs to the merging of the depletion regions, punch through is occurred [40].

2.7.7 Gate Induced Drain Leakage Current (I₇)

The silicon surface below the gate has nearly equal potential as the P-type substrate while the gate is influenced to form an accumulation layer by the side of the silicon surface. Owing to existence of accumulated holes by the side of the surface, the surface works like a P-region more deeply doped than the substrate. This causes the depletion layer next to the surface to be much thinner than somewhere else. The reduced depletion layer next to or close to the surface forms field crowding or an enhance in the local electric field and so enlarging the large field effects close to that region. While the negative gate bias is much more (i.e. drain at VDD and gate at negative or zero), the n+ drain area below the gate can be depleted and yet
complemented. Resulting into a sudden increase of more field effects such as avalanche growth and band-to-band tunnelling this generates extra field crowding and increases the peak field. Because of all these causes, in the drain region minority carriers are emitted below the gate. As the substrate is at a lower potential meant for minority carriers, finishing a path for the GIDL the minority carriers that have been accumulated at the drain depletion region under the gate are removed tangentially to the substrate. Higher potential among gate and drain (higher VDD) and thinner oxide thickness increases the electric field which in turn increases GIDL [20].

In an SRAM cell with an unlike leakage current method, the subthreshold leakage of the off transistors are the major cause of the leakage current. The next significant leakage current in the SRAM cells meant for presently existing technologies is the gate brings about drain leakage which is generally larger than one order of size smaller than the subthreshold leakage in the 45nm CMOS technology.

Figure 2.8 Leakage Current in Conventional 6T SRAM Cell
2.8 Power Dissipation Components in Digital CMOS Circuits

Power dissipation in CMOS circuits consists of three main components: switching power, short-circuit power and static power or leakage power. When a conducting path among supply and ground is appeared, short-circuit power consumption takes place. To get just about identical rise and fall time the pull-down and pull-up devices must be sized correctly. This power consumption component is important in precharge circuits and estimate dynamic circuits. To carry on this power dissipation component [38] small enough, careful design is essential to ignore [43]. The leakage power and switching power are the main components of power consumption. Due to charging and discharging of internal capacitances in the circuit switching power consumption is obtained. While the device is turned off, it results the leakage power dissipation. In an SRAM array the static power consumption is generally with a view to the leakage current in SRAM cells.

Figure 2.9 Static (Leakage) and Dynamic (Switching) Power for Different Technology Generations
The leakage power is obtained by multiplying the total leakage current in the memory array and the supply voltage \([44]\). Static power has begun to generate an important portion of the whole power consumption.

The static (leakage) power for different technologies are shown in Figure 2.9 \([45]\). It is clear that static power is significantly rising with scaling of technology. The ratio of leakage power to total power is likely to go beyond 50% in 45 nm technology compared to 10% in 90 nm designs.

**2.9 Architecture of an SRAM Unit**

The SRAM memory architecture contains a matrix of SRAM cells containing rows and columns. The selection of rows is decided by the row decoder and the selection of columns is decided by the column decoder. To decode rows, the number of address lines required is ‘\(m\)’ where \(M\) is the same to \(2^m\) and to decode columns, the number address lines required is ‘\(n\)’ where \(N\) is the same to \(2^n\) if the size of the memory is \(M \times N\). The sense amplifiers and the write circuits are connected between a pair of bit lines. In read and write operation, this interface through the bit lines to correspond with the cell. During write or read operation the suitable timing signals are generated by timing and control unit to activate the word line, sense amplifier and write driver circuit, respectively. The decoders used are simple NAND type \([46]\).

**2.10 Simulation Circuit of 6T SRAM Cell**

Figure 2.10 Simulation Circuit of 6T SRAM Cell
The Figure 2.10 shows the schematic circuit of 6T SRAM Cell. It is implemented on 45nm standard CMOS technology with power supply voltage of 1 V. The word line voltage is taken as 750 mV. To maintain good read and write stability proper sizing of the transistors has been done. The width (W) size of the access transistors are taken 200 nm each, The width of pull-up transistors are taken as 145 nm and width of pull-down transistors are considered as 250 nm. The length (L) of each transistors are taken as 45 nm.

Figure 2.11 DC Analysis of 6T SRAM Cell

DC Analysis is used to check what will happen if we turn on the circuit when no signal is applied to it. It is meant to check for time-domain checking, voltage, mesh current and branch voltage and branch current. The Figure 2.11 shows the DC analysis of 6T SRAM cell. In Figure 2.11 V1 is the output voltage of inverter 1 and v2 is the output voltage of inverter 2. Vin is the input voltage to SRAM cell. The Figure 2.11 signifies the SRAM cell principle. The setup is in pre-charge mode.
The Figure 2.12 shows the transient analysis of 6T SRAM Cell. This analysis shows the output voltage of two inverters vs time. V1 is the output voltage of inverter 1 and v2 is the output voltage of inverter 2. Here, it is observed that the voltage swing is 100% which indicates better stability and good performance of the cell. This type of stimulus is reasonable for voltage swing observation and measuring delay.

