

DESIGN AND STABILITY ANALYSIS OF A HIGH-TEMPERATURE SRAM

A Thesis

Presented to

The Graduate Faculty of The University of Akron

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

Tanvir Tanvir

December, 2012

DESIGN AND STABILITY ANALYSIS OF A HIGH-TEMPERATURE SRAM

Tanvir Tanvir

Thesis

Approved:

---

Co-Advisor  
Dr. Joan Carletta

---

Co-Advisor  
Dr. Robert J. Veillette

---

Committee Member  
Dr. Tom T. Hartley

---

Department Chair  
Dr. Alex De Abreu Garcia

Accepted:

---

Dean of the College  
Dr. George K. Haritos

---

Dean of the Graduate School  
Dr. George R. Newkome

---

Date

## ABSTRACT

This work discusses the design of a static random access memory (SRAM) capable of operating over a temperature range of 27°C to 120°C. A standard six-transistor SRAM cell is used to implement a 1Kword SRAM where each word is sixteen bits. The conventional array architecture is used. The control circuit is implemented using an asynchronous approach. Careful consideration is given to the stability of the SRAM cell and the performance of the memory. The stability margins considered include the read static noise margin, the hold static noise margin and the write static noise margin. These stability margins are discussed in detail and the effect of temperature on stability is discussed. Four designs are included: the Baseline Design, the Modified Design I, the Modified Design II and the Modified Design III. The device geometries of the SRAM cell are adjusted to have acceptable read and write static noise margins at 27°C and 120°C. The most stable of the designs shows a 10% increase in read stability margin with a 2.6% increase in read access time relative to the baseline design. It is observed that the stability was improved at the cost of slightly reduced speed. The designed SRAM has an acceptable stability over temperature.

## DEDICATION

I would like to dedicate this thesis to Mom, Pinku and Suhail.

## ACKNOWLEDGEMENTS

It is a pleasure to me to thank my committee members Dr. Joan E. Carletta, Dr. Robert Veillette and Dr. Tom Hartley for their encouragement and motivation throughout my Master's program. I owe my deepest gratitude to Dr. Joan E Carletta and Dr. Robert Veillette for all their support and patience that motivated me to finish my research and my thesis writing.

I would like to thank Dr. Alex De Abreu-Garcia, Faculty and students of Electrical and Computer Engineering Department for their encouragement and support throughout my degree program. I would specially thank Mrs. Gay Boden for all her extended help right from the first day.

I owe my heartfelt regards to my mom, my sister and my husband who have been my constant source of inspiration, motivation and encouragement.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER	
I INTRODUCTION .....	1
II ARCHITECTURAL OVERVIEW .....	4
2.1 SRAM Architecture .....	5
2.1.1 SRAM Array .....	5
2.1.2 SRAM Cell .....	7
2.1.3 Row Decoder, Column Decoder and Column Multiplexer .....	7
2.1.4 Control Circuit .....	8
2.1.5 Pre-Charge Circuitry .....	14
2.1.6 I/O and Sensing Block .....	14
2.2 Noise Margins .....	16
2.2.1 Hold Static Noise Margin .....	16
2.2.2 Read Static Noise Margin .....	17
2.2.3 Write Static Noise Margin .....	19
III SRAM DESIGN .....	24
3.1 SRAM Array .....	24

3.2 SRAM Cell.....	26
3.3 6-to-64 Row Decoder.....	34
3.4 4-to-16 Column Decoder .....	37
3.5 16-bit 16-to-1 Column Multiplexer .....	40
3.6 Pre-charge Circuitry.....	40
3.7 Control Circuit .....	42
3.6 I/O and Sensing Block .....	44
IV HIGH-TEMPERATURE OPERATION .....	49
4.1 Operation of a Transistor at High Temperature.....	49
4.2 High-Temperature Operation of SRAM .....	52
4.2.1 Baseline Design .....	53
4.2.2 Modified Design I.....	60
4.2.3 Modified Design II.....	62
4.2.4 Modified Design III .....	67
4.3 Trade-offs among Stability Margins and Access Times.....	70
V CONCLUSIONS AND FUTURE WORK .....	78
REFERENCES .....	81

## LIST OF TABLES

Table	Page
2.1 Truth table for SRAM modes of operation .....	9
3.1 Transistor channel dimensions for the SRAM cell for the Baseline Design .....	32
4.1 Transistor channel dimensions for the SRAM cell Modified Design I .....	60
4.2 Transistor channel dimensions for the SRAM cell Modified Design II .....	64
4.3 Transistor channel dimensions for the SRAM cell Modified Design III.....	67
4.4 Static noise margins for the SRAM cell for the four designs at 27°C and 120°C.....	71
4.5 SRAM access times for the four designs at 27°C and 120°C .....	71



## LIST OF FIGURES

Figure	Page
2.1 A conventional array architecture of an SRAM with $2^{P+Q}$ words of N bits each .....	6
2.2 A CMOS 6T SRAM memory cell .....	6
2.3 Timing diagram for a typical read cycle .....	12
2.4 Timing diagram for a typical write operation .....	13
2.5 Simple pre-charge circuit .....	14
2.6 I/O and sensing block circuitry associated with one data bit .....	15
2.7 Circuit to find the hold static noise margin (HSNM) .....	16
2.8 VTCs of the inverters showing the HSNM .....	18
2.9 Circuit to find the read static noise margin (RSNM) .....	18
2.10 VTCs of loaded inverters showing the RSNM .....	20
2.11 Circuit to find the write static noise margin (WSNM) .....	20
2.12 VTCs of loaded inverters showing the WSNM .....	21
2.13 Comparison of the read static noise margin and hold static noise margin .....	22
3.1 Block Diagram of 1K word SRAM .....	25
3.2 Organization of bits in a row .....	26
3.3 Circuitry in one column of the SRAM .....	27
3.4 SRAM cell .....	28
3.5 VTCs showing the hold static noise margin for the baseline design .....	31

3.6 VTCs showing the read static noise margin for the baseline design .....	32
3.7 VTCs showing the write static noise margin for the baseline design .....	33
3.8 Transient simulation showing the rise in voltage at node Q during a read operation..	33
3.9 Sub-blocks of the 6-to-64 row decoder.....	35
3.10 A 2-to-4 dynamic NOR decoder .....	36
3.11 Local pre-charge driver.....	36
3.12 Gate control logic for each individual row line and column select line .....	36
3.13 Row line driver .....	37
3.14 Sub-blocks of the 4-to-16 column decoder.....	38
3.15 Column select line driver .....	39
3.16 16-to-1 Column multiplexer/demultiplexer circuitry for one bit .....	39
3.17 RC model used for calculating the pulse width of the pre-charge signal .....	41
3.18 Control circuit logic .....	42
3.19 Delay chain used in the control circuit .....	44
3.20 Pre-charge buffer .....	44
3.21 Sense enable signal driver.....	45
3.22 I/O Block.....	45
3.23 Expanded circuit diagram of the tri-state inverter driving BL or BLbar .....	46
3.24 Differential sense amplifier.....	47
4.1 NMOS transistor in saturation .....	50
4.2 Variation of the drain-to-source current through a 4.0 $\mu\text{m}$ / 0.8 $\mu\text{m}$ NMOS transistor in saturation region with temperature .....	50
4.3 NMOS transistor in cut-off .....	51
4.4 Variation of the leakage current through a 4.0 $\mu\text{m}$ / 0.8 $\mu\text{m}$ NMOS transistor in sub-	

threshold region with temperature .....	51
4.5 VTC of an inverter at low and high temperature .....	52
4.6 Transistor level diagram of a SRAM cell .....	53
4.7 VTC showing HSNM for the Baseline Design at 27°C and 120°C .....	54
4.8 VTC showing RSNM for the Baseline Design at 27°C and 120°C .....	54
4.9 VTC showing WSNM for the Baseline Design at 27°C and 120°C .....	55
4.10 Simulation results showing write and read operations for the Baseline Design at 27°C.....	56
4.11 Simulation results showing write and read operations for the Baseline Design at 120°C.....	57
4.12 VTC showing the read SNM for Modified Design I at 27°C and 120°C.....	60
4.13 VTC showing the write SNM for Modified Design I at 27°C and 120°C .....	61
4.14 Simulation results showing write and read operations for the Modified Design I at 27°C.....	62
4.15 Simulation results showing write and read operations for the modified design I at 120°C.....	63
4.16 VTC showing read SNM for Modified Design II at 27°C and 120°C .....	64
4.17 VTC showing write SNM for Modified Design II at 27°C and 120°C.....	65
4.18 Simulation results showing write and read operations for the Modified Design II at 27°C.....	65
4.19 Simulation results showing write and read operations for the Modified Design II at 120°C.....	66
4.20 VTC showing the read SNM for Modified Design III at 27°C and 120°C.....	68
4.21 VTC showing the write SNM for Modified Design III at 27°C and 120°C.....	68
4.22 Simulation results showing write and read operations for the Modified Design III	

at 27°C.....	69
4.23 Simulation results showing write and read operations for the Modified Design III at 120°C.....	70
4.24 Variation of RSNM with cell ratio at 27°C and 120°C.....	72
4.25 Variation of WSNM with pull-up ratio at 27°C and 120°C.....	73
4.26 Variation of write time with pull-up ratio at 27°C and 120°C.....	73
4.27 Variation of read access time with cell ratio at 27°C and 120°C.....	74
4.28 Variation of read access time with RSNM at 27°C and 120°C.....	75
4.29 Variation of write time with RSNM at 27°C and 120°C.....	76
4.30 Variation of WSNM with RSNM at 27°C and 120°C.....	76

## CHAPTER I

### INTRODUCTION

Nowadays, quite a large number of high-temperature application-specific integrated circuits (ASICs) need embedded, on-chip memory to store program code or sensor data. Static random access memory (SRAM) can be used for this purpose. An ideal memory is stable, area efficient and fast, with low power consumption. A memory that can work in high-temperature environments, such as under the hood of an automobile, in an aircraft engine or down an oil well [8], requires both a special semiconductor process and a special design.

Silicon-on-Insulator (SOI) CMOS processes are suitable for the design of high-temperature circuits that can operate up to 250°C [11][8]. The use of an insulating substrate in an SOI process reduces the area of the MOSFET p-n junctions, which results in reduced parasitic capacitances and lower sub-threshold leakage currents. More significant for high-temperature applications, the use of an SOI process eliminates leakage currents to the substrate which increase greatly with temperature for bulk CMOS processes, and restrict the operating temperature range of the circuits [9].

At high temperatures, a memory will operate with increased leakage currents, decreased stability and reduced speed, regardless of the process used. The decrease in

stability is of particular concern, as it results in an increased chance of data corruption. Therefore, a high-temperature memory requires a robust design to ensure stability [5].

Implementing a SRAM used in a high-temperature micro-controller system for the application of down hole oil drilling is presented in [3]. The design presented in [3] was synchronous with a maximum operating frequency of 8.0 MHz and was aimed to operate at 275°C in a sub-micron SOI process. A methodology that could reduce the design time of SRAMs by partitioning into several critical sub-modules and characterizing them for performance was proposed in [3].

Design and implementation of a high-temperature electronic circuit, which consists of a processor core, on-chip cache memories and corresponding peripherals was presented in [5]. The design presented in [5] was synchronous with a maximum operating frequency of 100.0 MHz and was meant to operate at 200°C in a 180nm SOI process. The emphasis of the work in [5] is to design reliable high-temperature memory blocks at 200°C using SOI technology.

The effect of process and temperature variation along with transistor aging on the reliability of the 6T SRAM cell in a 32nm CMOS technology is analyzed in [15]. The static noise margin, write margin, access time and leakage are the performance metrics that were considered for the analysis. The emphasis of this work is to study the effect of supply voltage variation on the SRAM cell, along with other effects such as variations over temperature and process.

This thesis considers the design of an on-chip SRAM for a mixed-signal application-specific integrated circuit in an SOI CMOS process. High-temperature

synchronous SRAM designs using deep sub-micron processes intended for digital designs are presented in [5] and [15]. These memories could not be used as a part of sensor interfaces where analog circuitry is also required. For a mixed-signal design that includes both digital and analog designs on a single semiconductor die, the process used in this thesis and [3] would be more appropriate [12]. The work in [3] is focused on reduction of design time, and does not show how temperature and stability should be taken into account when designing the SRAM cell.

This thesis develops an SRAM design and verifies its operation over temperature through simulations using Cadence® Virtuoso® schematic editor and Cadence® Virtuoso® Spectre® circuit simulator. The SRAM design developed here uses an asynchronous approach; this avoids the problems associated with clock distribution [13] that may be experienced in a synchronous SRAM. The circuit is shown to operate correctly over a temperature range of 27°C to 120°C; the process is capable of operating at higher temperatures, but the MOSFET models are not validated above 120°C. A baseline SRAM was designed that was suitable at room temperature. Analysis verified reduced noise margins at high temperatures; the read static noise margin was particularly affected. Three redesigns were done to improve the noise margins at the cost of increased access times. In these redesigns, the size of the NMOS pass transistors in the six-transistor SRAM cells used was decreased, improving the RSNM both at room temperature and at high temperature. The simulations show that stability can be achieved over temperature at the expense of increased access times and decreased write static noise margin. The most stable of the designs has an RSNM of 0.52 V at 27°C and 0.42 V at

120°C. The read access time is 4.8 ns at 27°C and 6.8 ns at 120°C. The WSNM is 0.72 V at 27°C and 0.65 V at 120°C. The write time is 10.0 ns at 27°C and 15.5 ns at 120°C.

The main contribution is a study of the trade-offs in the design of a high-temperature asynchronous SRAM in an SOI CMOS process suitable for a mixed-signal application-specific integrated circuit such as a sensor interface with a digital back end. It is observed that the WSNM, the read access time and the write times are all degraded when the RSNM is improved. There is a direct trade-off between RSNM and WSNM, both at 27°C and at 120°C, throughout the range of designs considered.

The thesis is divided into the following chapters.

- Chapter II provides an architectural overview for a static random access memory and describes the function of all the blocks needed to implement it. Standard definitions for noise margins are also reported.
- Chapter III presents a baseline design for a six-transistor SRAM cell. It reports the noise margins based on simulations at room temperature. The chapter also includes the design details for the rest of the components needed for a 16 Kbit SRAM, organized as 1024 16-bit words.
- Chapter IV presents the high-temperature operation of the baseline SRAM cell. This chapter includes the design considerations for a high-temperature SRAM cell. Three modified designs are presented to show the tradeoff between stability and access time.
- Chapter V draws conclusions and makes recommendations for possible future work.



## CHAPTER II

### ARCHITECTURAL OVERVIEW

This chapter presents a standard architecture for a static random access memory (SRAM), and describes the function of all the blocks needed to implement the memory. It also gives standard definitions of noise margins, which are of primary importance in the SRAM design.

#### 2.1 SRAM Architecture

A conventional array architecture for an SRAM is presented in Figure 2.1. This consists of an array of memory cells, along with the circuitry needed to access the cells associated with any word of memory, based on the input address lines. A set of input control signals is used to control the access between the memory and the bidirectional data bus for the read and write operations. The following subsections describe all the components of the SRAM.

##### 2.1.1 SRAM Array

The cells of an SRAM are arranged in the form of a two-dimensional array with  $2^P$  rows and  $N \cdot 2^Q$  columns. Each cell holds a single bit of data. Each row in the array is organized into words, with each word having  $N$  bits. Each cell in the array is connected to one of the  $2^P$  row lines (RL), and to one of the  $N \cdot 2^Q$  pairs of column lines, also called the

*bit lines* (BL). A word occupies N cells in the array, all from the same row. A cell is selected by activating a particular row line and accessing a particular pair of bit lines.

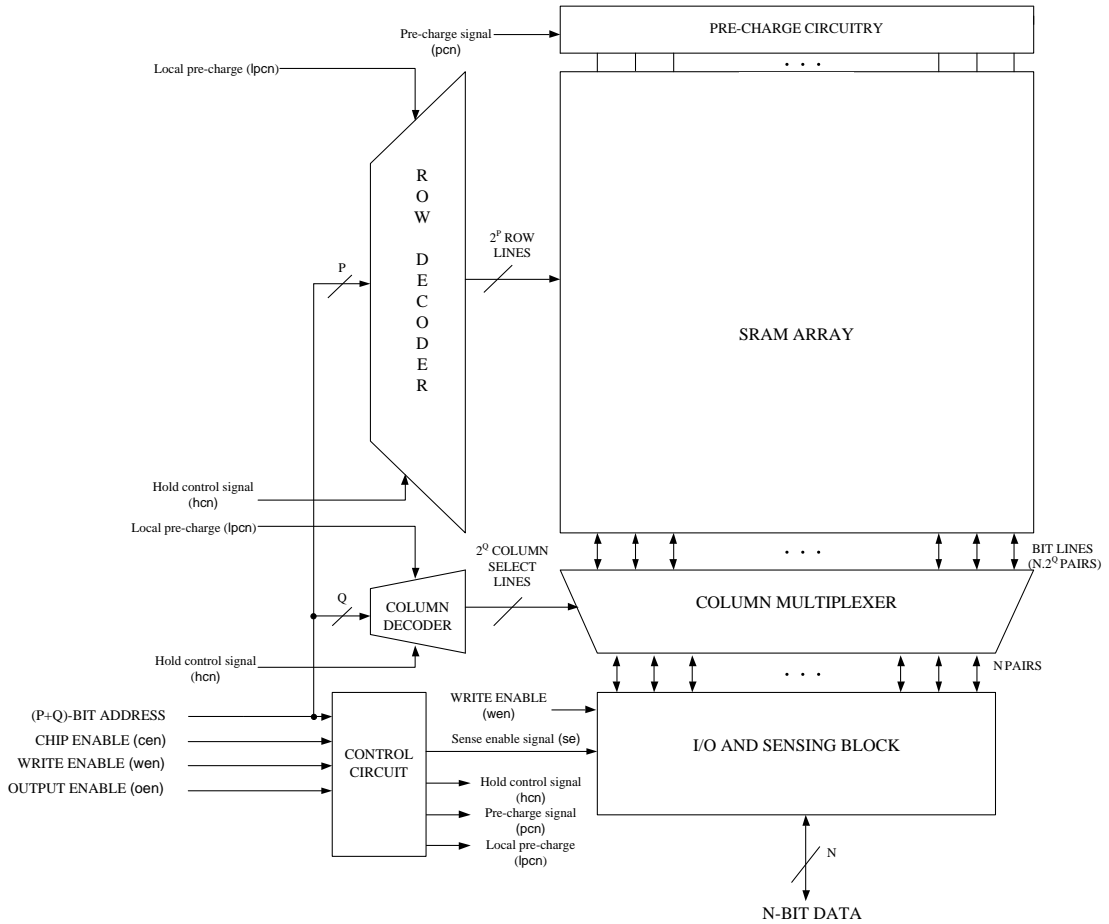


Figure 2.1 A conventional array architecture of an SRAM with  $2^{P+Q}$  words of N bits each

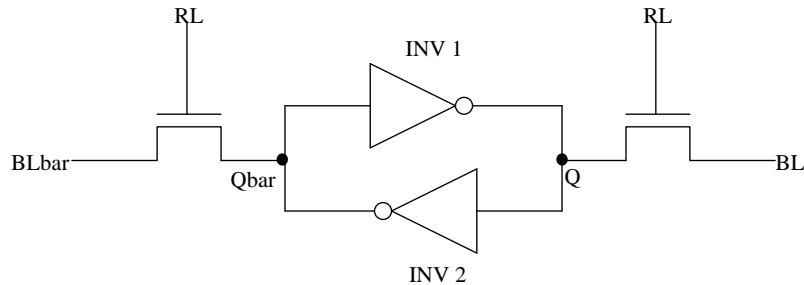


Figure 2.2 A CMOS 6T SRAM memory cell

### 2.1.2 SRAM Cell

The fundamental building block in an SRAM array is the SRAM cell which stores a single bit of data. Figure 2.2 shows a standard six-transistor (6T) SRAM cell [17]. The SRAM cell consists of two cross-coupled static CMOS inverters and two NMOS pass transistors. The circuit has two stable states, corresponding to storage of a “0” or a “1”. In the absence of any external influence, the circuit remains in a single state and holds that state as long as power is supplied. The state of the SRAM cell is changed or accessed by turning the pass transistors ON by setting the row line RL high, thus connecting the bit lines BL and BLbar with the cell. The NMOS pass transistors allow bidirectional flow of current between the bit lines and the cell.