Figure 2.13 Power Consumption Curve of 6T SRAM Cell
The major design constraint associated with the design and implementation of CMOS circuit is power consumption. Figure 2.13 shows the power consumption curve of 6T SRAM cell. To calculate power consumption this curve is generated at the time of simulation. Taking this curve power consumption is calculated by Cadence tool. From figure it is observed that power consumption increases with increase in supply voltage.

![Figure 2.14 Write '0' Leakage Current Curve of 6T SRAM Cell](image)

![Figure 2.15 Write '1' Leakage Current Curve of 6T SRAM Cell](image)
Leakage power consumption is attributed to the pull-up network or pull-down network that is off. Figures 2.14 and 2.15 show the write '0' and write '1' leakage current curve, respectively. This is generated from the drain node of off mode transistor at the write '0' and write '1' operation. At write '0' operation 0 V is applied at BL input and for write '1' operation 1 V is applied at BL input of the SRAM cell. Leakage power is obtained by multiplying the calculated leakage current and VDD which is placed in Table 2.1 to Table 2.4.

Figure 2.16 Write SNM Curve of 6T SRAM Cell

SNM is the standard way of analyzing SRAM cell stability. Figure 2.16 shows the write SNM curve and Figure 2.18 shows the read SNM curve of SRAM cell respectively. It is the plot of voltage transfer characteristic of inverter 2 (INV2) and the inverse VTC from inverter 1 (INV1). In Figure 2.16 voltage1 is the node "1" voltage and voltage2 is the node "2" voltage of SRAM cell. Node "1" and node "2" is the output of inverter 1 and inverter 2.
Simulation circuit of 6T SRAM cell is shown in Figure 2.17. The sense amplifier used in the circuit is to amplify the small voltage difference between BL and BL_bar.

Figure 2.18 Read SNM Curve of 6T SRAM Cell
2.11 Floating Gate MOSFET (FGMOS)

FGMOS is a device in which the second gate generally called floating gate is electrically isolated but capacitively coupled to input gates. Capacitors combining into this floating gate become effective gates of this transistor, where the strength of the gate depends upon the size of the capacitor [47]. The major advantage of using FGMOS is low power dissipation, tunability, compatibility, flexibility and controllability [48]. A control voltage present at one of the multi-input of FGMOS and facility to additional weighted inputs provides wide range of tunability to the circuit [49]. Electrons or holes can be introduced and can be stored in the floating gate by means of high electric fields; as a result change in the device threshold voltage can be read from the control gate (CG) [50]. Working of FGMOS set aside threshold voltage controllability without decreasing the feature size, thus operates at power supply voltage which are fine under the proposed functioning limit. Moreover, it consumes a smaller amount of power than the least essential power meant for a circuit designed through conventional MOSFET. In circuit terms, a floating gate is one that has no DC path to ground. Using regulated cascode configuration achieves high output impedance of the order of $g_m r_{out}^3$ where $g_m$ is device trans-conductance and $r_{out}$ is output resistance [51-53].

It is a conventional MOSFET in which the gate is capacitively coupled to the input using another poly-silicon layer. The equation that models the behaviour of floating gate voltage ($V_{FG}$) of FGMOS is given by equation (5) [49][54].

$$V_{FG} = \sum_{i=1}^{N} \frac{C_i}{C_T} V_i + \frac{C_{GS}}{C_T} V_S + \frac{C_{GD}}{C_T} V_D + \frac{Q_{FG}}{C_T}$$

(5)

$$C_T = C_{GD} + C_{GS} + C_{GB} + \sum_{i=1}^{N} C_i$$

$$C_i = \left( \frac{\varepsilon_{SiO_2}}{t_{SiO_2}} \right) A_i$$

Where $V_{FG}$ = Floating gate voltage

- $C_{GS}$ = Gate to source Capacitance
- $C_{GD}$ = Gate to drain Capacitance
An FGMOS can be fabricated through electrically isolating the gate of a standard MOS transistor, so that there are no resistive links to its gate. More than one secondary gates or inputs are deposited over the floating gate (FG) and are electrically isolated from it. These inputs are just capacitively joined to the floating gate, since the floating gate is fully enclosed with highly resistive material. Therefore, the floating gate is a floating node in terms of its DC operating point.

The floating gate, fabricated by the gate electrode (POLY1) layer, expands outer surface of the active area of the MOS transistor. The FG is enclosed with two SiO2
insulator layers and hence electrically cut off from the rest of the device. The inputs of
the device are placed on top SiO2 layer and are fabricated using second poly-silicon
(POLY2) layer as shown in Figure 2.19. The input electrode size gives the
capacitance value of capacitors which connect the FGMOS inputs through the floating
gate and they can be varied according to the need of the designer. The input
capacitance values are given by:

$$C_i = \left( \frac{\varepsilon_{SiO_2}}{t_{SiO_2}} \right) A_i$$

Where $\varepsilon_{SiO_2}$ is the permittivity of the SiO2, $t_{SiO_2}$ is the width of the SiO2 among the
floating gate and the effective inputs and $A_i$ is the area of each one input capacitor
plate.