### 2.1.3 Row Decoder, Column Decoder and Column Multiplexer

The address needed to access a word has  $(P+Q)$  bits; the  $P$  most significant bits of the address, referred to as the row address, indicate the row in which the word resides. The remaining  $Q$  bits, referred to as the column address, indicate a particular one of the  $2^Q$  words in the row to be accessed. The purpose of the row decoder is to activate one of the  $2^P$  row lines, depending on the  $P$ -bit row address. The purpose of the column decoder is to open up access to the  $N$  pairs of bit lines corresponding to a particular word, depending on the  $Q$ -bit column address. One of the  $2^Q$  column select lines is pulled high by the column decoder; this in turn activates the corresponding paths through the column multiplexer. The column multiplexer is bidirectional, and directs data to and from the bit lines. Thus, the row decoder, the column decoder and the column multiplexer work

together to select and access one N-bit word of the memory for a read or a write operation.

The row decoder and the column decoder can be implemented as dynamic NOR decoders. A dynamic NOR decoder requires a local pre-charge signal ( $lpcn$ ) that pulls all the decoder outputs high during a pre-charge phase; then, during an evaluation phase, depending on the address, all the outputs except one are pulled low. The local pre-charge signal ( $lpcn$ ) generated by the control circuit is also used to gate the row lines so that an active row line signal is passed to the array only during the evaluation phase, when only the one row line corresponding to the row being addressed is active. A hold control signal ( $hcn$ ) is also used that ensures that none of the row lines or the column select lines is active during a hold operation.

#### 2.1.4 Control Circuit

The memory read and write operations consist of sequences of steps. These sequences are governed by the control circuit. The control circuit is designed using an asynchronous approach. The control circuit provides the signals needed to control the row and column decoders, the pre-charge circuitry and the I/O and sensing block. The control circuit takes the chip enable signal ( $cen$ ), the write enable signal ( $wen$ ), the output enable signal ( $oen$ ) and the address as inputs, and generates the sense enable signal ( $se$ ), the pre-charge signal ( $pcn$ ), the local pre-charge signal ( $lpcn$ ) and the hold control signal ( $hcn$ ) with the required timing.

An SRAM has three modes of operation: Hold, Read and Write. The mode of operation is determined by the three external inputs chip enable signal ( $cen$ ), output

Table 2.1 Truth table for SRAM modes of operation

cen	wen	oen	Mode
0	0	0	Not Allowed
0	0	1	Write
0	1	0	Read
0	1	1	Hold
1	X	X	Hold

enable signal (**oen**) and write enable signal (**wen**) according to the truth table given in Table 2.1. Whenever **cen** is inactive, the SRAM is in a hold mode. The SRAM will also be in a hold mode if **cen** is active but **oen** and **wen** are both high (inactive). For a read or a write operation, it is required that the chip be selected, i.e., the chip enable signal (**cen**) must be low (active). For a write operation, the write enable signal (**wen**) is low (active) and the output enable signal (**oen**) is high (inactive). For a read operation, the output enable signal (**oen**) is low (active) and the write enable signal (**wen**) is high (inactive). It is required that the output enable signal (**oen**) and write enable signal (**wen**) should not be asserted simultaneously, i.e., both being low (active).

Whenever **cen** is inactive, the SRAM is in a hold mode, and the control circuit produces inactive signals for its outputs; that is, the local pre-charge signal (**lpcn**) and the pre-charge signal (**pcn**) are high (inactive) and the sense enable signal (**se**) is low

(inactive). The hold control signal ( $hcn$ ) may be active or inactive depending on the output enable signal ( $oen$ ) and the write enable signal ( $wen$ ), but none of the row lines or the column select lines are activated because of an inactive  $cen$ .

When  $cen$  is low (active) and  $wen$  and  $oen$  are both high (inactive), the SRAM is in a hold mode and the control circuit generates a high (inactive)  $pcn$  and a high (active)  $se$ . The  $lpcn$  may be low (active) if there is an address change, but none of the row lines or column select lines are activated because the hold control signal ( $hcn$ ) is low (active) and the row and the column driver outputs are disabled. As  $wen$  is inactive, the input block disconnects the SRAM array from the bidirectional data bus and no new read data is put on the bus as the word lines are forced low (inactive). In such a situation, the SRAM cell continues to hold the previously written data.

For a read operation, the bit lines  $BL$  and  $BLbar$  are pre-charged to  $V_{DD}$ . The row decoder and the column decoder activate a row line and a column select line, thus selecting  $N$  SRAM cells in the array. One of the bit lines of each selected column is pulled low and the other stays high. Then, the sense amplifiers are turned ON which amplifies the difference in potential between the bit lines, and the I/O block puts the  $N$ -bit data on the data bus.

The sequence for a read operation is initiated under either of two conditions. The first condition is when the output enable signal ( $oen$ ) transitions from high to low while the write enable signal ( $wen$ ) is high (inactive). The second condition is when there is a transition on the address bus while the output enable signal ( $oen$ ) is low (active) and the write enable signal ( $wen$ ) is high (inactive). For either condition, an inactive write enable

signal (**wen**) and an active output enable signal (**oen**) is the combination that distinguishes a read operation.

A general timing diagram for a read operation is shown in Figure 2.3. When **oen** transitions from high to low, the control circuit generates a high (inactive) **hcn**, a low (active) pulse on **lpcn** and a low (active) pulse on **pcn**. The **lpcn** pulse is used to activate the pre-charge circuitry that pulls up the row and column decoder outputs. The **pcn** pulse is used to activate the pre-charge circuitry that pulls the bit lines to  $V_{DD}$ . Once the **lpcn** pulse goes high, indicating that the pre-charge is done, the row and column decoders select a word by activating the corresponding row line and opening access to the corresponding pairs of bit lines. In response to the **pcn** transitioning back to its high (inactive) level, the control circuit also generates a high (active) **se**, which is used by the I/O and sensing block. This sequence is shown as Read Cycle 1 in Figure 2.3.

A read cycle is also initiated if there is a transition on the address bus while **oen** is low (active) and **wen** is high (inactive). A read cycle initiated by an address bus transition is shown in Figure 2.3 as Read Cycle 2.

For a write operation, the I/O block puts the data to be written and its complement on the bit lines, **BL** and **BLbar**. The row decoder and the column decoder activate a row line and a column select line, thus selecting **N** SRAM cells in the array. Once selected, the **N**-bit data is written in the **N** SRAM cells.

The sequence for a write operation is initiated under either of two conditions. The first condition is when the write enable signal (**wen**) transitions from high to low while the output enable signal (**oen**) is high (inactive). The second condition is when there is a

transition on the address bus while the write enable signal (**wen**) is low (active) and the output enable signal (**oen**) is high (inactive). For either condition, an inactive output enable signal (**oen**) and an active write enable signal (**wen**) is the combination that distinguishes a write operation.

A general timing diagram for a write operation is shown in Figure 2.4. When **wen** transitions from high to low, the control circuit generates a low (active) pulse on **lpcn** which pre-charges the outputs of the row and column decoders. Once the pre-charge is done, i.e., when **lpcn** goes high, the row and column decoders select a word by activating

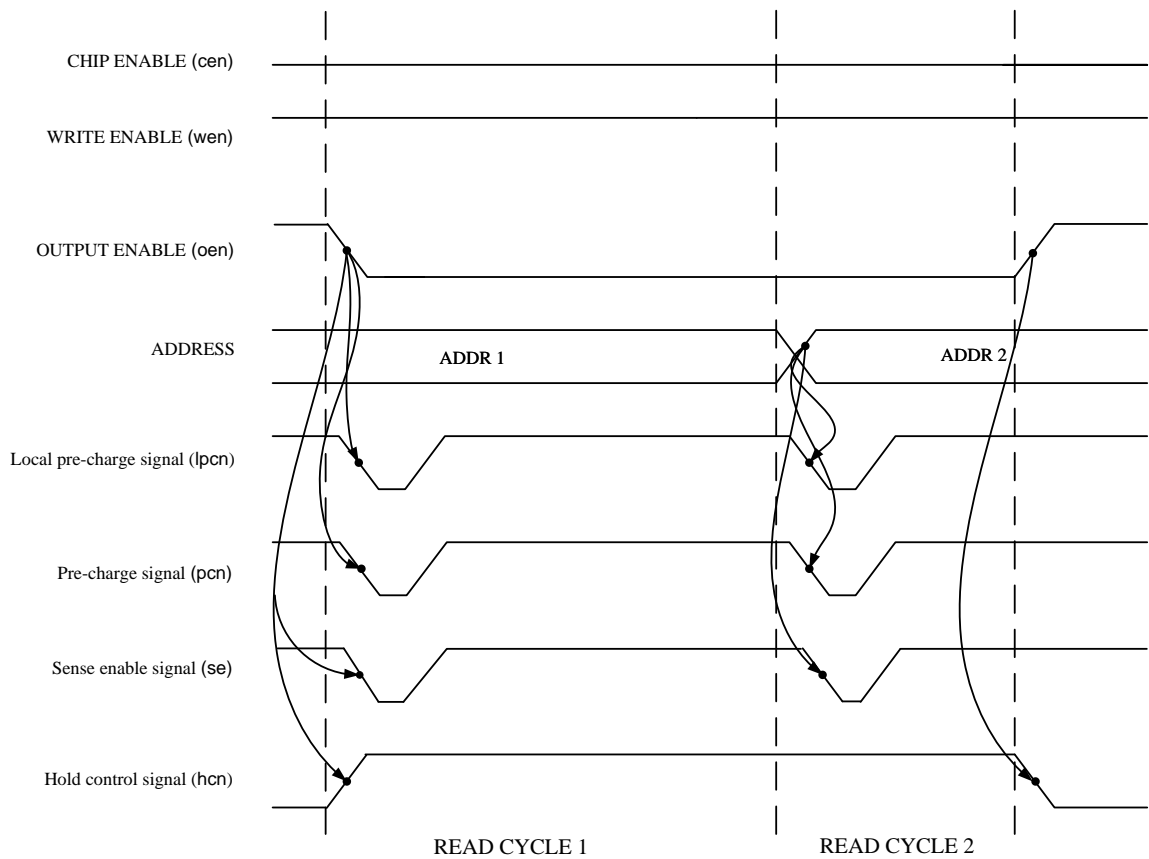


Figure 2.3 Timing diagram for a typical read cycle



the corresponding row line and opening access to the corresponding pairs of bit lines. The control circuit holds *hcn* and *pcn* high (inactive). In response to a high (inactive) *pcn*, the control circuit holds *se* low (inactive). This sequence is shown as Write Cycle 1 in Figure 2.4.

A write cycle is also initiated if there is a transition on the address bus while *wen* is low (active) and *oen* is high (inactive). The control circuit generates signals similar to Write Cycle 1. This sequence is shown as Write Cycle 2 in Figure 2.4.

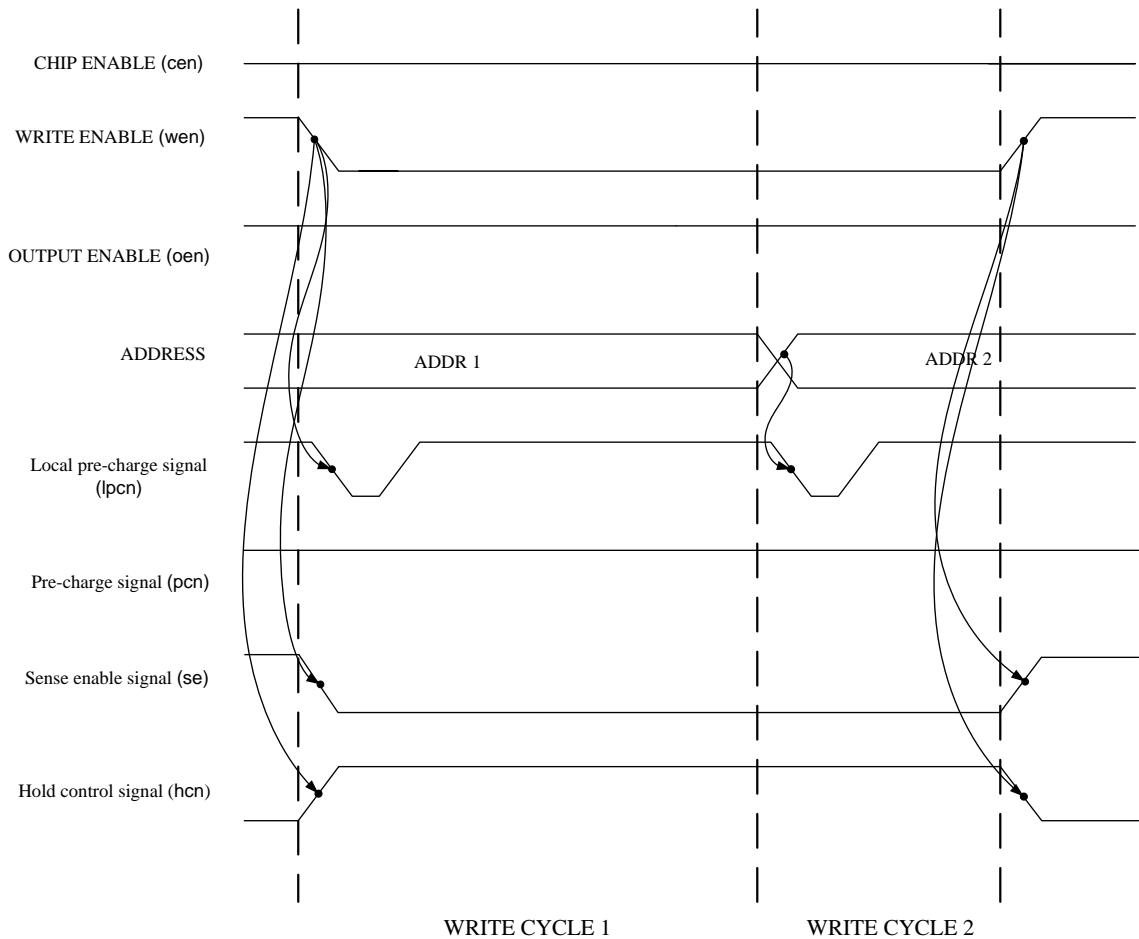


Figure 2.4 Timing diagram for a typical write operation

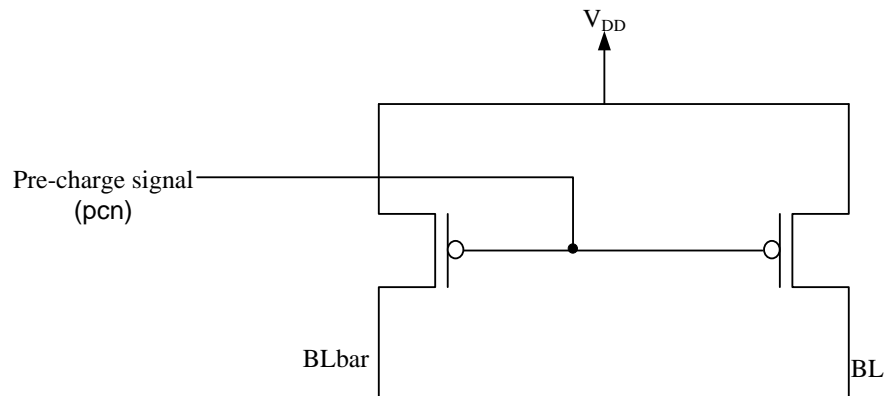


Figure 2.5 Simple pre-charge circuit

### 2.1.5 Pre-Charge Circuitry

As discussed in the previous section, the control circuit generates a pulse on the pre-charge signal (**pcn**) that causes the bit lines to be pulled high at the start of every read operation. One simple implementation of the pre-charge circuitry uses a PMOS transistor for each bit line, as shown in Figure 2.5. When the pre-charge signal (**pcn**) goes low, the PMOS transistors are ON and the bit lines are pulled up to  $V_{DD}$ . Every column has a pair of PMOS pull-up transistors.

### 2.1.6 I/O and Sensing Block

The I/O and sensing block consists of an array of sense amplifiers and an I/O block. Figure 2.6 shows the circuitry associated with a single bit of the data bus. The sense amplifier comes into play only during a read operation. The SRAM cell output is a pair of bit lines that are differential in nature. The sense amplifier is enabled by the sense enable signal (**se**) generated by the control circuit. When **se** is high (active), the sense

amplifier amplifies the analog differential voltage between the bit lines. The amplification results in a single-ended digital output that represents one bit read from the SRAM array.

As the name indicates, the I/O block controls the flow of data between the bit lines and the data bus. It can be viewed as an array of switches that routes the data bits and their complements from the data bus to the bit lines (BL and BLbar) during a write operation, and routes the data bits from the sense amplifiers to the data bus during a read operation. When *wen* is asserted, the I/O block puts each data bit to be written and its complement on the corresponding pair of bit lines. When *se* is asserted, the I/O block puts the data bits read from the SRAM array on the data bus.

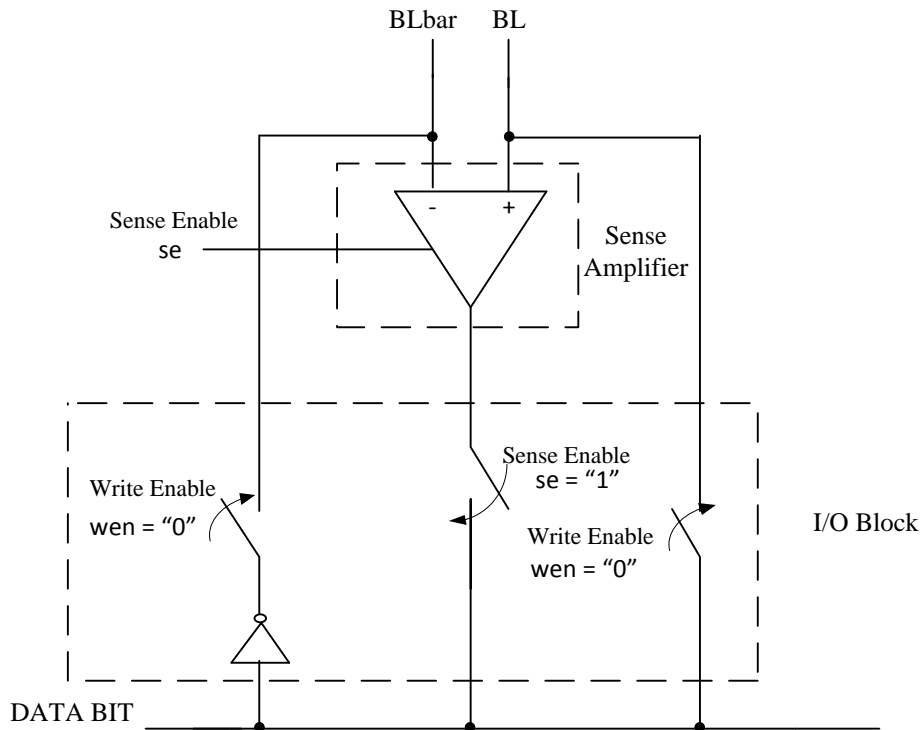


Figure 2.6 I/O and sensing block circuitry associated with one data bit

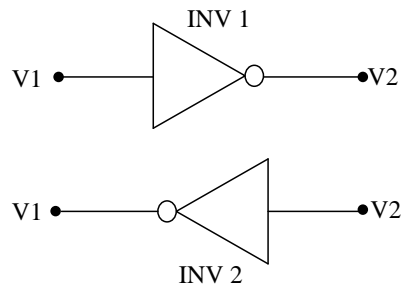


Figure 2.7 Circuit to find the hold static noise margin (HSNM)

## 2.2 Noise Margins

Stability is the most important criterion in the design of an SRAM, because the SRAM cell should be able to hold its data under all read and hold conditions. If the SRAM cell is not stable, it may result in a loss of data. The most important measures of stability for an SRAM cell are the hold static noise margin (HSNM) and the read static noise margin (RSNM). These static noise margins may be defined as the minimum noise voltage levels that could possibly force the cell to an undesired state during a hold or a read operation, respectively [14]. The SRAM is also required to be writable; the write static noise margin (WSNM) may be defined as the minimum noise voltage level that could prevent the SRAM cell from being driven to the desired state during a write operation.