The basic structure of an N-channel FGMOS transistor with N-input voltages VG1, VG2,...,VGN and the symbolic representation of such device is shown in Figure 2.19
(a) and Figure 2.19 (b) [55] respectively. It is analogous to the conventional MOSFET
by means that floating gate is the same to the gate of conventional MOSFET. The
floating gate is produced by means of the first poly-silicon layer above the n-channel
whereas the multiple-input gates are generated by means of the second poly-silicon
layer and all are placed above the floating gate. To the multiple input gates, this
floating gate is capacitively coupled. The voltage of floating gate is not controlled
directly but with control gates during capacitor coupling. The equivalent threshold
voltage can be changed by some programming techniques since noticed from the
control gates to have a low threshold voltage (lVth) MOSFET. Relatively complex
programming circuits and higher programming voltage are its disadvantages. Now we
can choose any of above techniques or combination of techniques to achieve low
power in analog circuit design despite knowing the fact that designing such low
voltage/low power analog technique is a challenging task and also there are no well
established design technologies and theories in this area. Low power and low voltage
techniques based on FGMOS have been adopted for the proposed circuit and is within
the scope of this thesis report.
Designing with FGMOS is not an easy task for a number of reasons that will be subsequently described. In normal situation inside a circuit a floating node signifies an ‘error’ owing to the fact that the preliminary condition is unidentified except it is by some means fixed [49]. This produces two problems of unlike in nature. First, it is not right to simulate these circuits. Secondly, an unidentified amount of charge may possibly stay entrapped at the floating gate through the fabrication method that will outcome in an unfamiliar initial situation for the floating gate voltage. Hence, a model proposed by Yin et al. has been utilized for circuit simulation [54].

Figure 2.20 Equivalent Schematic for N-channel N-Input FGMOS

The equivalent schematic for N-channel N-Input FGMOS is shown in Figure 2.20. Here, the resistors are connected in parallel with the input capacitors that is used to simulate the circuit [54], [56].

Figure 2.21 Model for FGMOS
The model for FGMOS is shown in Figure 2.21. In the model, the voltages $V_1$, $V_2$, ..., $V_N$ are the input voltages which are coupled to gate by resistor-capacitor parallel combination. $V_F$ and $V_0$ are floating gate voltage and substrate voltage respectively. $M$ is a conventional MOSFET. A typical value of $C_i$ that has been taken in this thesis is 1fF for the model shown above and that of the resistance value is taken as 1TΩ.

2.12 Why FGMOS?

To design circuits that (a) can work at levels of power supply voltage that are well under the proposed working limits for a meticulous technology and (b) consume low power than the least essential circuit power designed by means of only MOS devices within the same technology and same performance. The two major design objectives can be low voltage and low power (LV/LP). These two can be realized by following four altered sub-goals.

1. Decrease the complexity of circuit: Because the circuitry gets easier, smaller amount of current divisions are essential, as a result the power consumption reduces. Other advantages related to the response of frequency are due to less number of internal nodes.

2. Simplification of the signal processing: Utilizing FGMOS transistors, complex jobs are simpler to realize. These can be applied in nonlinear signal processing to minimize the demand of the voltage.

3. Shifting the levels of the signal: The devices can be biased within the most suitable operating region meant for a broad collection of input signals, by changing the useful threshold voltages consequently in the FGMOS transistors. This can be achieved with no additional level shifters, while in several cases it may be harmful [57].

4. Make possible fine-tuning: Within low voltage, tuning becomes an extra issue context, where deviations are further crucial because they may bias the devices away of their proposed working area. FGMOS transistors raise the amount of degrees of freedom existing to adjust the circuits. The three basic properties of
FGMOS transistors are used to achieve low voltage and low power design. These are:

i. Flexibility: To realize both linear and complex functions within a compact and simple manner leads the subsequent circuits simplification.

ii. Controllability: As the useful threshold voltage (Vth) of each single transistor may be controlled independently according to require working range.

iii. Tunability: Since the FGMOS is a device of multiple inputs, it also can be designed to tune by putting additional inputs.

2.13 Schematic of FGSRAM Cell

The Figure 2.22 shows the schematic circuit of FGSRAM cell. In respect of conventional SRAM cell structure, the cross coupled inverter is replaced by the two inputs floating gated PMOS and NMOS as shown in Figure 2.22. In both cases 1st input of the floating gate is used for inverting operation and 2nd input is connected with bias voltage by which we can set or modify the threshold voltage of both the
PFGMOS and NFGMOS [54], [58]. The WL is the word line voltage. V1 and V2 are the two output node voltages of the inverters. The P_bias Voltage is taken as -10mV and N_bias voltage is taken as 10mV. Here, supply voltage and word line voltage are taken as 1 V and 750 mV respectively. To maintain good read and write stability proper sizing of the transistors has been done and taken as the same sizing as conventional 6T SRAM cell.

Figure 2.23 Simulation Circuit of FGSRAM Cell

Figure 2.24 DC Analysis of FGSRAM Cell
Figure 2.23 shows the simulation circuit of FGSRAM cell and the Figure 2.24 shows the DC analysis of FGSRAM cell. With the help of DC analysis, we come to know about the response of circuit when it is simply turned on without applying any signal. This illustrates about minimum power required to drive the circuit and its limits and how much current consumption is there.

Figure 2.25 Transient Analysis of FGSRAM Cell

For a circuit involving energy storing elements like capacitors or inductors, the transient time means the time during which those elements are either storing energy or releasing energy. Figure 2.25 shows the transient analysis of FGSRAM cell. From Figure 2.25, it is observed that like transient analysis of 6T SRAM cell it is also showing 100% voltage swing so that read and write operation will not be disturbed.
Figure 2.26 Power Consumption Curve of FGSRAM Cell

Figure 2.26 shows the power consumption curve of FGSRAM cell. This curve is generated at the time of calculating the overall power consumption of FGSRAM cell. The value of calculated power for read and write operation is placed in Table 2.1 to 2.4.

Figure 2.27 Write '0' Leakage Current Curve of FGSRAM Cell
Figure 2.27 and 2.28 shows the write '0' and write '1' leakage current curve respectively. This curve is generated by Cadence Virtuoso 45 nm technology at the time of calculating leakage current. After calculating the leakage current the leakage power is calculated by multiplying it with VDD and placed in Table 2.1 to 2.4.

Figure 2.28 Write '1' Leakage Current Curve of FGSRAM Cell

Figure 2.29 Write SNM Curve of FGSRAM Cell
Figure 2.30 shows the read simulation circuit of FGSRAM Cell. The sense amplifier is used here to sense the logic levels from a bit line which represents a data bit ('1' or '0') stored in a memory cell on the chip, and amplify the small voltage swing to recognizable logic levels so the data can be interpreted properly at the output terminal of the chip. For write '0' operation the voltage at BL is given as zero and BL-bar is given as one. It is given just reverse for write '1' operation. During the write and read operation the leakage is measured through the bulk/substrate of off transistor.
Figure 2.31 Read SNM Curve of FGSRAM Cell

Figure 2.29 and Figure 2.31 shows the write SNM and read SNM curve respectively. This SNM curve is plotted by Cadence tool to measure the write and read stability of the FGSRAM cell.

2.14 Simulation Results of 6T SRAM Cell and FGSRAM Cell

All the circuits of conventional 6T SRAM cell and FGSRAM cell are simulated in Cadence Virtuoso environment on 45 nm technology.

<table>
<thead>
<tr>
<th>SRAM Cell</th>
<th>Leakage Power</th>
<th>Delay</th>
<th>Power Consumption</th>
<th>SNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>6T SRAM</td>
<td>16.2 pW</td>
<td>248.2 ps</td>
<td>13.9 µW</td>
<td>84.8 mV</td>
</tr>
<tr>
<td>FGSRAM</td>
<td>3.4 nW</td>
<td>296.9 ps</td>
<td>228.5 nW</td>
<td>424.2 mV</td>
</tr>
</tbody>
</table>

Table 2.1 shows the write ‘0’ simulation output of the 6T SRAM cell and FGSRAM cell. It is observed from the Table 2.1 that 98.4% power consumption is decreased and 80% SNM is improved in FGSRAM cell than 6T SRAM cell. But leakage power and delay is increased by 99.5% and 16.4% respectively.
Table 2.2: Write ‘1’ Operation of 6T SRAM Cell and FGSRAM Cell

<table>
<thead>
<tr>
<th>SRAM Cell</th>
<th>Leakage Power</th>
<th>Delay</th>
<th>Power Consumption</th>
<th>SNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>6T SRAM</td>
<td>7.4 pW</td>
<td>139.6 ps</td>
<td>13.8 µW</td>
<td>84.8 mV</td>
</tr>
<tr>
<td>FGSRAM</td>
<td>3.7 nW</td>
<td>152.6 ps</td>
<td>228.5 nW</td>
<td>424.2 mV</td>
</tr>
</tbody>
</table>

Table 2.2 shows the write ‘1’ simulation output of 6T SRAM cell and FGSRAM cell. It is observed from the Table 2.2 that 98.3% Power Consumption is decreased and 80% SNM is improved in FGSRAM cell than 6T SRAM cell. But leakage power and delay is increased by 99.9% and 8.5% respectively.

Table 2.3: Read '0' Operation of 6T SRAM Cell and FGSRAM Cell

<table>
<thead>
<tr>
<th>SRAM Cell</th>
<th>Leakage Power</th>
<th>Delay</th>
<th>Power Consumption</th>
<th>SNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>6T SRAM</td>
<td>16.2 pW</td>
<td>249.8 ps</td>
<td>17.5 µW</td>
<td>106.06 mV</td>
</tr>
<tr>
<td>FGSRAM</td>
<td>3.4 nW</td>
<td>297.6 ps</td>
<td>4.0 µW</td>
<td>247.4 mV</td>
</tr>
</tbody>
</table>

Table 2.3 shows the read ‘0’ simulation output of 6T SRAM cell and FGSRAM cell. It is observed from the Table 2.3 that 77% power consumption is decreased and 57% SNM is improved in FGSRAM cell than 6T SRAM cell. But leakage power and delay is increased by 99.8% and 16% respectively.