### 2.2.1 Hold Static Noise Margin

Consider the 6T SRAM cell shown in Figure 2.2. During a hold operation, the row line is low and the NMOS pass transistors are OFF. This disconnects the bit lines from the cross-coupled inverters. The stability of the SRAM cell during a hold operation

can be analyzed by considering the static voltage transfer characteristics (VTCs) of the two inverters that make up the cross-coupled pair [6]. The two inverters INV1 and INV2 are shown in Figure 2.7. An example pair of voltage transfer characteristics for the two inverters, plotted on the same axes, is shown in Figure 2.8. The characteristics intersect at three points, two of which are the stable operating points corresponding to the storage of a “0” or a “1”. All the simulations done as a part of the thesis used Cadence Spectre® Circuit Simulator.

The hold static noise margin (HSNM) can be obtained [14] by finding the largest square that can be drawn, aligned with the V1 and V2 axes, between the VTCs of the two inverters. The edge dimension of the square represents the maximum noise voltage for which the SRAM cell is guaranteed to remain in its original state. The presence of noise, which can be modeled as noise voltage sources in series with the outputs of the inverters, shifts the VTCs and therefore affects their points of intersection. If noise voltages shift the VTC of INV1 to the left and the VTC of INV2 upward by more than the HSNM, the intersection corresponding to stable operating point 1 ceases to exist. If the SRAM cell had been holding a bit at that operating point, the bit would be lost. Noise voltages may affect stable operating point 2 in a similar way. The SRAM cell is guaranteed stable during a hold operation if the noise voltages are lower than the HSNM. For this example, the HSNM is 0.88 V, or 29% of  $V_{DD}$ .

### 2.2.2 Read Static Noise Margin

Consider again the 6T SRAM cell shown in Figure 2.2. At the beginning of a read operation, the bit lines are pre-charged to  $V_{DD}$ . Then the row line is activated, thus

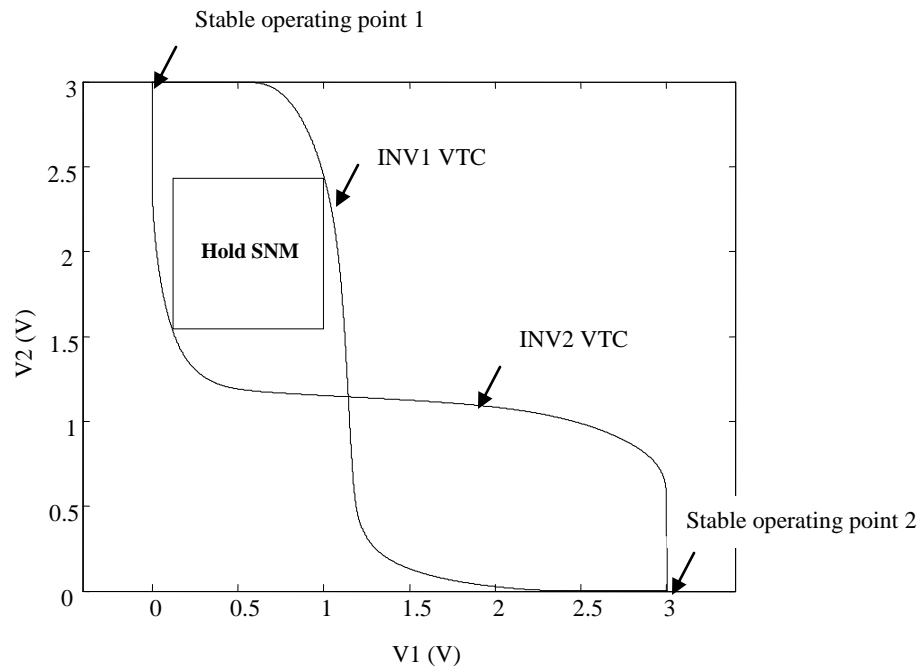


Figure 2.8 VTCs of the inverters showing the HSNM

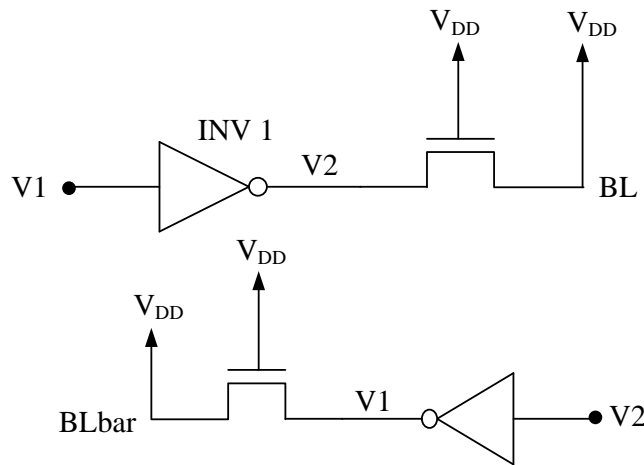


Figure 2.9 Circuit to find the read static noise margin (RSNM)

connecting the bit lines with the cross-coupled inverters. This connection causes loading of the inverters, which changes their voltage transfer characteristics. The simplified circuit for finding the read static noise margin (RSNM) is shown in Figure 2.9. Here, the

inverters are loaded by the NMOS pass transistors with the bit lines held at  $V_{DD}$ . The stability of the SRAM cell during a read operation can be analyzed by considering the static VTCs of the two loaded inverters.

Figure 2.10 shows an example set of VTCs for the two inverters, plotted on the same axes. The read static noise margin (RSNM) is the dimension of the largest square that can be drawn, aligned with the  $V_1$  and  $V_2$  axes, between the VTCs of the loaded inverters. The edge dimension of the square represents the maximum noise voltage for which the SRAM cell is guaranteed to remain in its original state, i.e., without any change in data stored, during a read operation.

From Figure 2.10, it can be observed that the VTCs do not extend all the way to 0 V; this is because the inverters are loaded by the NMOS pass transistors connected to the bit lines, which are modeled as being held at  $V_{DD}$ . For this reason, the RSNM of an SRAM cell is smaller than the HSNM of that cell. The RSNM for this example is 0.38 V, while the HSNM was previously found to be 0.88 V.

This method for calculating RSNM assumes that the bit lines are held at a constant voltage  $V_{DD}$ . In actuality, the bit lines are pre-charged to, but not held at,  $V_{DD}$ ; one of them will drop in voltage during the read. For this reason, the calculated RSNM is a conservative value.

### 2.2.3 Write Static Noise Margin

The write static noise margin (WSNM) can be used to analyze the writability of the SRAM cell. Refer again to the 6T SRAM cell in Figure 2.2. To write a “1”, the row line is activated while BL is pulled high and BLbar is pulled low. This

connection causes loading of the inverters, which changes their VTCs. The simplified circuit used for finding the WSNM when writing a “1” is shown in Figure 2.11. The writability of the SRAM cell can be analyzed by considering the static VTCs of the two loaded inverters.

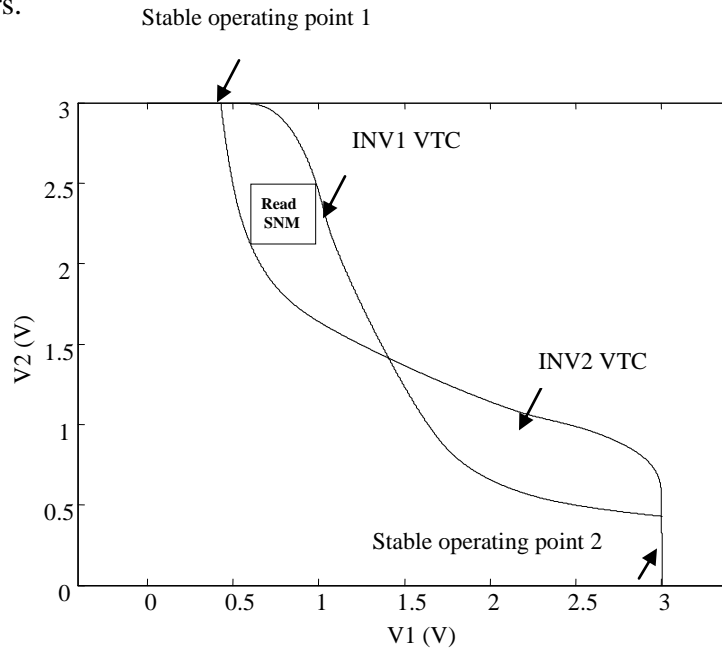


Figure 2.10 VTCs of loaded inverters showing the RSNM

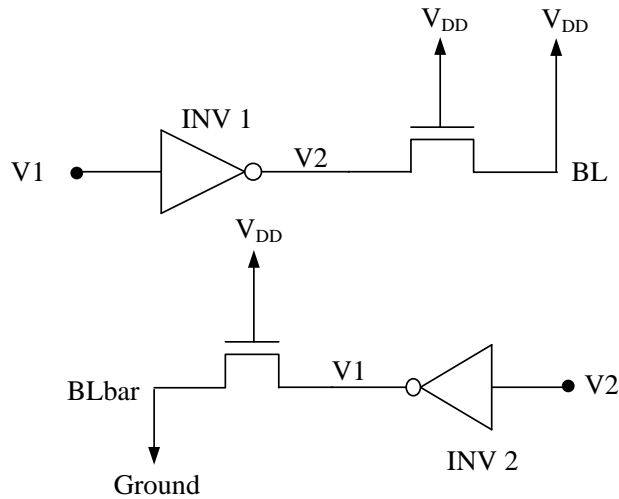


Figure 2.11 Circuit to find the write static noise margin (WSNM)



Figure 2.12 shows an example set of VTCs for the two inverters, plotted on the same axes. For a successful write operation, the VTC curves should cross only at one point; in such a case the cell is monostable, and must end up at the desired state. The write static noise margin is the minimum noise voltage level that, acting in a coordinated way on the two inverters so as to shift the VTC of INV1 down and the VTC of INV2 to the right, would cause there to be a second stable operating point. In this case, the WSNM can be obtained by sliding a 45° line upwards starting from the lower right and finding a location where it reaches its first local minimum; the edge dimension of the square drawn at this location gives the WSNM.

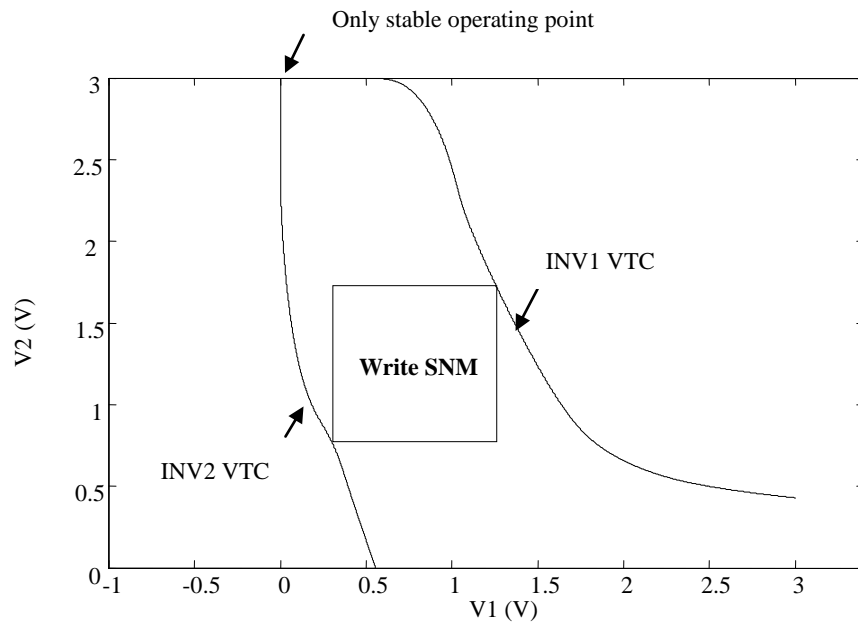


Figure 2.12 VTCs of loaded inverters showing the WSNM

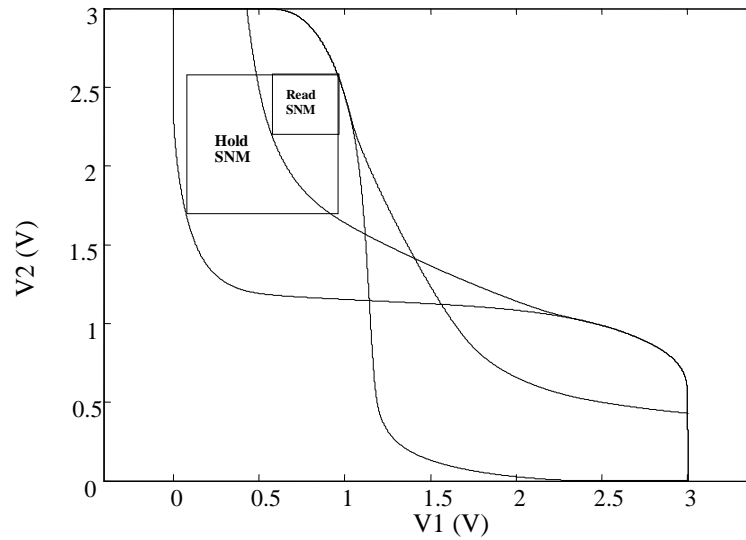


Figure 2.13 Comparison of the read static noise margin and hold static noise margin

If the noise voltage level exceeds the WSNM, the VTCs may not cross only at the desired operating point, but may have more than one intersection. In such a case, if the cell were initially in the “0” state, it would not reach the “1” state as desired. The write static noise margin for this example is 0.94 V.

In general, the HSNM is greater than the RSNM. Figure 2.13 shows a comparison of the VTCs that give the HSNM and the RSNM. A smaller RSNM shows that the SRAM cell is more vulnerable to data corruption during a read operation than during a hold operation. Therefore, it is necessary to focus on the RSNM to ensure that the SRAM is stable during a read operation.

This chapter has given an overview of all the components in the SRAM architecture. It has also discussed the noise margins and writability of the SRAM cell. In the next chapter, a complete SRAM, including the SRAM cell and all the associated

addressing, control and interface circuitry, is designed. The SRAM cell's noise margins are also analyzed.

## CHAPTER III

### SRAM DESIGN

In this chapter, the complete design of a 1Kword SRAM is given. Each word is 16 bits long. The SRAM has a single port, i.e., it is capable of one read or one write at a time. In the design process, careful consideration is given to the stability of the SRAM cell. The read and write access times of the memory are recorded. This chapter documents the design of all the components required to implement the SRAM.

Figure 3.1 gives a high-level block diagram of a 16Kbit SRAM. The 16Kbit SRAM consists of an array of SRAM cells to store the bits, along with decoding, multiplexing and control circuitry to select and access the bits in any part of the array. The 16Kbit SRAM is arranged in the form of a matrix with 64 rows and 256 columns. A 6T SRAM cell is used to store each bit. The 256 cells in each row are organized into sixteen words of sixteen bits each. A total of ten address lines are required to access the 1024 words. The six most significant bits of the address are used to select one of the 64 rows, and the remaining four bits of the address are used to select one of the sixteen words in the selected row. The design is explained in detail in the subsequent sections.

#### 3.1 SRAM Array

The SRAM array is a two-dimensional array of 6T SRAM cells. The sixteen words in a single row of the array are organized such that the most significant bits of all

sixteen words are grouped together, followed by the next most significant bits of all sixteen words, and so on, down to the least significant bits. The interleaving of the sixteen words stored in a single row of the array is shown in Figure 3.2. Thus, the array has sixteen columns, one for each bit in a word, with each column holding sixteen bits, one for each word in a row.

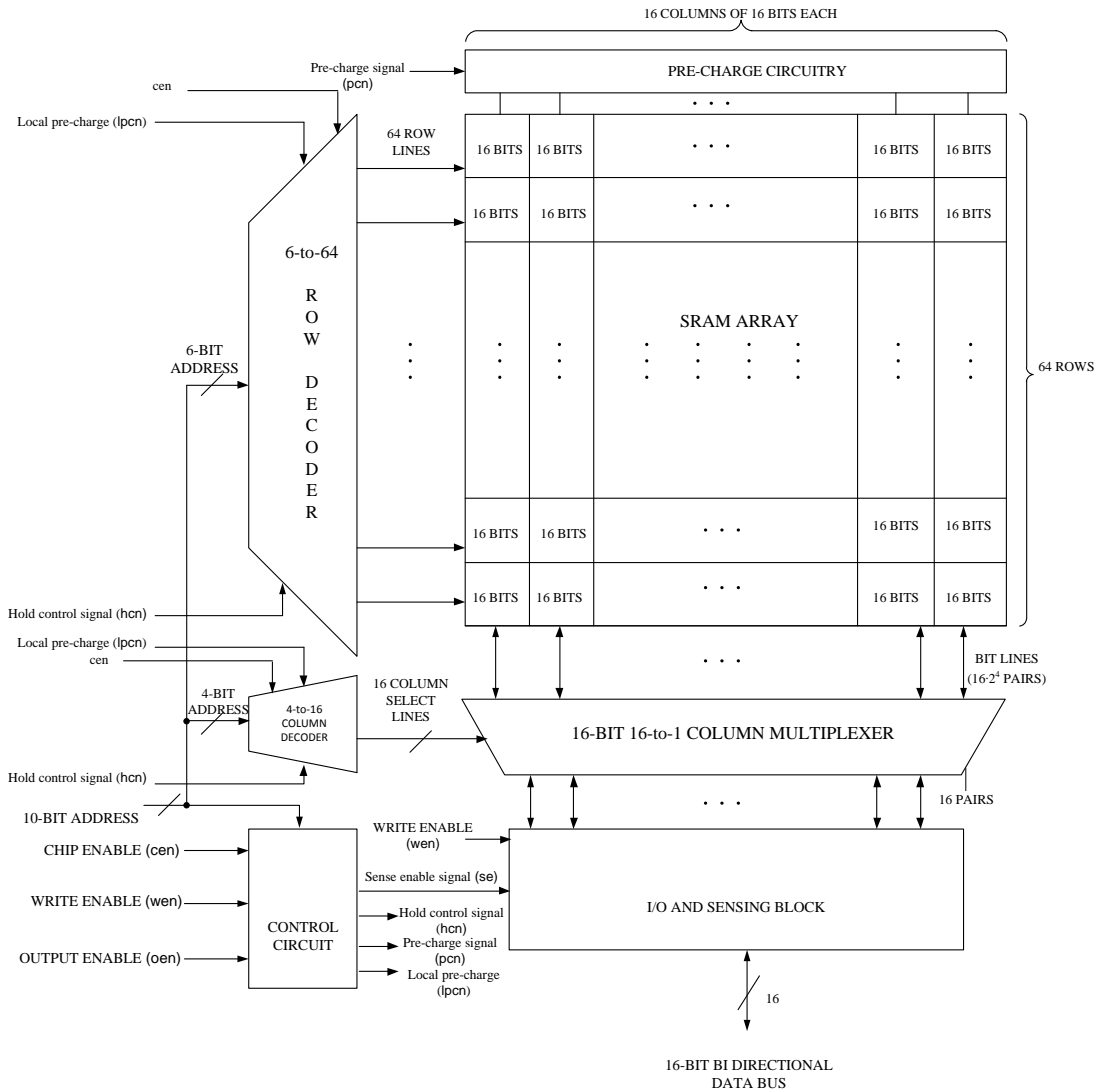


Figure 3.1 Block diagram of 1Kword SRAM

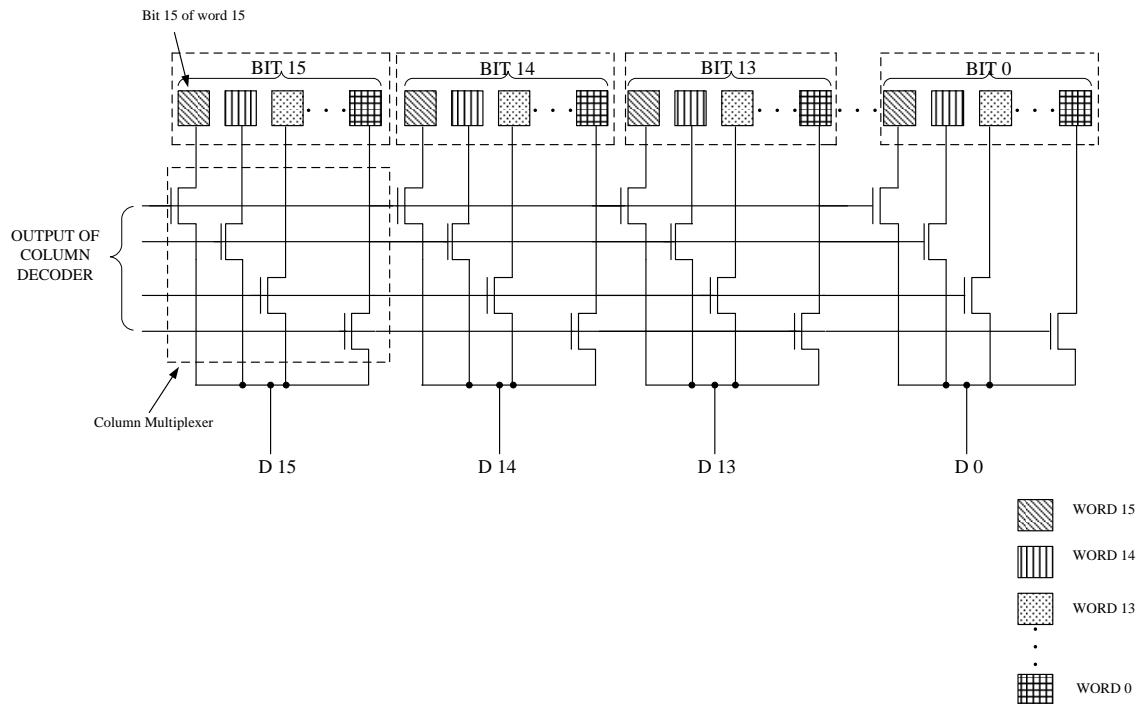


Figure 3.2 Organization of bits in a row

The circuitry in one of the 256 columns of the SRAM array is shown in Figure 3.3. This includes the SRAM cells, the address decoders along with the drivers, the column multiplexer and the I/O and sensing block. There is one sense amplifier for every sixteen columns. Each block is described in detail in the subsequent sections.