Table 2.4: Read ‘1’ Operation of 6T SRAM Cell and FGSRAM Cell

<table>
<thead>
<tr>
<th>SRAM Cell</th>
<th>Leakage Power</th>
<th>Delay</th>
<th>Power Consumption</th>
<th>SNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>6T SRAM</td>
<td>7.4 pW</td>
<td>139.5 ps</td>
<td>17.5 µW</td>
<td>106.06 mV</td>
</tr>
<tr>
<td>FGSRAM</td>
<td>3.4 nW</td>
<td>153.2 ps</td>
<td>4.0 µW</td>
<td>247.4 mV</td>
</tr>
</tbody>
</table>

Table 2.4 shows the read ‘1’ simulation output of 6T SRAM cell and FGSRAM cell. It is observed from the Table 2.4 that 77% power consumption is decreased and 57% SNM is improved in FGSRAM cell than 6T SRAM cell. But leakage power and delay is increased by 99.9% and 9% respectively.
It is observed from the Table 2.1 to Table 2.4 that in FGSRAM cell the overall power consumption decreases, stability improves than 6T SRAM cell but leakage power and delay increases than 6T SRAM cell. Hence, to achieve less leakage power and to make the cell faster the low power techniques can be applied. The delay comparison of write and read operation for 6T SRAM and FGSRAM cell is given in Figure 2.32. It is observed that both for write and read operation delay in FGSRAM cell is more than 6T SRAM cell.

![Delay Comparison (ps)](image1)

**Figure 2.32 Delay Comparison of 6T SRAM and FGSRAM Cell for Write and Read Operation**

![SNM Comparison (mV)](image2)

**Figure 2.33 SNM Comparison of 6T SRAM and FGSRAM Cell for Write and Read Operation**
Figure 2.33 shows the SNM comparison of 6T SRAM and FGSRAM Cell for Write and Read Operation. It is observed that both for write and read operation SNM improves in FGSRAM cell than 6T SRAM cell.

Table 2.5: Hold Operation of 6T SRAM Cell and FGSRAM Cell

<table>
<thead>
<tr>
<th>SRAM Cell</th>
<th>Leakage Power</th>
<th>Power Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>6T SRAM</td>
<td>798.5 nW</td>
<td>798.5 nW</td>
</tr>
<tr>
<td>FGSRAM</td>
<td>24.8 nW</td>
<td>24.8 nW</td>
</tr>
</tbody>
</table>

The Table 2.5 shows the hold operation of 6T SRAM and FGSRAM cell. Here it is observed that in FGSRAM cell leakage power and power consumption is reduced by 97% and 98.4% respectively than 6T SRAM cell.

Table 2.6: Result Comparison with Existing Technique

<table>
<thead>
<tr>
<th>Operation</th>
<th>6T SRAM Ref.(58)</th>
<th>6T SRAM (Thesis work)</th>
<th>FGSRAM Ref.(58)</th>
<th>FGSRAM (Thesis work)</th>
<th>FGSLLEC SRAM (Thesis work)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating voltage</td>
<td>980 mV</td>
<td>1V</td>
<td>980 mV</td>
<td>1 V</td>
<td>1 V</td>
</tr>
<tr>
<td>Total circuit average power</td>
<td>36.23 µW</td>
<td>15.7 µW</td>
<td>6.055 µW</td>
<td>2.1 µW</td>
<td>1.8 µW</td>
</tr>
<tr>
<td>Read-write delay</td>
<td>0.78 ns</td>
<td>0.2 ns</td>
<td>0.31 ns</td>
<td>0.22 ns</td>
<td>0.19 ns</td>
</tr>
<tr>
<td>Leakage Current</td>
<td>550 nA</td>
<td>11.2 pA</td>
<td>1.89 nA</td>
<td>3.475 nA</td>
<td>1.3PA</td>
</tr>
</tbody>
</table>

2.15 Literature Survey

This section describes various attempts reported by researchers to decrease the power dissipation in SRAM or to develop low power and energy efficient SRAM. These investigations cover SRAMs operated at low voltages reducing power dissipation, SRAMs using techniques like power gating in which the circuits are switched off when they are not needed, SRAMs (drowsy) where the power supply voltage is
reduced to a lower value during standby mode and SRAMs based on adiabatic techniques. Reducing the power supply voltage decreases the leakage power exponentially and dynamic power quadratically. Although power supply voltage scaling too limits the signal swing and so decreases noise margin. Again, insistent technology scaling in the sub-100 nm area improves the sensitivity of the circuit parameters to process variation (PV). Leakage currents are mainly due to gate leakage current and sub threshold leakage current. High-K gate technology decreases the gate leakage current. Forward body biasing methods and dual Vth techniques are used to reduce sub threshold leakage current. In subthreshold SRAMs, power supply voltage is lower than the transistor threshold voltage and the sub threshold leakage current is the operating current.

Wada et al. [35] have thought about organization, configuration and process parameters for the access time formulation of on chip cache. The disadvantages of this model are, it doesn’t consider tag paths and interrelate delays. For every delay stage it considers a step input waveform, and more simplified circuit models and delay models.

Tsuguo Kobayashi [59] has suggested a current controlled sense amplifier that provides development in excess of Teruo Seki’s [60] latched sense amplifier in term of performance and power dissipation. It does not need any decoupling at its input, because of that it has high impedance input variation stage, which keep away from lower noise margin in order to the attached pass-gate transistors and it leads to a voltage fall that rejects the existing input voltage variation.