### 3.2 SRAM Cell

Consider the 6T SRAM cell shown in Figure 3.4. The stability of the SRAM cell is the most important factor that has to be considered when choosing the device geometries. The SRAM cell should be capable of withstanding certain levels of noise during hold and read operations; as described in Chapter 2, the hold static noise margin (HSNM) and the read static noise margin (RSNM) denote these noise voltage levels. For

large memory arrays, it is desirable to use the smallest transistors possible while still maintaining acceptable noise margins.

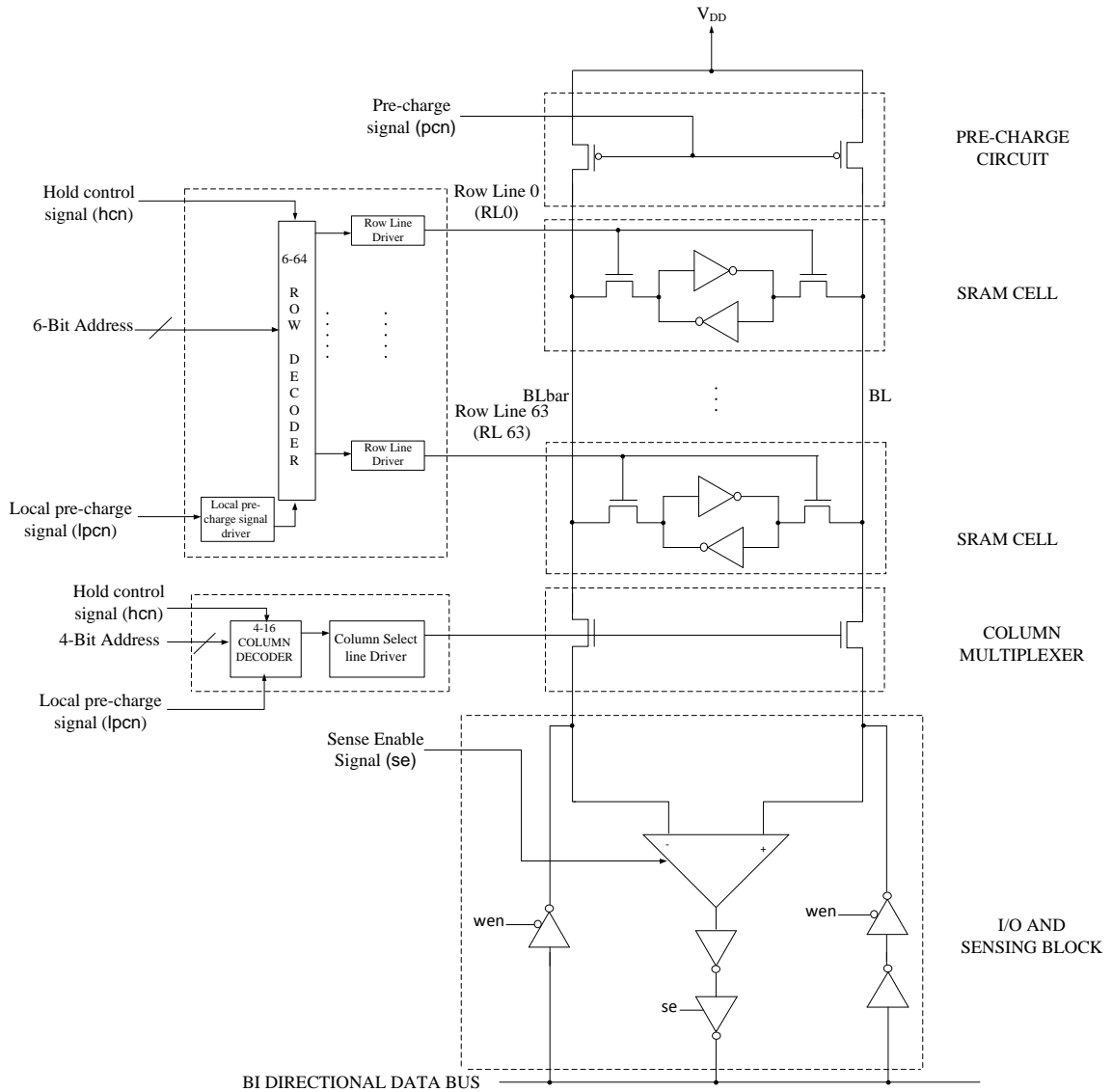


Figure 3.3 Circuitry in one column of the SRAM

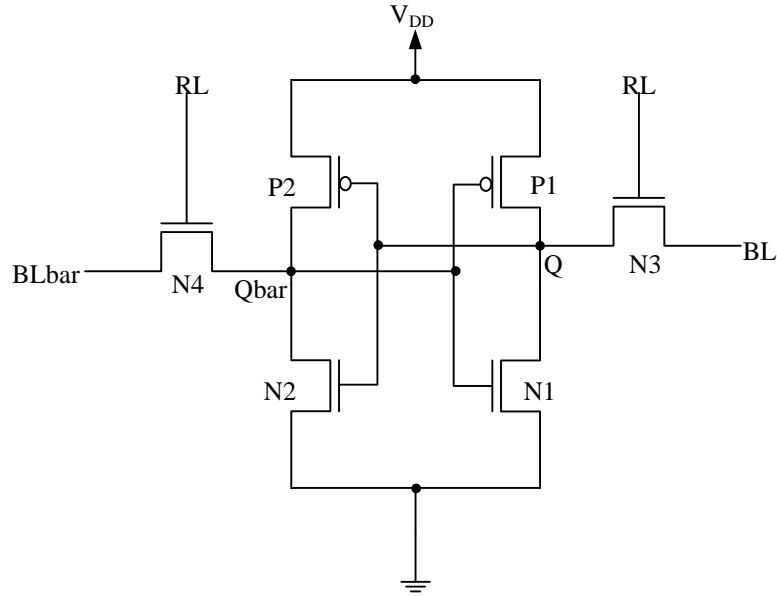


Figure 3.4 SRAM cell

Let us consider the SRAM cell in Figure 3.4 when the node Q holds a “0” and Qbar holds a “1”. During a read operation, the bit lines BL and BLbar are initially pre-charged to  $V_{DD}$ . When the row line RL is asserted, the bit lines are connected to the cross-coupled inverters through the NMOS pass transistors N3 and N4. The bit line BLbar remains high, reinforced through N4 and P2 while BL is discharged through N3 and N1. As BL is being discharged, the voltage at node Q temporarily rises before falling back to 0 V. This is because N1 and N3 form a voltage divider network which makes the voltage at Q a certain fraction of the voltage at BL during the discharge. The temporary rise in voltage at the storage node Q that holds a “0” depends on the relative sizes of N3 and N1. If N1 is made strong (large  $W/L$ ) compared to N3, the rise in voltage at the node Q will be relatively small. If this rise in voltage at node Q exceeds the voltage threshold of the P2-N2 inverter, the output voltage of the P2-N2 inverter may fall; this may result in a



*read upset*. Read upset is a situation where the cell changes state because of the rise in voltage at the “0” storage node during a read. To avoid a read upset, the rise in voltage at the node Q must be less than the switching threshold of the P2-N2 inverter. This constraint is called the Readability [13]. Readability under nominal conditions implies that the SRAM cell has a positive read static noise margin.

To meet the readability constraint, the device geometries must be such that N1 is stronger (larger  $W/L$ ) than N3. As the cell is symmetric, this implies that N2 is also stronger (larger  $W/L$ ) than N4. The cell ratio (CR) may be defined as the ratio  $W/L$  of the NMOS pull down divided by the ratio  $W/L$  of the NMOS pass transistor [10]. That is,

$$CR = \frac{\left(\frac{W}{L}\right)_{N1,N2}}{\left(\frac{W}{L}\right)_{N3,N4}} \quad (3.1)$$

Typically, for CR greater than 1.2, the voltage at the storage node does not rise above the inverter threshold. As a rule, N1 and N2 are made 1.2 to 2 times as wide as N3 and N4 [13].

Similarly, the SRAM cell is designed such that data is written reliably. During a write operation, suppose that Q holds a “0” and Qbar holds a “1”, and the data to be written is a “1”. To write a “1”, the I/O block drives BL to  $V_{DD}$  and BLbar to ground, and the row line is asserted. The voltage at the node Qbar must be pulled below the switching threshold of the P1-N1 inverter in order to change the state of the SRAM cell. This constraint is called Writability [13]. Writability under nominal conditions implies that the SRAM cell has a positive write static noise margin.

When writing a “1”, the node Qbar is discharged through N4; however, until the cell switches state, it also continues to be charged through the PMOS transistor P2. While both N4 and P2 are ON, they form a voltage divider network. If properly designed, the voltage at the node Qbar falls below the switching threshold of the P1-N1 inverter and feedback works to turn P2 OFF and N2 ON, so that Qbar is pulled down completely. Thus, the state of the cell changes and the desired data is written in the cell.

For a successful write operation, the NMOS pass transistor should be made stronger than the PMOS pull up. Therefore, N4 is made stronger (larger  $W/L$ ) than P2 and N3 is made stronger (larger  $W/L$ ) than P1. The pull-up ratio of the cell (PR) may be defined as the ratio  $W/L$  of the PMOS pull-up transistor divided by the ratio  $W/L$  of the NMOS pass transistor [10]. That is,

$$PR = \frac{\left(\frac{W}{L}\right)_{P1,P2}}{\left(\frac{W}{L}\right)_{N3,N4}} \quad (3.2)$$

The lower the PR, the easier it will be to pull down the node Qbar, thus making the cell more writable. To pull Qbar below the threshold of the P1-N1 inverter, the PR value has to be below approximately 1.8 [16]. As a rule, typical values for PR range from 0.4 to 1.8 [13].

The following are the steps followed to design an SRAM cell.

1. The pull down and the pass transistors are selected such that the CR is greater than 1.2 [16]. Typically, CR is in the range 1.2 to 2.

2. Once the sizes of the pull down and the pass transistors are selected, the size of the PMOS pull up transistors is selected such that the PR is less than 1.8 [16]. Typically, PR is in the range 0.4 to 1.8.

Once the SRAM cell dimensions are set, the HSNM, RSNM and WSNM can be determined from the VTCs as explained in Chapter 2.

Following the above steps to design an SRAM cell, a baseline SRAM cell was designed with  $CR=1.25$  and  $PR=0.47$ . The device geometries are shown in Table 3.1. The hold, the read and the write static noise margins obtained for the baseline device geometries are shown in Figures 3.5, 3.6 and 3.7. The HSNM is 0.88 V, the RSNM is 0.18 V and the WSNM is 1.08 V. It is clear that the baseline device geometries do not give a good RSNM. The Baseline Design will be modified in Chapter 4 to improve the RSNM while still maintaining an acceptable WSNM.

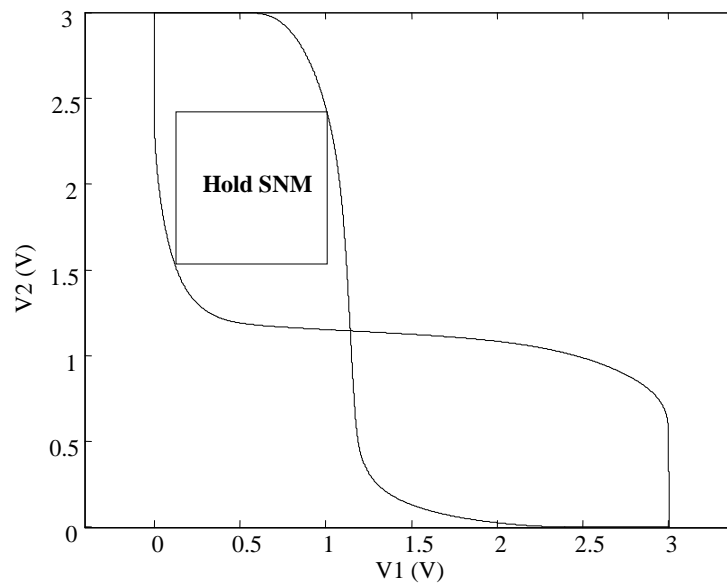


Figure 3.5 VTCs showing the hold static noise margin for the Baseline Design

Table 3.1 Transistor channel dimensions for the SRAM cell for the Baseline Design

Transistors	Channel Dimensions ( $W \times L$ )
P1, P2	1.4 $\mu\text{m} \times 0.5 \mu\text{m}$
N1, N2	6 $\mu\text{m} \times 0.8 \mu\text{m}$
N3, N4	4.8 $\mu\text{m} \times 0.8 \mu\text{m}$

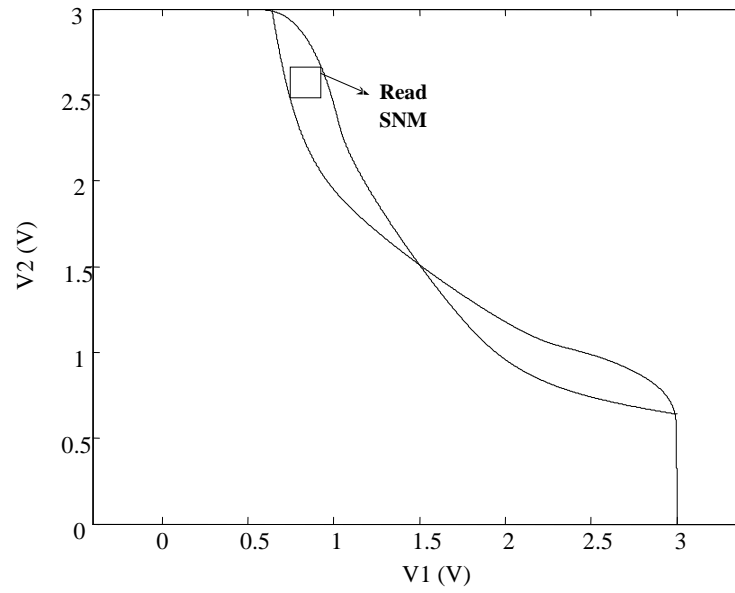


Figure 3.6 VTCs showing the read static noise margin for the Baseline Design

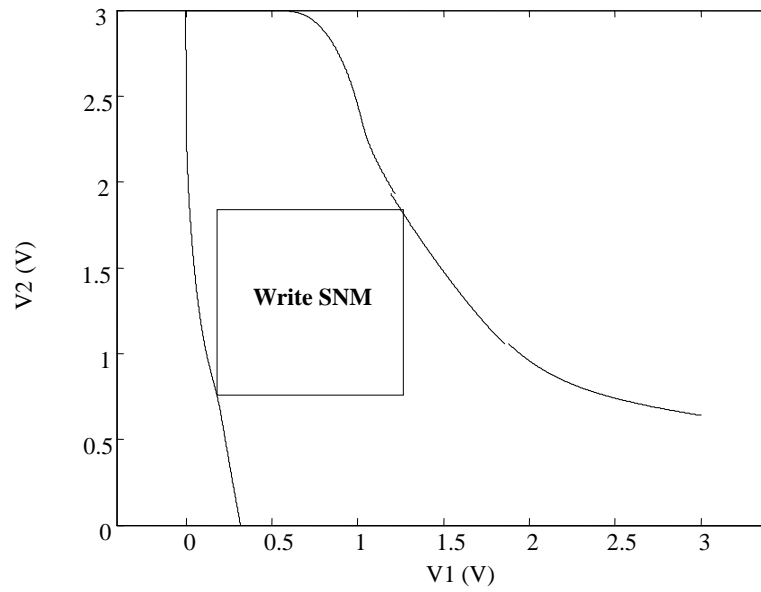


Figure 3.7 VTCs showing the write static noise margin for the Baseline Design

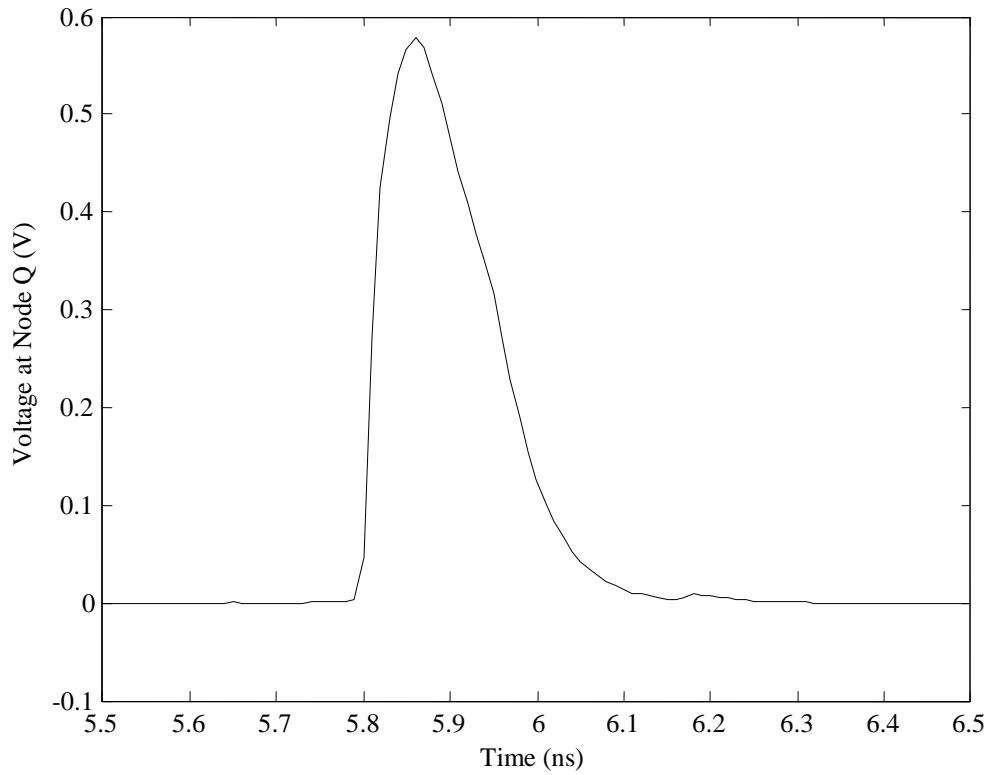


Figure 3.8 Transient simulation showing the rise in voltage at node Q during a read operation

A transient simulation was done on an array with 64 rows and 16 columns, i.e., one word per row, to observe the rise in voltage at the storage node Q that stores a “0” during a read operation. The only parasitics included in the analysis were the capacitances of the transistors. The simulation result is shown in Figure 3.8. The rise in voltage at the node Q is 0.57 V. The switching threshold voltage of the inverter is around 0.78 V. The rise in voltage at the node Q is less than the switching threshold of the inverter by 0.21 V. By comparison the RSNM from Figure 3.6 is only 0.18 V. Note that the RSNM is a worst-case value that accounts for the most destabilizing combination of noise voltages and assumes the bit lines do not vary from  $V_{DD}$  when loaded.