Kenneth and Amrutur [61] proposed SRAM applying half swing pulse mode gate family which does not influence recital although it is using decreased input signal swing. Since it decreases bit lines signal swing, pre-decode lines and word line, having very low power dissipation.

Amrutur and Horowitz [62] presented a model for delay, power and area of SRAM. It is assumed that wire delay changes contact speed of SRAM. They thought that with technology scaling we require to reform wires to maintain the wire delay in the
amount of the delay of the gate. This technique can be incorrect to calculate performance of the modern design of submicron cache.

Navid Azizi et.al [63] presents an asymmetrical SRAM design to minimize gate leakage. An NMOS access transistor within an asymmetric SRAM cell is introduced among the right gate of pull-down transistor storage node to disconnect the storage node from the gate of the pull-down transistor in a 6T SRAM cell. This decreases the leakage current in gate by decreasing the gate voltage of the leakage transistor pull down network considering the cell is storing '0'. Approach of asymmetric sizing has been applied to get raised write and read margins inside SRAM by Keunwoo et.al [64].

Nalam.S et.al [65] presents regarding an asymmetrical design for 6T SRAM worked with small voltage. This SRAM is presented to make use of two phase writing and variation sensing of split bit line. Nalam.S et.al in reference [66] utilize irregular sizing for 5T SRAM to attain superior read stability than 6T.

Behnam Amelifard [67] proposed a method of dual Vth SRAM design for 6T SRAM to decrease leakage power dissipation though retaining their performance. Also for great on-chip cache in 180nm technology they compare different dual Vth power and cell performance of design options.

S. Ohbayashi et.al [68] presents a design of 65 nm SoC embedded 6T SRAM for making with write and read cell circuits of stabilizing. Here every word line is attached to a number of normally-on replica access transistors (RATs) that have similar layout structure as pass transistor of SRAM cell and minimizes the word line level compared to VDD. The word line level of read support circuit is managed by the access transistors. Therefore, the selected word line level changes through VDD and threshold voltage of access transistor. During the write cycle, write assist circuit (WAC) enables switching and connecting the columns in such a way that reduced voltage is obtained by capacitive voltage division. Hence, this design does not require two power sources.

Makoto Yabuuchil et.al [69] came out with 45 nm embedded SRAM with low standby-power and better immunity aligned with temperature and process variations.
The circuit designed by them overcomes the reduction of word line voltage at -40 degree centigrade within the fast NMOS and slow PMOS (FS) condition which is owing to dependence of temperature of mobility and threshold voltage of SRAM logic circuit and access transistor. It also overcomes the decrease in the word line voltage at slow NMOS and slow PMOS (SS) state. The replica access transistors are set up to WL source driver and extra passive resistance elements realized making use of N+ poly-silicon gate thus the word line voltage level is controlled as it depends on the ratio between passive resistance and access transistor. Since, the gate length in the SS situation becomes longer owing to process variations a lower resistance (R) results on slow process corner. Resistive element is less sensitive when compared to that of MOS transistors. Hence, temperature dependence on WL is suppressed. A divided memory cell array scheme is proposed by the authors to lower the supply voltage during write operation. This is done by reducing the wiring capacitance. The SNM of 0.327 μm² cell without read assist circuit is reported to be 150 mV and with read assist circuit is 214 mV. The SNM of the 0.245 μm² cell is reported to be 208 mV making use of read assist circuit.

Baker Mohammad et.al [70] proposed a circuit to achieve better noise margin of the 6T SRAM cell by decreasing the effect of the cell parametric variation in the low voltage operation method. This technique raises the write margin and SRAM SNM applying the particular voltage supply and through least impact to area of the chip timing and complexity. The method set familiar through the authors support both the corner detection on-chip to become accustomed with the SRAM actions to silicon and software controllability to trade off yield, performance and power. Delay elements are used to adjust the new voltage level value. Hence, its level and granularity is maximised by the speed of the delay element. But the memory access speed will be decreased to reduce of the WL voltage. However, the authors have proposed that the impact of timing can be decreased if the control signal permits the decreased control of voltage swing just on the fast corners and the SNM is largely be fond of to change the cell. Simulation results shows that in the case study using 45nm data the VDD min is reduced from 1 V to 0.8 V which results in active power of 36%.
Tadayoshi Enomoto et. al [71] presented a low leakage power 180 nm CMOS SRAM in which the power gating technique is employed with the help of a personal controllable voltage level circuit with which the ground level and the supply voltage level is controlled through active and standby mode. The effects represented shows that the evaluated standby leakage power of the 1K-b SRAM memory cell arrangement significantly reduced to 5.4% that of the conventional SRAM memory cell block at 1.8 V supply voltage while the degradation of speed and area overhead were insignificant and the write operating margin was raised.

Debasis Mukherjee and Hemanta Kr. Mondal [72] explained calculation of static noise margin, read noise margin and write noise margin of an SRAM cell by the butterfly curve method. Here, they explained how the memory cell stability depends on the values of the cell-ratio and pull-up ratio.