### 3.3 6-to-64 Row Decoder

A 6-to-64 row decoder takes six bits of the address as input and selects one of the 64 rows. The different blocks involved in the row address decoding are shown in Figure 3.9. The row decoder logic is implemented as a dynamic NOR decoder. A basic 2-to-4 dynamic NOR decoder is shown in Figure 3.10. The operation of a dynamic NOR decoder follows two stages: a pre-charge stage and an evaluation stage. The pre-charge stage uses  $lpcn$  generated by the control circuit which is low (active) for 2.5 ns. The pulse on  $lpcn$  turns the PMOS transistors ON and pulls the 64 decoder output lines to  $V_{DD}$ . Once the pre-charge is done, the control circuit sets  $lpcn$  high (inactive), turning the PMOS transistors OFF. The row lines stay high because of the capacitively stored charge in the decoder output lines. In the evaluation stage, NMOS transistors corresponding to the input address are turned ON and all the decoder output lines except one are discharged to ground.

For a 6-to-64 row decoder, the local pre-charge signal (*lpcn*) has to drive the gates of 64 PMOS pre-charge transistors, which represent a large capacitive load. So, a local pre-charge driver is used for buffering to quickly turn ON the 64 pre-charge transistors. The drivers are generally designed such that the MOSFETs in each inverter are wider than those in the preceding one by the same factor or stage ratio. The local pre-charge driver is shown in Figure 3.11. The minimum sized inverter represented as X is constructed of PMOS and NMOS transistors with channel dimensions  $1.2 \mu\text{m} \times 0.8 \mu\text{m}$ . A stage ratio of 3.5 was found through transient simulation to result in rise and fall times around 0.5 ns and a propagation delay of 3.5 ns.

To ensure that only the desired row line is driven high during a read or a write operation and none of the row lines in the SRAM array are selected during a hold operation, the outputs of the 6-to-64 row decoder are gated with the hold control signal (*hcn*). The gating circuitry is shown in Figure 3.12. Each gated output of the row decoder is connected via a long wire to the 256 pairs of NMOS pass transistors present in the

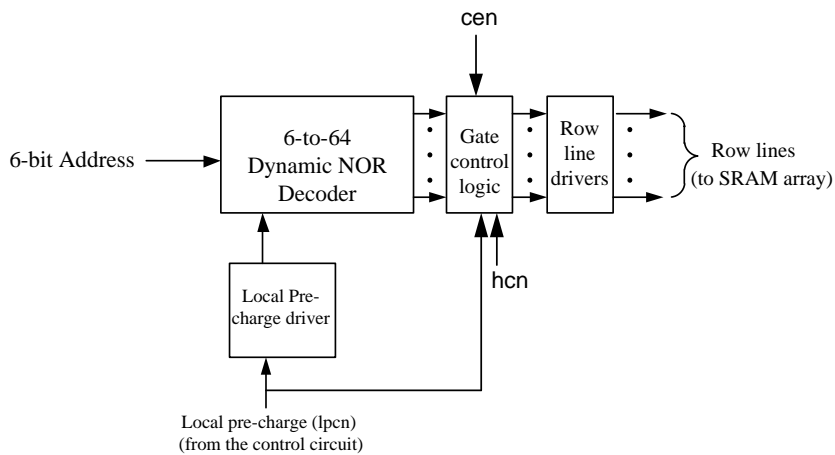


Figure 3.9 Sub-blocks of the 6-to-64 row decoder

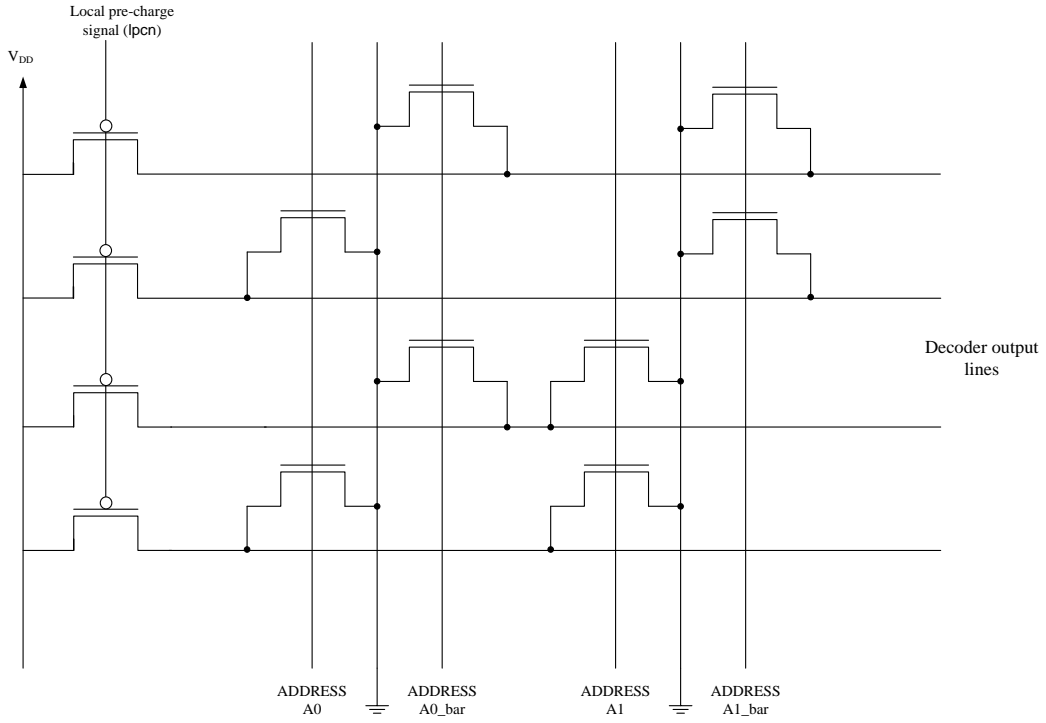


Figure 3.10 A 2-to-4 dynamic NOR decoder

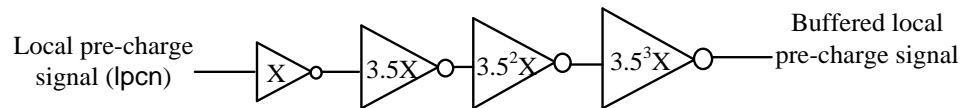


Figure 3.11 Local pre-charge driver

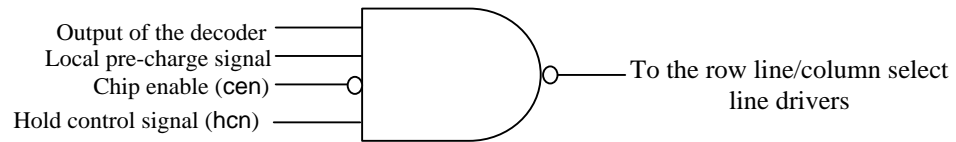


Figure 3.12 Gate control logic for each individual row line and column select line



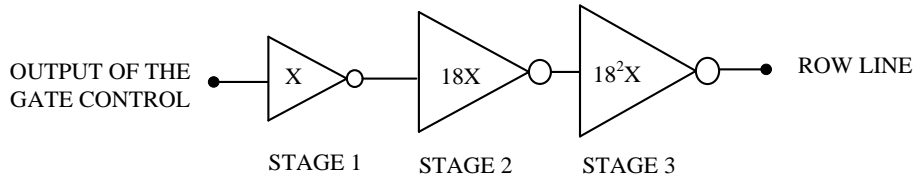


Figure 3.13 Row line driver

SRAM cells in the row. The row line thus represents a large RC load to the row decoder. Therefore, row drivers are used to drive the large row lines as shown in Figure 3.13.

A row with 256 SRAM cells was simulated. The load considered in the simulation consisted only of the gate capacitances in the device models for 512 MOSFETs. The output of the gate control is inverted so the row line driver must consist of an odd number of inverters. In this design, the row line driver consists of three inverters as shown in Figure 3.13. A stage ratio of 18 was found to give rise and fall times around 0.5 ns and a propagation delay of around 2 ns, which was deemed sufficiently fast.

#### 3.4 4-to-16 Column Decoder

The column decoder takes four bits of the address as input bits and selects one of the 16 columns. The different blocks involved in the column address decoding are shown in Figure 3.14.

For a 4-to-16 column decoder, the  $\text{lpcn}$  from the control circuit can drive the sixteen PMOS pre-charge transistors directly, so a local pre-charge driver was not used. A column driver is required so that the row line and the bit lines are selected at the same time and it is designed such that the column select line has rise and fall times comparable to those of the row line. The output of the column decoder has to drive 32 NMOS

transistors, which represent a capacitive load to the decoder. Therefore, the column drivers are used.

A column driver is shown in Figure 3.15. As the gate control gives an inverted output, the column driver must use an odd number of inverter stages. The column driver uses three stages and each stage is 17 times as wide as the previous stage. Transient simulation shows that, with a stage ratio of 17, the three inverter stages effectively drive the capacitive load of the column select line of the column multiplexer, with rise and fall times of around 0.5 ns and a propagation delay of 3.5 ns. The transient simulations included the gate capacitance of 32 NMOS transistors each of size  $4.5 \mu\text{m} \times 0.5 \mu\text{m}$  and the wire model of the column select line was not included.

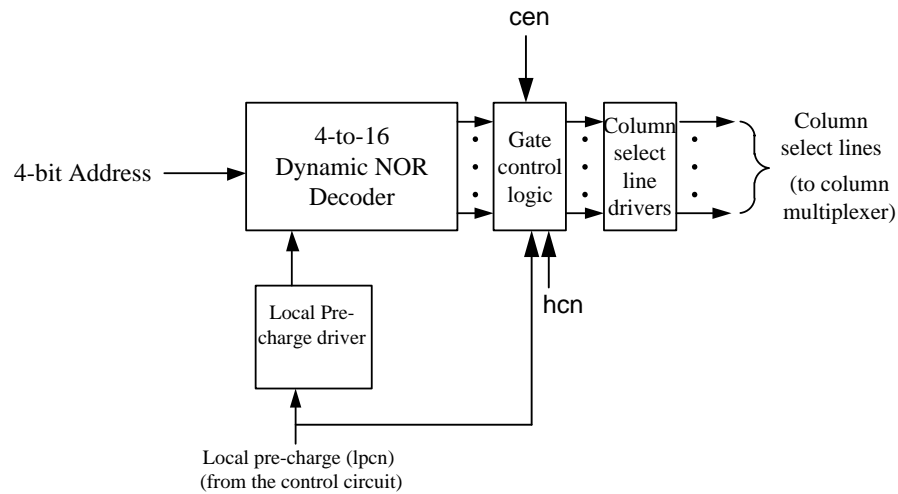


Figure 3.14 Sub-blocks of the 4-to-16 column decoder

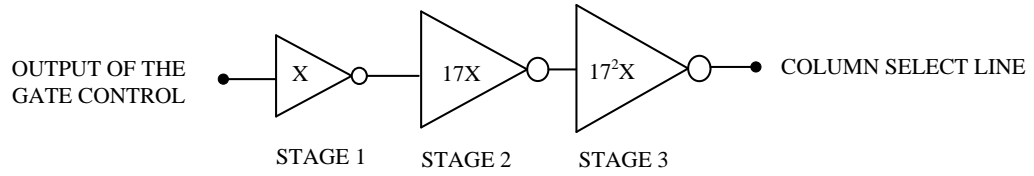


Figure 3.15 Column select line driver

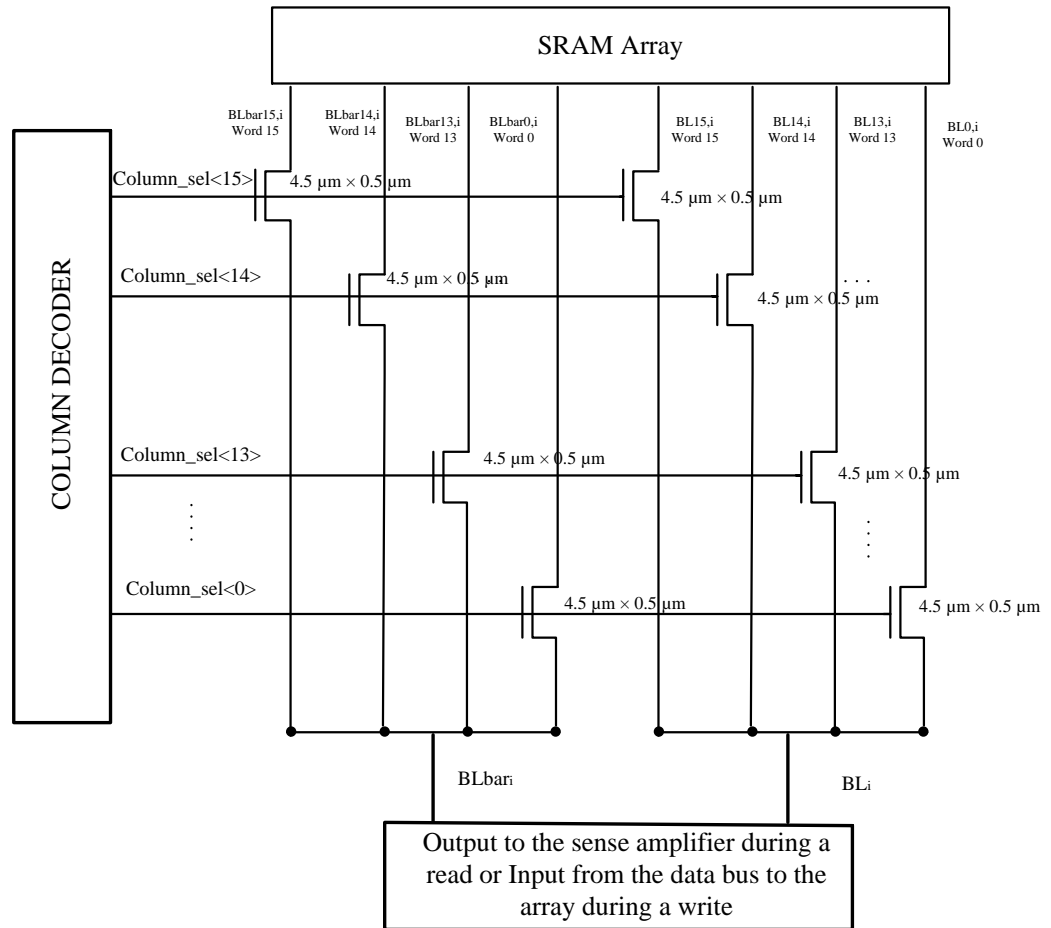


Figure 3.16 16-to-1 Column multiplexer/demultiplexer circuitry for one bit

### 3.5 16-bit 16-to-1 Column Multiplexer

The column multiplexer uses NMOS pass transistors to connect the appropriate bit in each column to the corresponding bit of the data bus. During a read operation 16 bits are to be read, one from each column. The NMOS pass transistors act as a 16-to-1 multiplexer to select the proper bit from among the bits of all 16 words stored in a given row. There is a sense amplifier for every sixteen columns that is used to read out the data provided by the column multiplexer.

During a write operation the 16-bit data word is to be written, one bit in each column. The NMOS pass transistors act as a 1-to-16 demultiplexer. Depending on the column address, only one column is selected and the data word to be written is directed from the data bus onto the desired bit lines. In this design there are sixteen 16-bit multiplexers. The 16-to-1 column multiplexer/demultiplexer circuitry for one bit is shown in Figure 3.16. The size of each NMOS pass transistor is  $4.5 \mu\text{m} \times 0.5 \mu\text{m}$ .

### 3.6 Pre-charge Circuitry

During a read operation, the bit lines BL and BLbar are pre-charged using a pair of PMOS transistors. The control circuit generates a pulse on pcn during a read, which turns the PMOS transistors ON so that the bit lines can be pulled to  $V_{DD}$ . During a write operation, pcn is held high (inactive) and the PMOS transistors are OFF. The PMOS transistor size was selected considering the resistance and the capacitance all through the length of the bit lines. The size of the selected PMOS transistor is  $2.4 \mu\text{m} \times 0.5 \mu\text{m}$ .

To calculate the minimum pulse width required for the pre-charge signal so that the bit lines are pre-charged to  $V_{DD}$ , the resistance and capacitance of the metal bit lines

used for the layout were considered to form a lumped RC model. The RC model of one of the bit lines of a column is shown in Figure 3.17. The bit line is divided into segments between the SRAM cells of the array. Each segment is modeled with two resistors and one capacitor in a T configuration. The length and width of each wire segment, measured

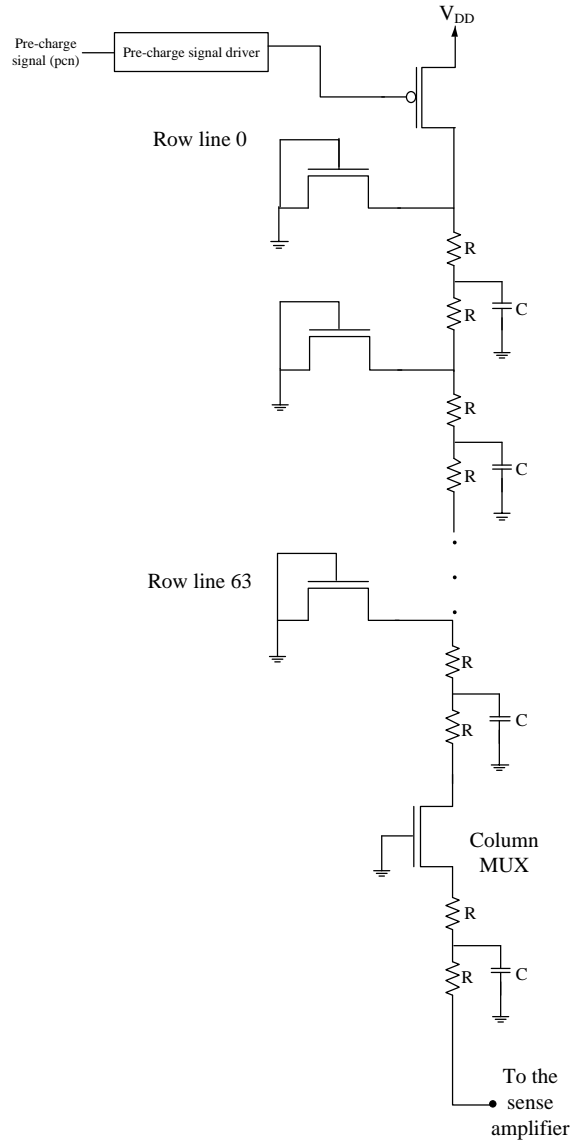


Figure 3.17 RC model used for calculating the pulse width of the pre-charge signal

from the layout is  $10\ \mu\text{m}$  and  $1.5\ \mu\text{m}$ , respectively. The sheet resistance and the plate capacitance values for each wire segment are assumed to be  $0.031\ \Omega/\text{sq}$  and  $0.018\ \text{fF}/\mu\text{m}^2$ . Therefore, each resistance  $R$  shown in the diagram is  $0.103\ \Omega$  and each capacitance  $C$  is  $0.270\ \text{fF}$ . A transient simulation showed that a pre-charge signal pulse width of  $2.5\ \text{ns}$  is sufficient to pull all the capacitive nodes on the bit lines, which were initialized to  $0.0\ \text{V}$ , to  $V_{\text{DD}}$ .

### 3.7 Control Circuit

The control circuit takes in the chip enable signal ( $\text{cen}$ ), the output enable signal ( $\text{oen}$ ), the write enable signal ( $\text{wen}$ ) and a ten-bit address and generates a local pre-charge signal ( $\text{lpcn}$ ), a hold control signal ( $\text{hcn}$ ), a pre-charge signal ( $\text{pcn}$ ) and a sense enable signal ( $\text{se}$ ) with the required timing. Figure 3.18 shows the schematic of the

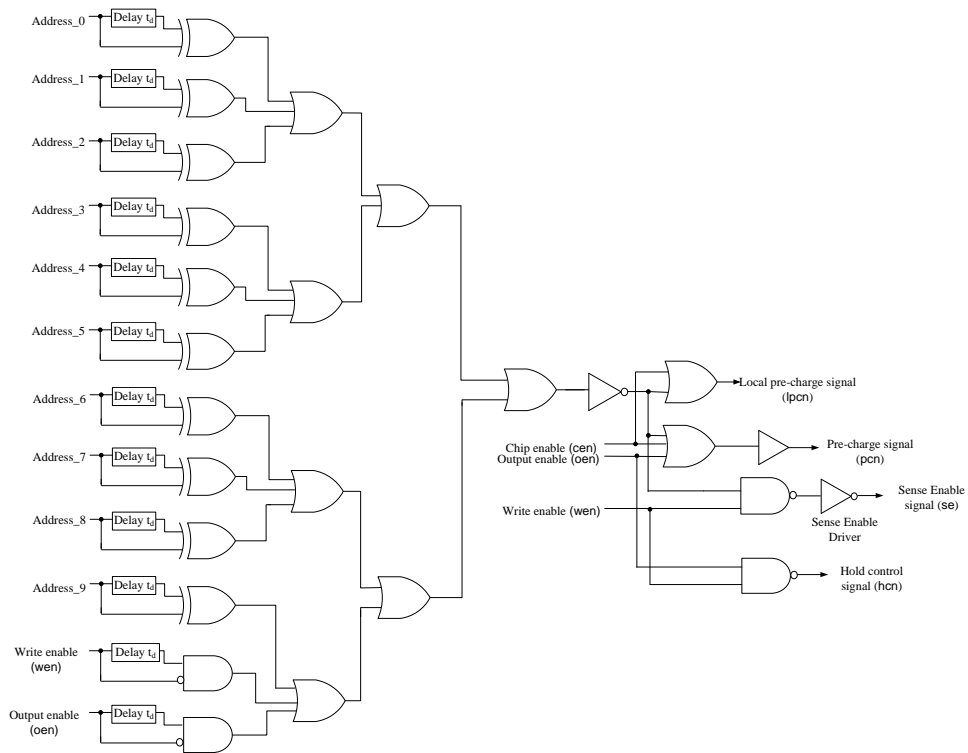


Figure 3.18 Control circuit logic

control circuit. As described in Chapter 2, the control circuit generates a pulse whenever there is a change in the input address or a transition in the output enable signal ( $oen$ ) from high (inactive) to low (active) or a transition in the write enable signal ( $wen$ ) from high (inactive) to low (active). This pulse is used to generate  $lpcn$ ,  $pcn$  and  $se$  with the required timing.

The widths of the pulses  $pcn$ ,  $lpcn$  and  $se$  generated by the control circuit in response to a change in the address or a transition in the  $oen$  or the  $wen$  from high (inactive) to low (active) are determined by the delay elements included in the control circuit. Each delay element is implemented using a chain of inverters. The delay element used is shown in Figure 3.19, where each inverter is annotated with its dimensions in microns. The inverter chain was sized to obtain pulse widths of 2.5 ns. A transient simulation was done by sweeping the pulse width of the pre-charge signal from 1.0 ns to 3.0 ns and it was observed that with a pulse width of 2.5 ns the bit lines were pulled to  $V_{DD}$ .

The functionality of the control circuit has been described in Chapter 2. During every read operation, the low (active) pulse on  $pcn$  pulls the bit lines to  $V_{DD}$  and then goes high (inactive) before the row line goes high (active). The  $pcn$  signal from the control circuit is not directly given to the PMOS pull up transistors in the SRAM array. A pre-charge buffer is used to drive the load of 512 PMOS transistors. The buffered pre-charge signal rises (i.e., goes inactive) just before the row line is asserted. The pre-charge buffer is shown in Figure 3.20. Four inverter stages are used and every stage in the driver

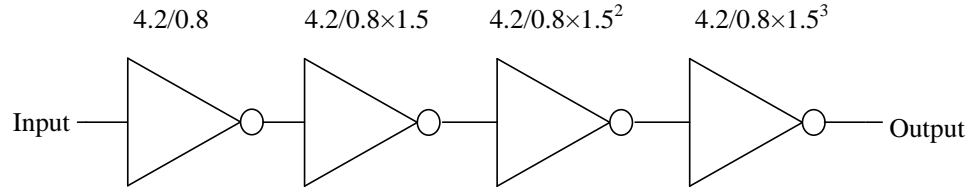


Figure 3.19 Delay chain used in the control circuit

is 3.4 times wider than the previous stage. X represents an inverter constructed of MOSFETs with dimensions  $1.2 \mu\text{m} \times 0.8 \mu\text{m}$ . Similarly, the sense enable signal from the control circuit has to drive the load of sixteen sensing blocks, so a sense enable signal driver is used. The sense enable driver is shown in Figure 3.21. The driver consists of three inverter stages and the transistors in every stage in the driver are 7 times wider than those in the previous stage. The buffered sense enable signal rises (i.e., goes active) 1 ns after the row line goes high (active) so that the sensed output is not passed to the data bus before a differential voltage can be detected between the bit lines.

### 3.8 I/O and Sensing Block

The I/O and sensing block consists of the input / output block and a sense amplifier. The input / output block is a data control block that routes the data to and from the 16 bit bi-directional data bus. During a write operation, data flows in from the data

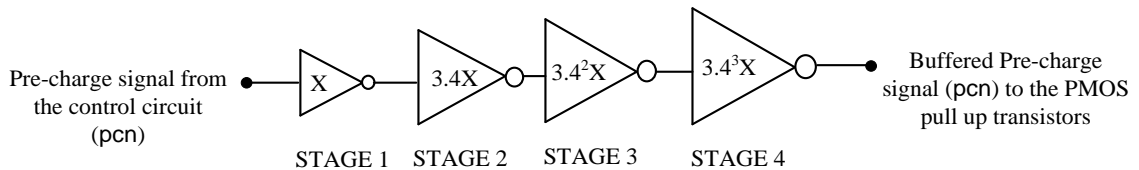


Figure 3.20 Pre-charge buffer



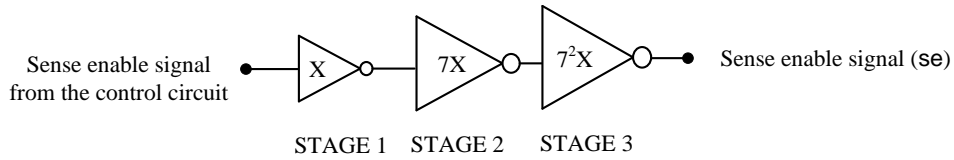


Figure 3.21 Sense enable signal driver

bus to the bit lines BL and BLbar. During a read operation, the data from the SRAM cell is read by sensing the difference in potential between the BL and BLbar and the data is put on the data bus. The direction of flow of data during a write or a read operation is controlled using a set of tri-state inverters.

There is a 16-bit data bus that allows us to write or read 16 bits at a time and therefore sixteen tri-state inverters are used for a read and sixteen pairs of tri-state inverters for a write operation. Figure 3.22 shows the general structure of the I/O block circuit. A transistor-level tri-state inverter circuit is shown in Figure 3.23.

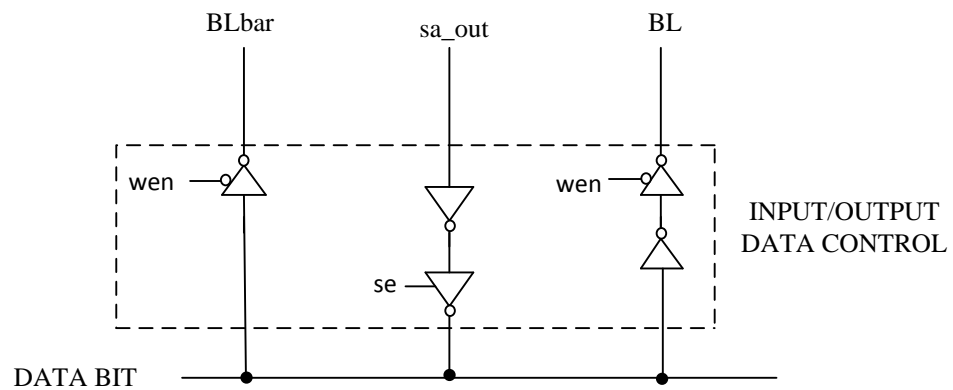


Figure 3.22 I/O Block

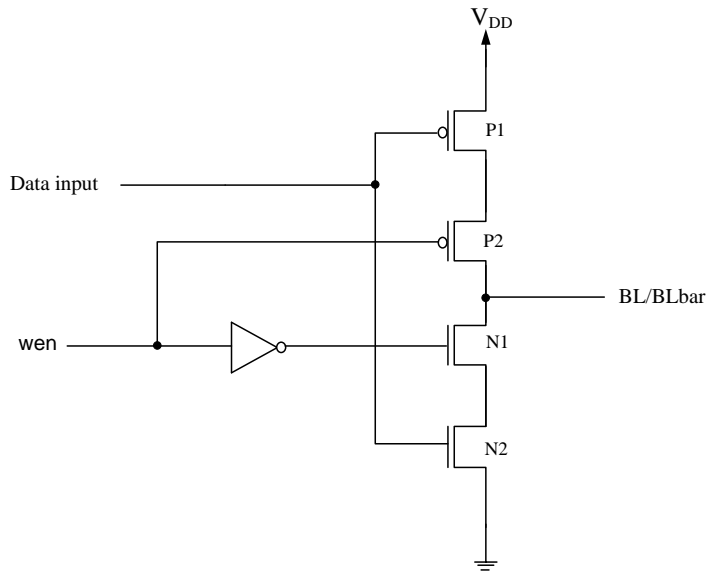


Figure 3.23 Expanded circuit diagram of the tri-state inverter driving BL or BLbar

During a write operation a low (active) **wen** enables the tri-state inverters that drive the bit lines BL and BLbar. During a read or a hold operation, when **wen** is high (inactive), those tri-state inverters are disabled and their outputs are in the high-impedance state.

Similarly during a read operation, the **oen** is low (active), and the control circuit generates a high (active) sense enable signal (**se**) which enables the tri-state inverter that drives the data bus. Thus, when **se** is high (active), the output of the sense amplifier is put on the data bus. During a write or a hold operation, **se** is low (inactive) and the output of that tri-state inverter is in a high-impedance state.

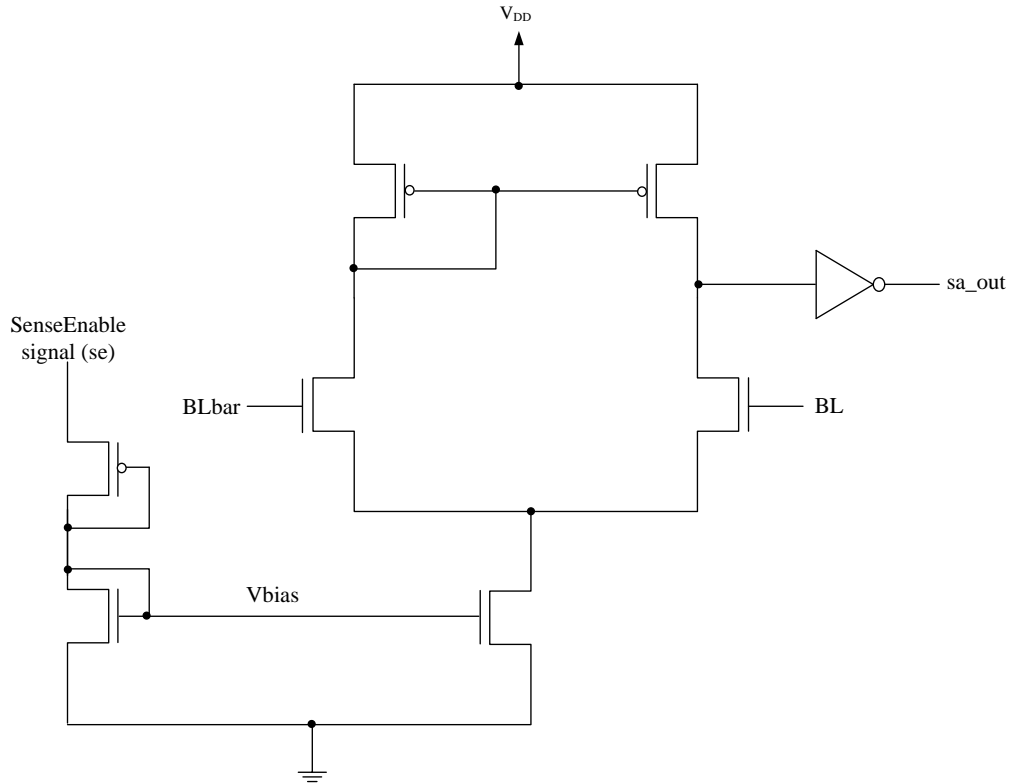


Figure 3.24 Differential sense amplifier

The sense amplifier is implemented using a differential amplifier. The sense amplifier is shown in Figure 3.24. The differential amplifier senses a small voltage difference across the bit lines and provides an amplified signal as its output.

During a read operation, the control circuit generates a high (active) sense enable signal (se). When se is high (active), the sense amplifier amplifies the difference in potential between the bit lines and the data is read. The sense amplifier is designed such that the bias voltage  $V_{bias}$  is applied to the NMOS tail current source only when se is high (active). When se is low (inactive), the bias voltage,  $V_{bias}$  is less than the threshold voltage of tail the NMOS current source and the output of the sense amplifier is

disconnected from the bi-directional data bus by the tri-state inverter as shown in Figure 3.22.

Thus, the complete design of a 1Kword SRAM has been discussed. In the design process, the stability of the SRAM cell was discussed and it is found that the design can be improved to get an improved read static noise margin. The improved design, and the analysis involved in the operation of the SRAM at high temperatures is discussed in Chapter 4.

## CHAPTER IV

### HIGH-TEMPERATURE OPERATION

The goal of this chapter is the design of an SRAM cell that is stable over a range of temperatures while maintaining acceptable access times. Three different designs are presented to show a range of trade-offs between stability and access time. Noise margins for all three designs are included.

#### 4.1 Operation of a Transistor at High Temperature

The operation of a transistor at high temperature is affected by degradation in mobility, a drop in threshold voltage, and an increase in intrinsic free carriers [17]. The performance of a transistor over temperature can be measured by observing the variation of the operating current and the leakage current through the transistor with temperature.

Figure 4.1 shows an NMOS transistor in saturation. During a read or a write operation, when the row line first goes high (active), one of the NMOS pass transistors is in saturation with the voltages shown. Figure 4.2 shows the variation of operating current with temperature for the NMOS transistor in saturation. From Figure 4.2, it is clear that a degradation in mobility decreases the operating current of the MOSFET with the given terminal voltages, resulting in a decrease in the speed of a read or write operation.

Figure 4.3 shows an NMOS transistor in cut-off. In the SRAM cell, when the row line first goes low (inactive), one of the NMOS pass transistors is in cut-off with the voltages shown. Figure 4.4 shows a sharp increase in leakage current as temperature increases from 27°C to 120°C, which results from both the increase in thermally-produced free carriers and the decrease in the threshold voltage [17].

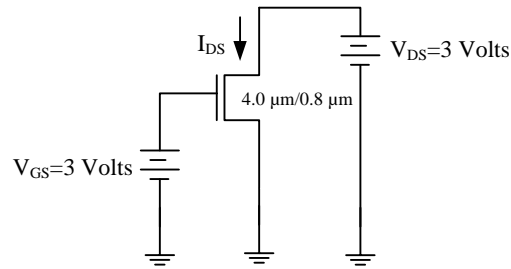


Figure 4.1 NMOS transistor in saturation

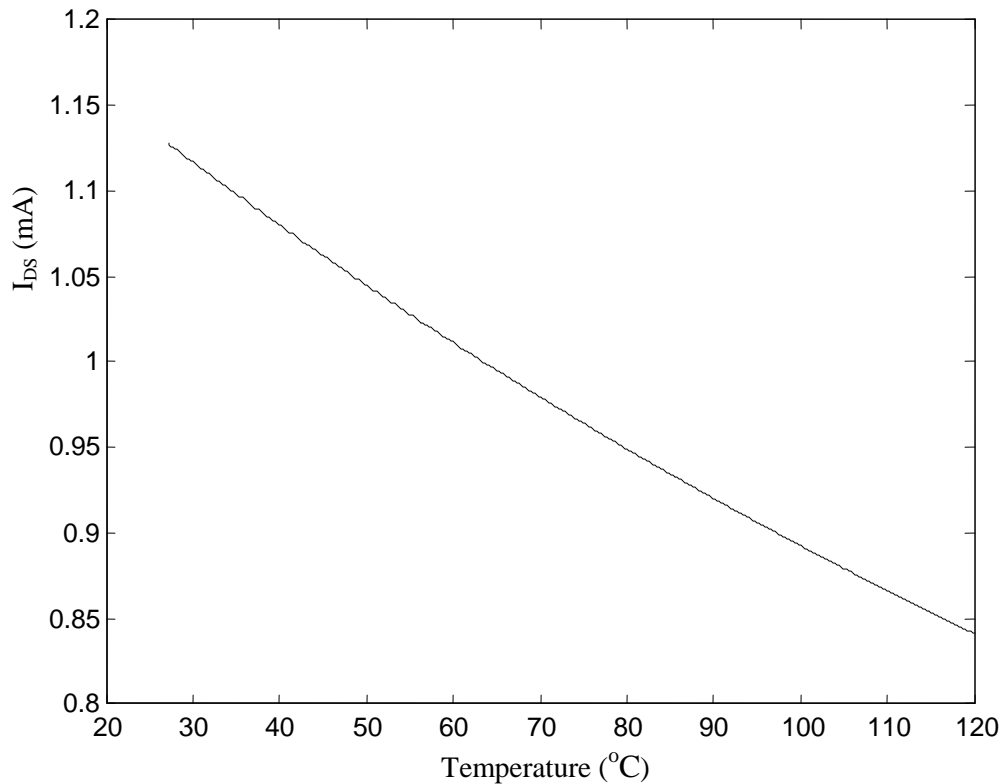


Figure 4.2 Variation of the drain-to-source current through a  $4.0 \mu\text{m} / 0.8 \mu\text{m}$  NMOS transistor in saturation with temperature

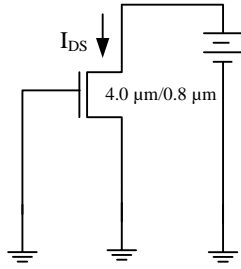


Figure 4.3 NMOS transistor in cut-off

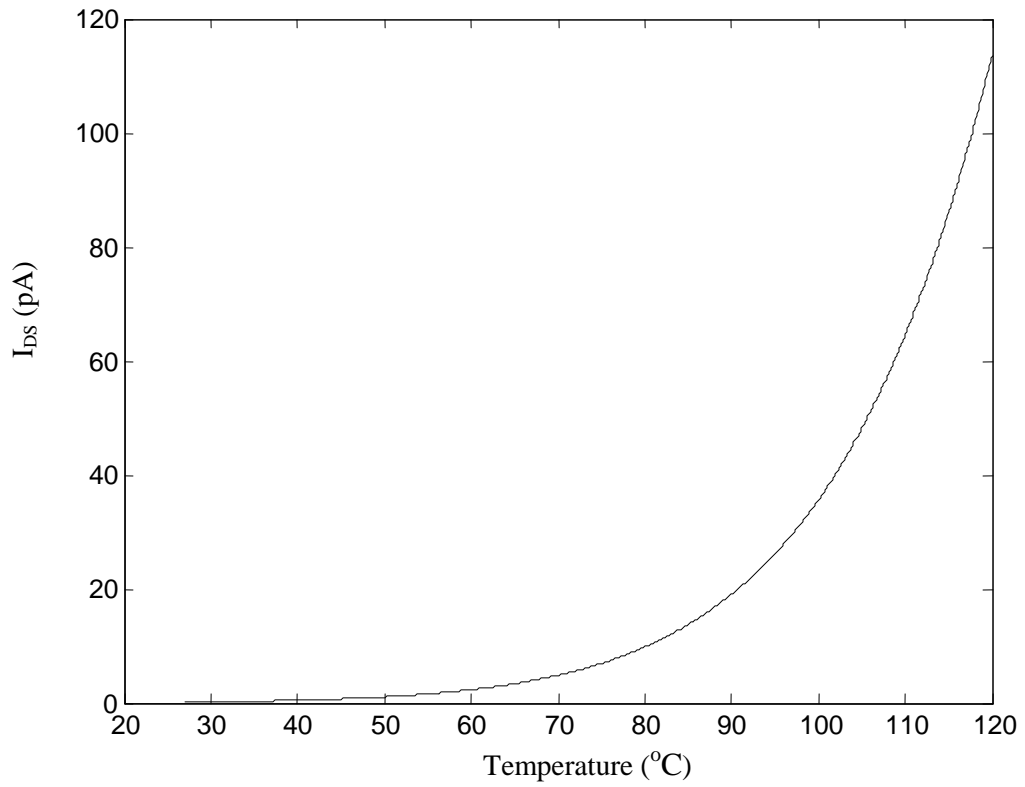


Figure 4.4 Variation of the leakage current through a  $4.0 \mu\text{m} / 0.8 \mu\text{m}$  NMOS transistor in sub-threshold region with temperature

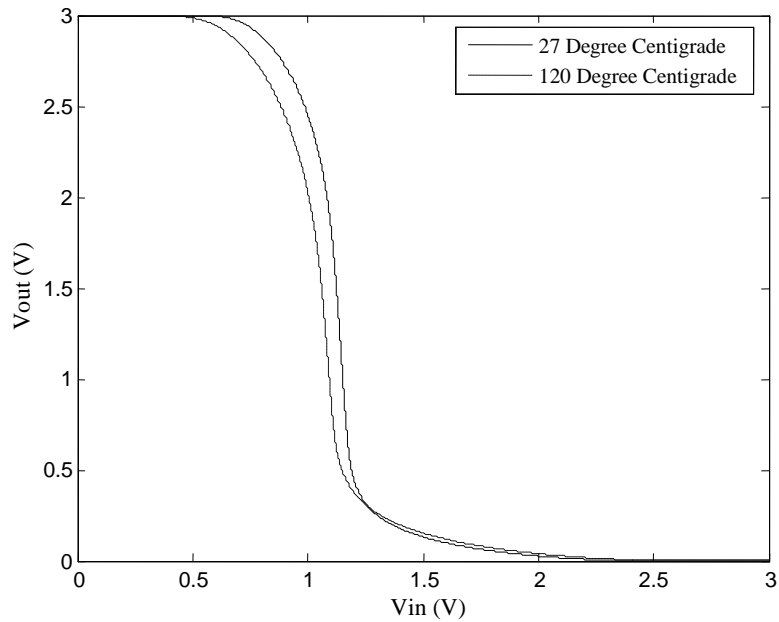


Figure 4.5 VTC of an inverter at low and high temperature

Figure 4.5 shows the voltage transfer characteristic of an inverter at room temperature and at high temperature. At high temperature, the VTC drops from the supply voltage at a lower input voltage, because the threshold voltage of the NMOS transistor has decreased. Similarly, the VTC rises earlier from ground potential as the input voltage drops from the supply voltage, because the threshold voltage of the PMOS transistor has decreased.