Byung-Do Yang [73] proposed the energy reusing technique both during writing and reading of 6T SRAM different that of [64] in which of the reusing energy is done just in writing operation. The author states no degradation of static noise margin (SNM), keeping of write power of 84% and read power of 17%. But the operation speed is limited to 145 MHz. The SRAM cell has an area overhead of start up circuit of precharge-recycling.

Anh-Tuan Do et. al [74] proposed a novel current-mode sense amplifier. For both local and global sensing stages it broadly uses the cross-coupled inverters, and also to achieve high-speed and very low-power properties all together. The design suggested by authors may work with a highest frequency of 1.25 GHZ at 1 V supply voltage and of 0.2 V minimum supply voltage. These qualities of the offered circuit build it a wise selection for modern systems of more complexity where power consumption and reliability is of main fear.

Rakesh Chandankhede [75] has presented a combined current controlled sense amplifier with a latch that provides advance in performance and power consumption. Here, when enable signal (EN) is small, the differential voltage (logic) on the BL and BL_Bar develops on the output of the latch though it does not increase because the
pull-down transistor is cut off and as a result equal logic holds at rest till EN goes high. At this time low voltage line releases to ground as soon as EN signal goes high.

### 2.16 Analytical Summary of Relevant Works

<table>
<thead>
<tr>
<th>References</th>
<th>Proposed work</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>Organization, configuration and process parameters for the access time formulation of on chip cache.</td>
<td>For every delay stage it considers a step input waveform, and more simplified circuit models and delay models.</td>
</tr>
<tr>
<td>60</td>
<td>Proposed latched sense amplifier in terms of performance and power dissipation.</td>
<td>Keep away from lower noise margin in order to the attached pass-gate transistors and it leads to a voltage fall that rejects the existing input voltage variation.</td>
</tr>
<tr>
<td>61</td>
<td>SRAM applying half swing pulse mode gate family which does not influence recital although it is using decreased input signal swing.</td>
<td>It decreases bit lines signal swing, pre-decode lines and word line, having very low power dissipation.</td>
</tr>
<tr>
<td>62</td>
<td>Presented a model for delay, power and area of SRAM. They thought that with technology scaling we require to reform wires to maintain the wire delay in the amount of the delay of the gate.</td>
<td>This technique can be incorrect to calculate performance of the modern design of submicron cache.</td>
</tr>
<tr>
<td>63</td>
<td>Presents an asymmetrical SRAM design to minimize gate leakage. An NMOS access transistor within an asymmetric SRAM cell is introduced among the right gate of pull-down transistor storage node to disconnect the storage node from the gate of the pull-down transistor in a 6T SRAM cell.</td>
<td>This decreases the leakage current in gate by decreasing the gate voltage of the leakage transistor pull down network considering the cell is storing '0'.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>67</td>
<td>Proposed a method of dual Vth SRAM design for 6T SRAM to decrease leakage power dissipation though retaining their performance.</td>
<td>This work compares different dual Vth power and cell performance of design options.</td>
</tr>
<tr>
<td>68</td>
<td>Presents a design of 65 nm SoC embedded 6T SRAM for making with write and read cell circuits of stabilizing.</td>
<td>During the write cycle, write assist circuit (WAC) enables switching and connecting the columns in such a way that reduced voltage is obtained by capacitive voltage division. Hence, this design does not require two power sources.</td>
</tr>
<tr>
<td>69</td>
<td>Came out with 45nm embedded SRAM with low standby-power and better immunity aligned with temperature and process variations.</td>
<td>The SNM of the 0.245 μm² cell is reported to be 208 mV making use of read assist circuit.</td>
</tr>
<tr>
<td>Page</td>
<td>Proposed a circuit to achieve better noise margin of the 6T SRAM cell by decreasing the effect of the cell parametric variation in the low voltage operation method.</td>
<td>Simulation results shows that in the case study using 45 nm data the VDD is reduced from 1 V to 0.8 V which results in active power of 36%.</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>72</td>
<td>To calculate static noise margin, read noise margin and write noise margin of an SRAM cell by the butterfly curve method.</td>
<td>Explained how the memory cell stability depends on the values of the cell-ratio and pull-up ratio.</td>
</tr>
<tr>
<td>74</td>
<td>Proposed a novel current-mode sense amplifier.</td>
<td>This provides a wise selection for modern systems of more complexity where power consumption and reliability is of major concern.</td>
</tr>
<tr>
<td>75</td>
<td>Has presented a combined current controlled sense amplifier with a latch that provides advance in performance and power consumption.</td>
<td>When enable signal (EN) is small, the differential voltage (logic) on the BL and BL_Bar develops on the output of the latch though it does not increase because the pull-down transistor is cut off and as a result equal logic holds at rest till EN goes high.</td>
</tr>
</tbody>
</table>
2.17 Existing Leakage Reduction Technique

To decrease the leakage current, the designers used many low power techniques. Here, Some existing leakage reduction techniques are presented.