#### 4.2 High-Temperature Operation of SRAM

Consider the 6T SRAM cell shown in Figure 4.6. Let us assume that the SRAM cell holds a “1”, i.e., the node Q holds a “1” and Qbar holds a “0”. When a row line is asserted for a read operation, N4 can be viewed as an NMOS transistor in saturation. From Figure 4.2, it can be said that the operating current through N4 will decrease with



temperature. This decrease in operating current may result in an increased read access time. Similarly, during a write operation, when writing a “0”, N3 and N4 can be viewed as transistors in saturation and because of a reduction in the operating current at high temperatures, the write access time is increased.

During a hold operation, the row line is not asserted and the NMOS pass transistors are OFF and can be viewed as transistors in cut-off. From Figure 4.4, it is clear that the leakage current increases with temperature, but the leakage is still small enough that it will result in only a slight decrease in the HSNM. Therefore, leakage current has not been considered while calculating the HSNM.

#### 4.2.1 Baseline Design

Using the Baseline device geometries given in Table 3.1, the SRAM cell was simulated and the VTCs that give the HSNM, the RSNM and the WSNM at 27°C and 120°C are shown in Figures 4.7, 4.8 and 4.9.

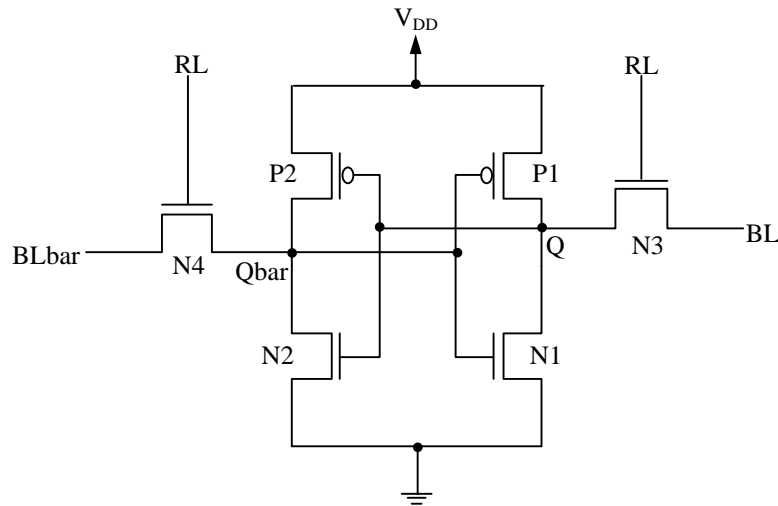


Figure 4.6 Transistor level diagram of a SRAM cell

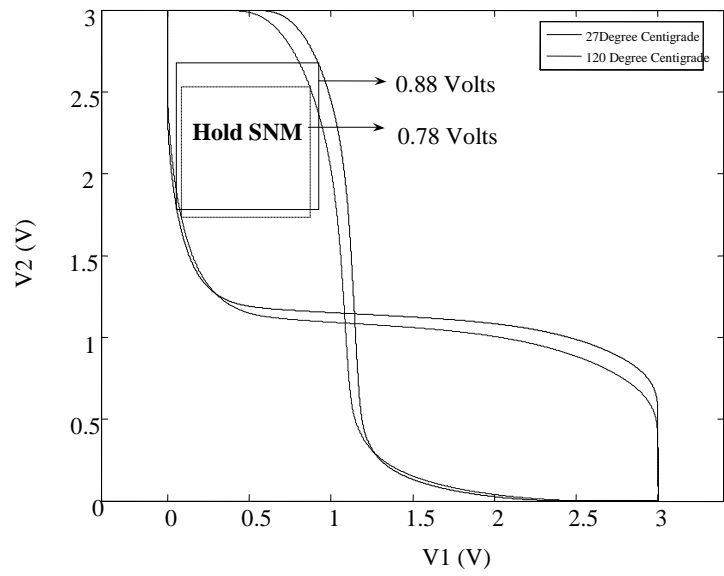


Figure 4.7 VTC showing HSNM for the Baseline Design at 27°C and 120°C

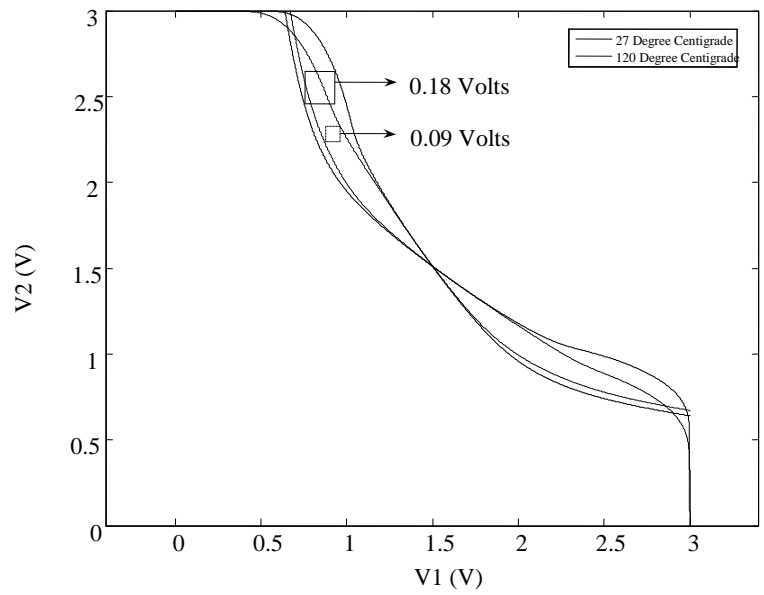


Figure 4.8 VTC showing the RSNM for the Baseline Design at 27°C and 120°C

At high temperatures, the threshold voltage of the NMOS transistor is smaller, so the VTC begins to fall from  $V_{DD}$  for a lower input voltage. The threshold voltage of the PMOS transistor is smaller, so the VTC begins to rise above 0.0 V for input voltage less than  $V_{DD}$ . The result is that the HSNM, RSNM and WSNM are all reduced. Figure 4.8 shows a comparison of the RSNM at 27°C and 120°C. The RSNM is 180 mV at 27°C and 90 mV at 120°C. This means that, at 120°C, a noise voltage of just 90 mV during a read operation could cause the SRAM cell to change state, resulting in a read upset. Therefore, the RSNM obtained for the Baseline Design at 120°C is not acceptable because there is a high probability for data corruption.

The Baseline SRAM cell was simulated to determine the read and write access times. A transient simulation was done by simulating an array of 64 rows and 16 columns

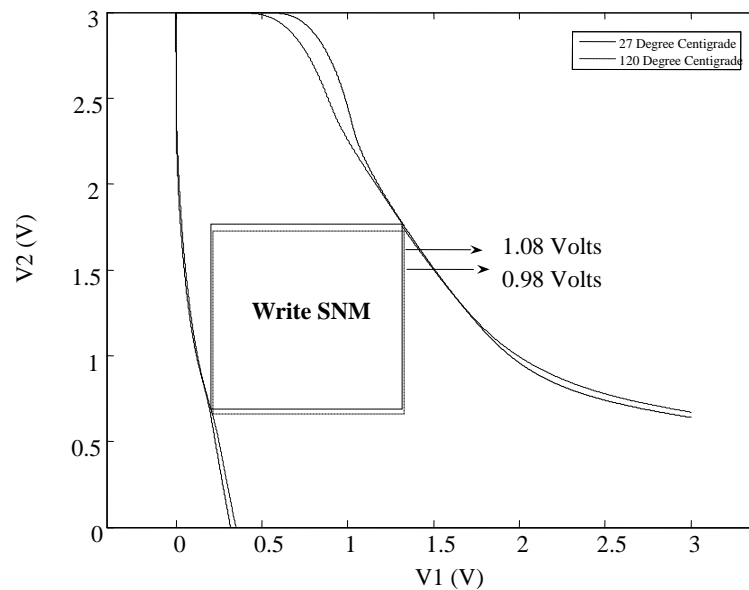


Figure 4.9 VTC showing WSNM for the Baseline Design at 27°C and 120°C

of SRAM cells along with the control circuit, pre-charge drivers, decoders, column multiplexer and sense amplifier. The 10-bit address input selected a particular row and a column, thus selecting an SRAM cell. First, a “1” was written into the selected SRAM cell by putting the data and its complement on the bit lines. Then, the data was read from the same SRAM cell.

Figure 4.10 shows a transient simulation of the array during the write and read operations at 27°C. The write time is the delay between a write request and successful data written into the cell. It is the time delay between the write enable signal (wen) going

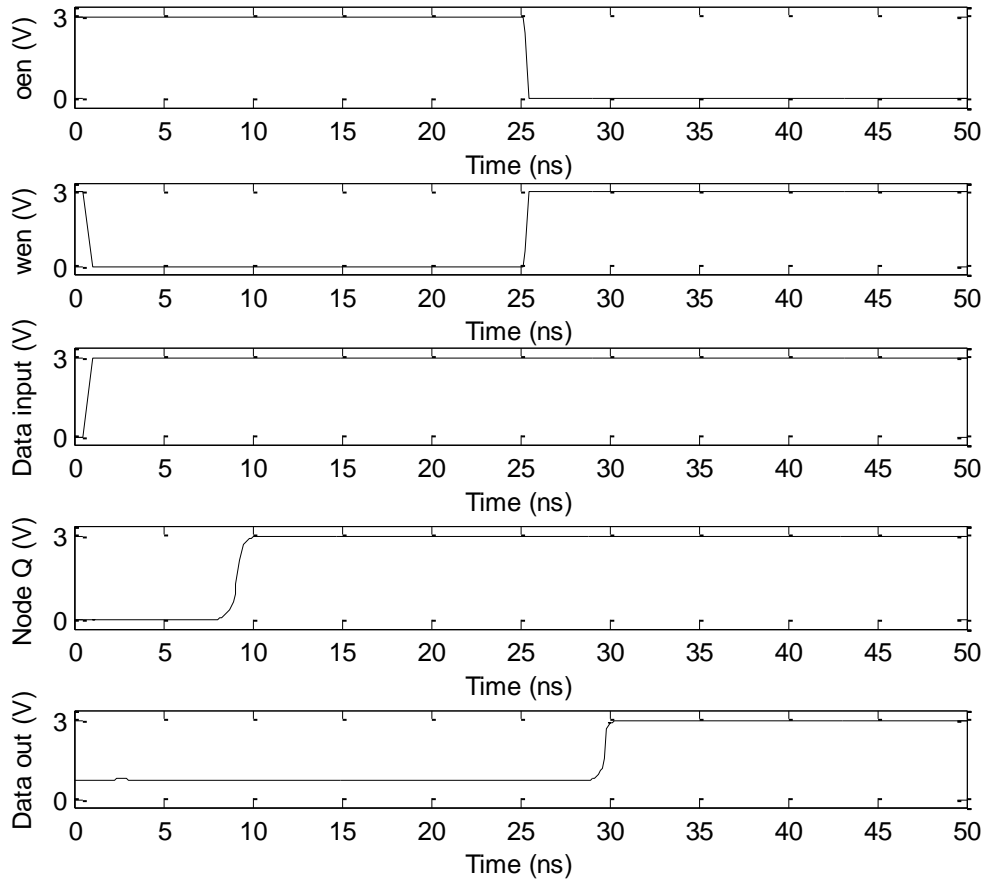


Figure 4.10 Simulation results showing write and read operations for the Baseline Design at 27°C

low (active) and the transition of the voltage at the storage node Q, or the time delay between a given valid address and a successful valid data written into the node Q of the cell assuming a low (active) wen.

The read access time is the delay between the read request and a successful data access on the data bus. It is the time delay between the output enable signal (oen) going low and a transition in the voltage (Data out) at the output of the tri-state inverter in the output block or the time delay between a given valid address and a successful data access

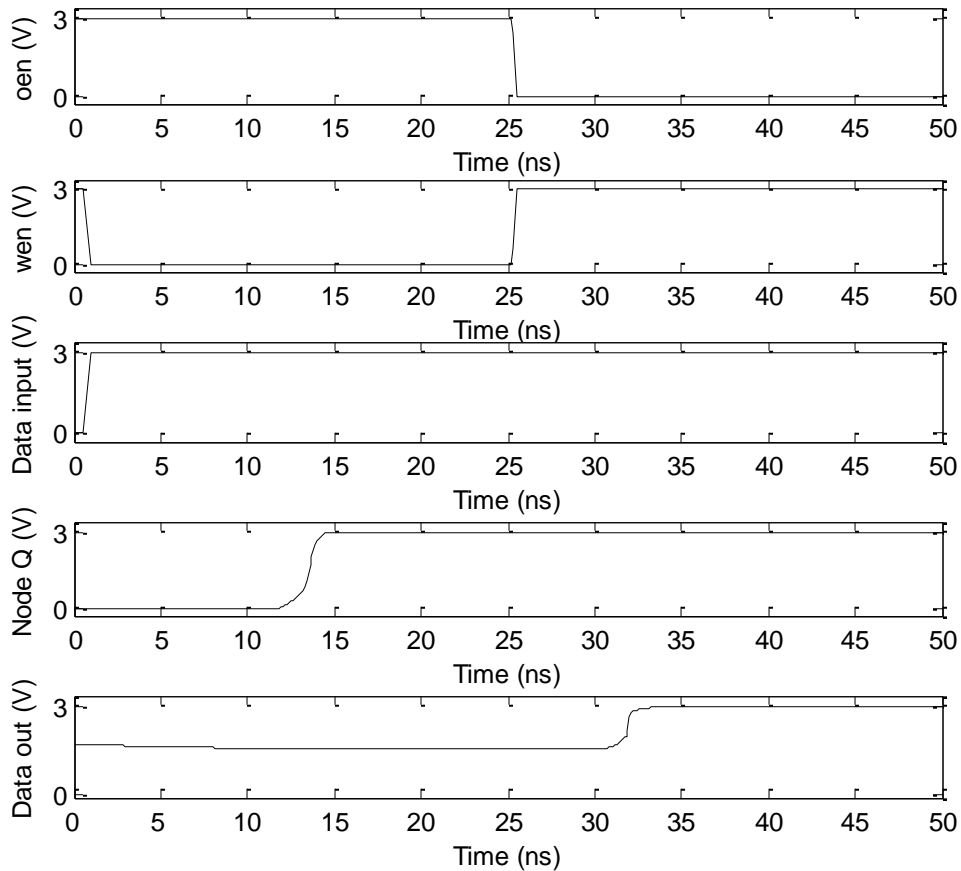


Figure 4.11 Simulation results showing write and read operations for the Baseline Design at 120°C

on the output of the tri-state inverter assuming a low (active)  $\text{oen}$ . In the following figures, Data input refers to the input data to the tri-state inverters in the input block and Data out is the output of the tri-state inverter in the output block. The Data out signal does not start from 0.0 V but takes some value midway between 0.0 V and 3.0 V, because when  $\text{se}$  is low (inactive), the output of the tri-state inverter is high impedance.

Throughout this work, the write and read access times are measured assuming the convention that a signal transition is considered to have occurred at the instant when the signal has moved 90% of the way from its initial value to its final value. For example, the write time for a “1” would be measured between the time when  $\text{wen}$  falls to 10% of  $V_{DD}$  and the time when node Q rises to 90% of  $V_{DD}$ . Using this convention, the write time and the read access time were calculated from the transient simulation results to be 9 ns and 4.5 ns at 27°C.

Similarly, the SRAM cell was simulated at 120°C, as shown in Figure 4.11, and the write time and the read access time were 13.2 ns and 6.5 ns. This shows that, because of a decrease in the operating current with temperature, the write and read access times have increased. The same simulation set up will be used in all the transient simulations in all the subsequent sections.

The SRAM cell is most vulnerable to unintended changes of state during a read operation. To assure that the SRAM cell is stable, the readability constraint has to be satisfied. As explained in Chapter 3, to meet the readability constraint, the rise in voltage at the storage node that stores a “0” must be less than the switching threshold of the other inverter to avoid a read upset.

Considering this, design modifications are done to get an improved RSNM at room temperature and at high temperature. At high temperatures, the threshold voltage of the transistors is reduced and the VTCs may come close to each other resulting in a very small RSNM. To avoid this situation, the device dimensions can be modified to give a better RSNM. This can be achieved by using a smaller access transistor and a larger pull down transistor. By doing so, the rise in voltage at the storage node Q can be limited and the read static noise margin can be improved.

When the size of the NMOS pass transistor is reduced, the RSNM of the SRAM cell increases at the cost of increased read access time and a decrease in the write static noise margin. During a write operation, the node Qbar that stores a “1” has to discharge through the NMOS pass transistor. By making the NMOS pass transistor weak (small  $W_L$ ), the current through the transistor is reduced, thereby decreasing the WSNM.

In the following three sub-sections, three modified SRAM cell designs are given. Each design uses the same PMOS pull-up and NMOS pull-down transistors. So, the HSNM for all the designs remains the same, and the VTCs that give the HSNM have not been shown. Each design uses successively smaller NMOS pass transistors, so that the load on the inverters during a read operation is reduced to result in a successive improvement in RSNM.