2.17.1 Sleepy Transistor Technique

Figure 2.34 shows the sleepy transistor technique in which the power rails or ground are cut off by sleepy transistors [76]. Motoh et al. suggested a multi threshold voltage CMOS (MTCMOS) technique that adds high threshold voltage sleepy transistors among VDD and pull-up networks and among Ground and pull-down networks [41], [77]. The sleepy transistors are switched off as the logic circuits are in the standby mode. With the help of the sleep transistors in the standby mode, dramatically leakage power is decreased by cutting off the logic circuits. Throughout the sleepy mode as pull down and pull-up network is hanging, it drops its state. Therefore, the sleepy transistors inserted additionally will raise delay and area. Also, this technique is known for power gating [78].
2.17.2 Forced Stack Technique

Figure 2.35  (a) Forced stack Technique  (b) Forced stack inverter

Figure 2.35 shows the forced stack technique. In this method, each transistor in the network is replaced with two transistors in series having half the width of original transistor. Here, the duplicated transistors make small reverse bias among the source and the gate as together the transistors are switched off [79], [80]. It achieves significant current minimization with a view to the dependence of subthreshold current on gate bias. In Figure 2.35 (b) When A = 0, then both M1 and M2 transistors are turned off. So, the voltage of node VX is more than Gnd due to internal resistance of M2. Hence, M1 has a negative drain-source voltage which degrades the DIBL effect. It overcomes the disadvantages by the sleepy technique with retaining the state and it takes additional wake-up time.
2.17.3 Sleepy Stack Technique

Sleepy stack technique is shown in Figure 2.36. In this technique both sleepy and stack approaches are combined [76], [80-81]. It divides the transistors equally and then the sleepy transistors are inserted in parallel to the divided transistors. The sleepy transistors of the sleepy stack technique work alike to the sleepy transistors applied in the sleepy transistor technique. Sleepy transistors are switched off and suppressing the leakage current though saving the state. So, in the active mode delay is reduced. This technique [82] may potentially decreases the circuit delay in two ways. First, since the sleep transistors are on for all time during active mode, the sleepy stack technique attains faster switching time than the forced stack technique and secondly by using the multi threshold technique.
2.17.4 Lector Technique

Figure 2.37 shows the lector technique. In this technique, between pull-down and pull-up network two leakage control transistors (NMOS and PMOS) are used [83], [84], [80]. These two transistors are connected in such a way that for any input arrangement one is near cut off for all time. Thus, there creates a high resistance path among the supply rails and the ground. So, the leakage power is reduced. The concept of lector technique is that, “in a path among supply voltage and ground a state with more than one transistor off is less leaky than a state with only a single transistor off among any supply and ground path”. Therefore, among the supply voltages through the lector circuit the lector technique leads to a current limited resistive path to decrease the leakage power dissipation.
2.17.5 ABC-MTCMOS Technique

Auto back gated multi threshold voltage (ABC-MTCMOS) technique is shown in Figure 2.38. It makes use of reverse bias to decrease the leakage power [68]. This is a state protecting technique which protects the circuit logic state. Reverse body bias (RBB) decreases the leakage power with increase in threshold voltage without losing the state of the circuit. A MTCMOS method presents low leakage and high recital function by using high speed, low threshold voltage (lVth) transistors in active mode and low leakage, high threshold voltage transistors in sleep mode which decreases the leakage power dissipation of an SRAM cell [85-86].
2.17.6 Sleepy Keeper Technique

Figure 2.39 shows the sleepy keeper technique. The advanced version of sleep technique is sleepy keeper technique [87]. In this method, an extra high threshold PMOS transistor is connected in parallel with sleepy transistor (PMOS) and an extra hVth NMOS transistor is connected in parallel with sleepy NMOS transistor. In standby mode, sleepy transistors are in switched off state. Therefore, while sleep signal is asserted, the high threshold voltage NMOS transistor tied in parallel with the sleep transistor (PMOS) is the only source of VDD to the pull up network and the hVth transistor (NMOS) tied in parallel with the sleep transistor (NMOS) gives the path to attach the pull down network with the ground. Thus the main advantage of this technique is that it decreases the leakage power by maintaining the circuit logic state.
2.17.7 Zigzag Approach

Zigzag technique is shown in Figure 2.40. This method reduces the area overhead caused by the extra sleep transistor used in sleep technique [88]. By putting the alternate sleep transistor, this overhead can be decreased by choosing the particular preselected input vector. During sleepy mode, logic is ‘0’ input and every logic input turn round its state and the output is '1'. Hence, the zigzag method makes use of few sleep transistors than the sleepy technique.

2.17.8 Input Vector Control (IVC) Technique

Amount of leakage power has strong dependence on the input arrangements. To control the leakage power this technique utilizes the input vector [89]. The input vector that provides the low leakage is selected by the automation system and in sleep mode applied to the circuit.
2.17.9 Galeor Technique

Figure 2.41 explains the galeor technique. This technique decreases the leakage current passing through the circuits [90]. Two gated leakage transistors are introduced between the pull up network and pull down network of the existing circuit such that extra inserted transistors gate are tied to their particular drain regions. The output voltage swing according to the threshold failure caused by extra MOS transistors is reduced by this technique.

After literature survey and analysis of simulation results it is observed that in FGSRAM cell power consumption is reduced and SNM improves but leakage current and delay improves than 6T SRAM cell. As our main focus is leakage reduction so in the next chapter low power technique sleepy is applied in the FGSRAM cell.