Table 4.1 Transistor channel dimension for the SRAM cell Modified Design I

Transistors	Channel Dimensions
	$W \times L$
P1, P2	1.4 $\mu\text{m} \times 0.5 \mu\text{m}$
N1, N2	6 $\mu\text{m} \times 0.8 \mu\text{m}$
N3, N4	3 $\mu\text{m} \times 0.8 \mu\text{m}$

#### 4.2.2 Modified Design I

Table 4.1 gives a modification to the device geometries that improves the RSNM at room temperature and at high temperature. The size of the NMOS pass transistor is reduced from  $4.8 \mu\text{m} \times 0.8 \mu\text{m}$  as in Baseline Design to  $3 \mu\text{m} \times 0.8 \mu\text{m}$ . The reduction in the size of the NMOS pass transistor reduces the load on the inverter during a read

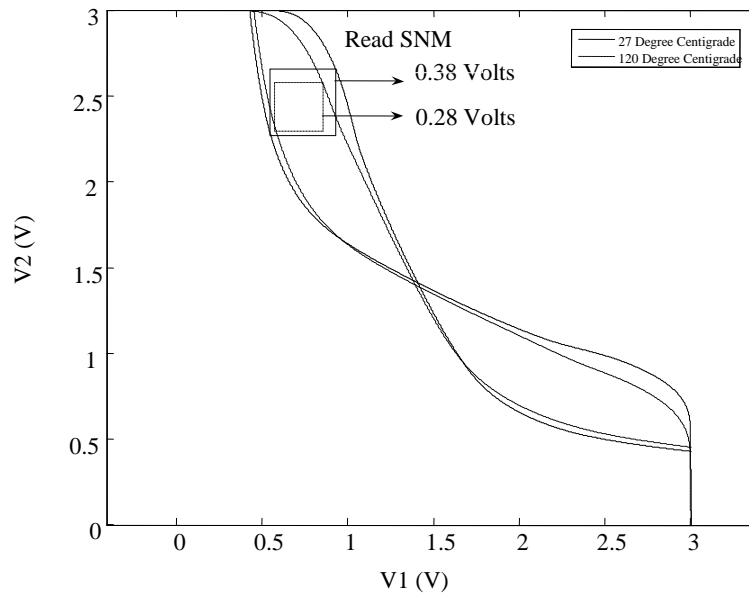


Figure 4.12 VTC showing read SNM for Modified Design I at 27°C and 120°C



operation. This results in a voltage transfer characteristic that gives an improved RSNM. The SRAM cell design with reduced NMOS pass transistors is called Modified Design I.

The VTCs that give the RSNM and the WSNM at 27°C and 120°C are shown in Figures 4.12 and 4.13. From Figure 4.12, the read static noise margin is 380 mV at 27°C and 280 mV at 120°C. There is an improvement of about 200 mV at 27°C and 190 mV at 120°C relative to the Baseline Design. However, the write static noise margin has decreased by 180 mV at 27°C and by 200 mV at 120°C relative to Baseline Design.

The SRAM cell Modified Design I was simulated to determine the write and read access times. Figure 4.14 shows a simulation of an SRAM cell during write and read operations at 27°C. The write and read access times were calculated to be 9.1 ns and 4.6 ns at 27°C. The write time increased by 0.1 ns at 27°C and at 120°C, respectively and the read access time increased by 0.1 ns at 27°C and at 120°C, relative to the Baseline Design.

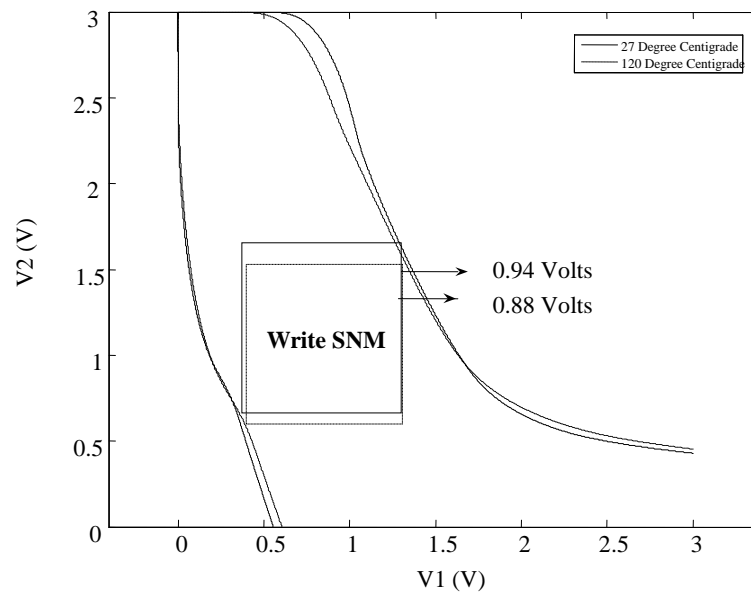


Figure 4.13 VTC showing the write SNM for Modified Design I at 27°C and 120°C

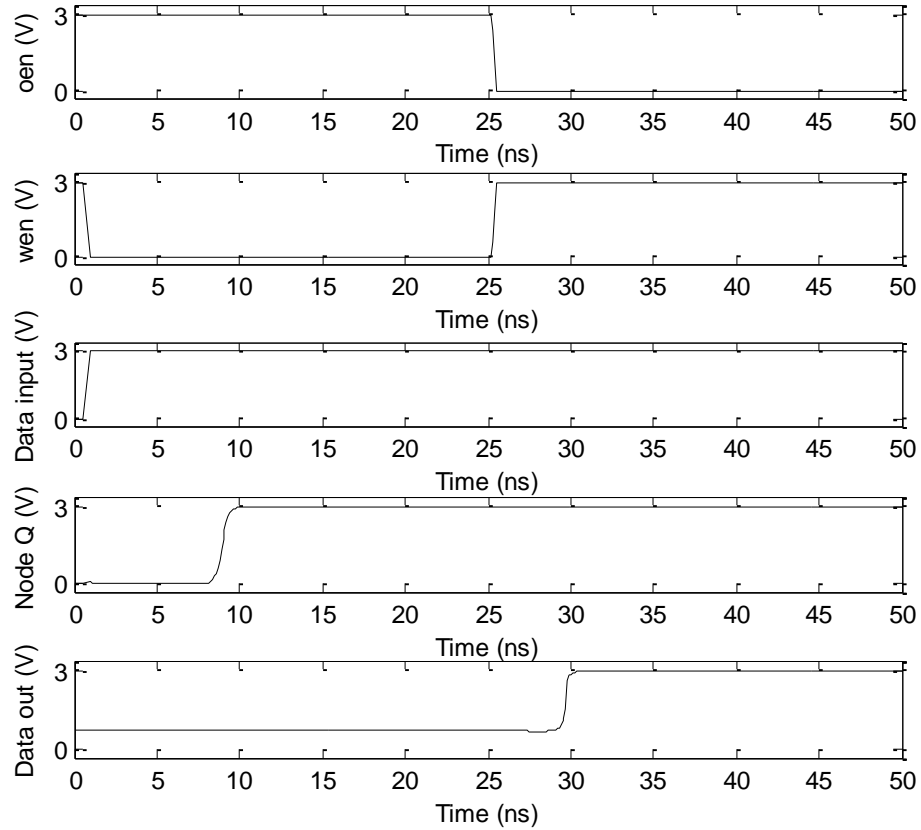


Figure 4.14 Simulation results showing write and read operations for the Modified Design I at 27 °C

The decrease in the width of the NMOS pass transistors in Modified Design I resulted in an increase in the RSNM at the cost of decreased WSNM and slightly increased write and read access times. This trend is shown again by the next two modified designs to be presented.

#### 4.2.3 Modified Design II

The read static noise margin was further improved by further decreasing the size of the NMOS pass transistors. In this design, called Modified Design II, the size of the

NMOS pass transistors was reduced to  $2.4 \mu\text{m} \times 0.8 \mu\text{m}$ . The device geometries for the Modified Design II are shown in Table 4.2.

Figures 4.16 and 4.17 show the overlapped VTCs that give RSNM and WSNM for Modified Design II at  $27^\circ\text{C}$  and  $120^\circ\text{C}$ . From Figure 4.16, the read static noise margin is 470 mV at  $27^\circ\text{C}$  and 370 mV at  $120^\circ\text{C}$ . There is an improvement of about 90 mV at  $27^\circ\text{C}$  and  $120^\circ\text{C}$  relative to Modified Design I. However, the write static noise margin has decreased by 90 mV at  $27^\circ\text{C}$  and by 100 mV at  $120^\circ\text{C}$  relative to Modified Design I.

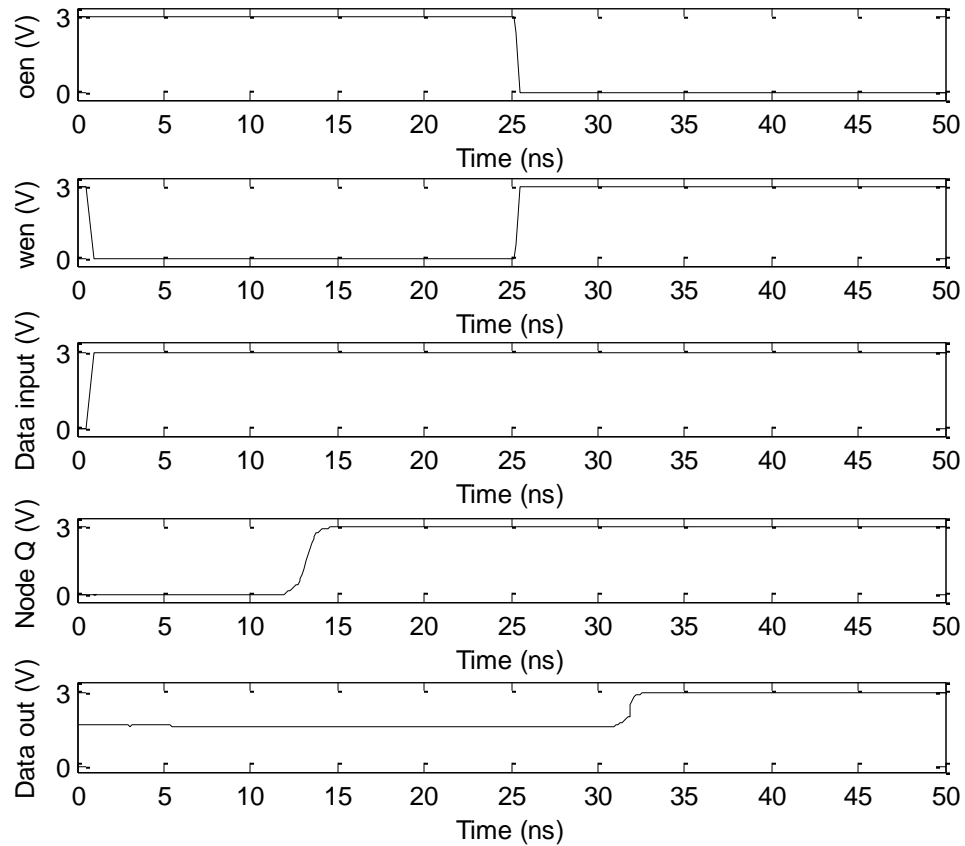


Figure 4.15 Simulation results showing write and read operations for the Modified Design I at  $120^\circ\text{C}$

Table 4.2 Transistor channel dimension for the SRAM cell Modified Design II

Transistors	Channel Dimensions $W \times L$
P1, P2	1.4 $\mu\text{m} \times 0.5 \mu\text{m}$
N1, N2	6 $\mu\text{m} \times 0.8 \mu\text{m}$
N3, N4	2.4 $\mu\text{m} \times 0.8 \mu\text{m}$

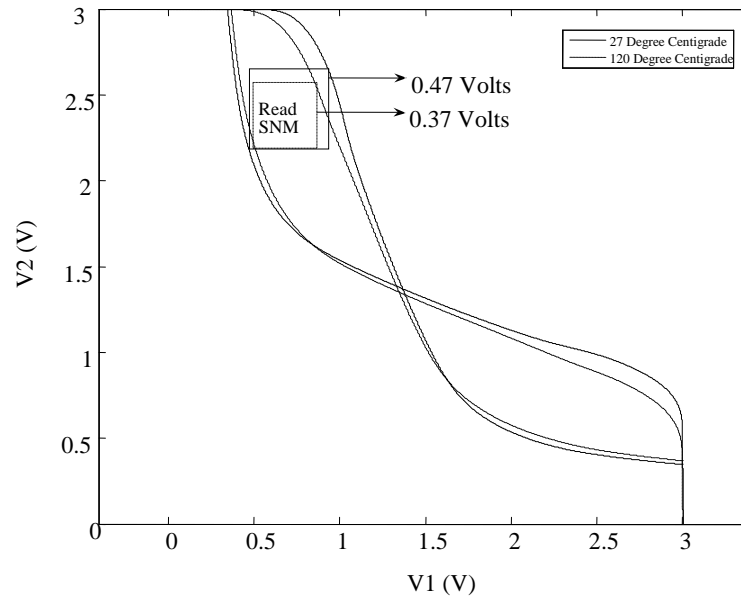


Figure 4.16 VTC showing read SNM for Modified Design II at 27°C and 120°C

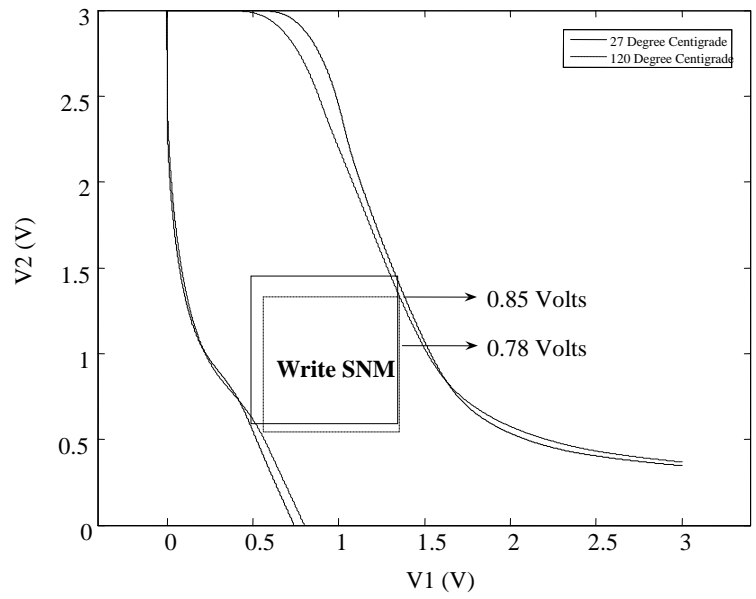


Figure 4.17 VTC showing write SNM for Modified Design II at 27°C and 120°C

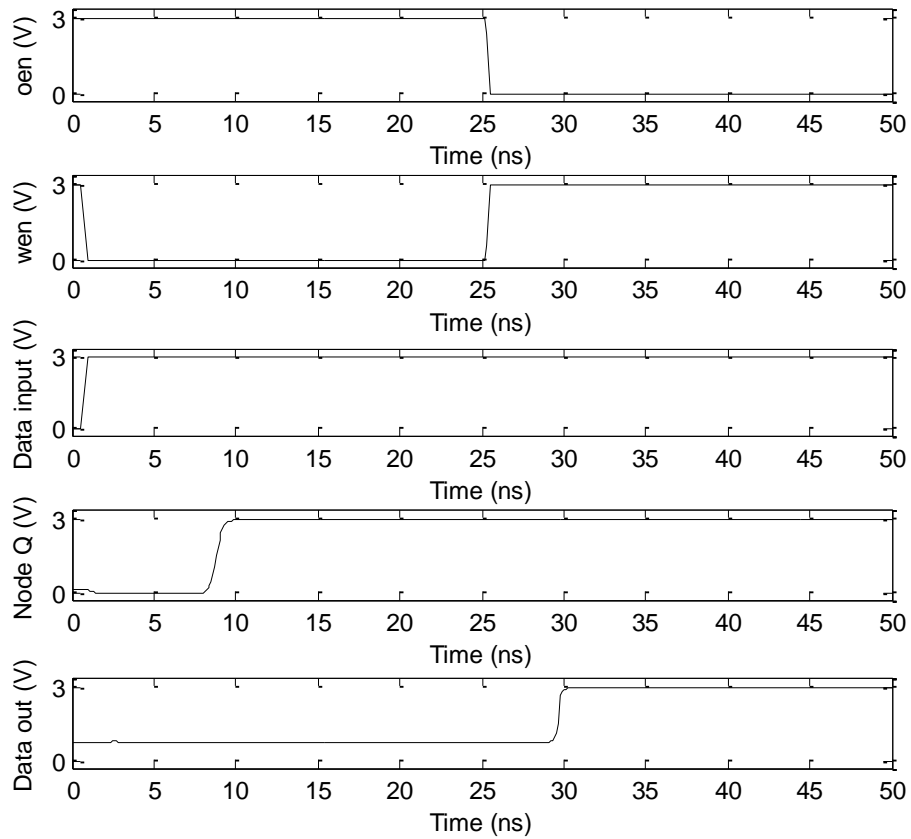


Figure 4.18 Simulation results showing write and read operations for the Modified Design II at 27 °C

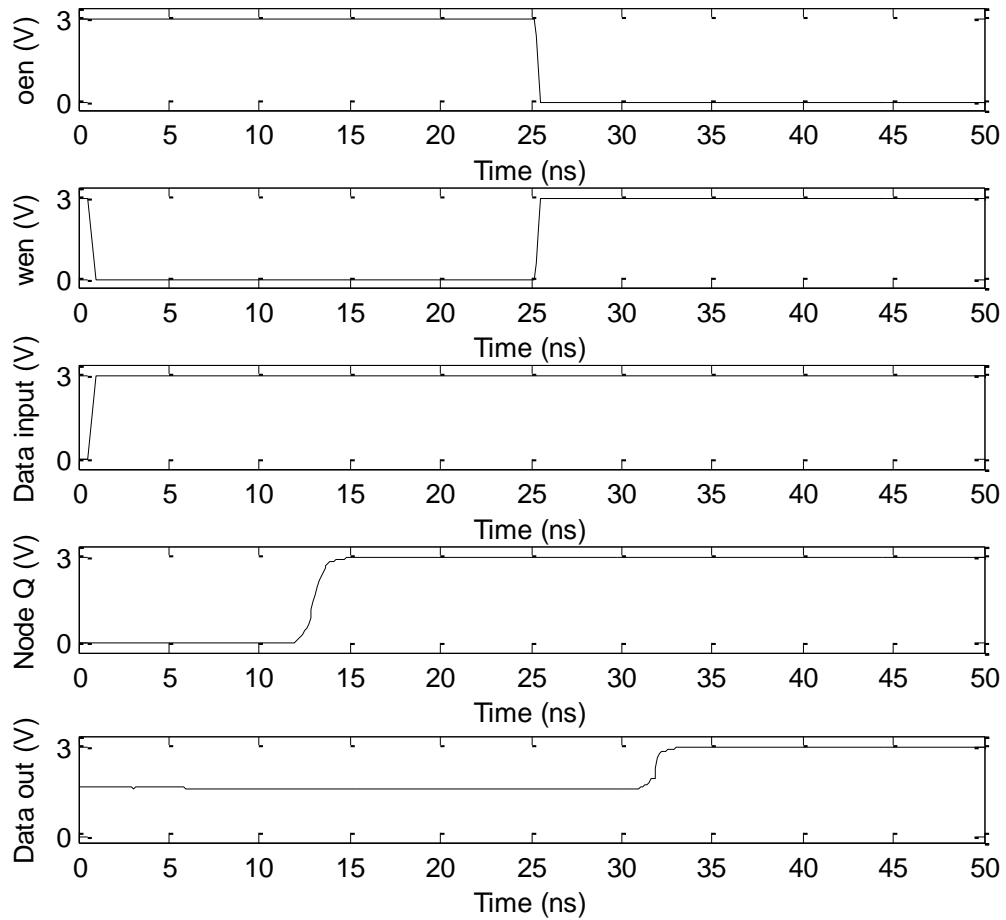


Figure 4.19 Simulation results showing write and read operations for the Modified Design II at 120°C

Similar to the simulation described in the previous section, the SRAM cell Modified Design II was simulated to determine the write and read access times. The simulations at 27°C and at 120°C are shown in Figures 4.18 and 4.19. The write and read access times were calculated to be 9.4 ns and 4.7 ns at 27°C, and 13.8 ns and 6.7 ns at 120°C. The write time increased by 0.3 ns at 27°C and by 0.5 ns at 120°C, respectively, and the read access time increased by 0.1 ns at 27°C and at 120°C, relative to the Modified Design I.

Table 4.3 Transistor channel dimensions for the SRAM cell Modified Design III

Transistors	Channel Dimensions $W \times L$
P1, P2	1.4 $\mu\text{m} \times 0.5 \mu\text{m}$
N1, N2	6 $\mu\text{m} \times 0.8 \mu\text{m}$
N3, N4	2.0 $\mu\text{m} \times 0.8 \mu\text{m}$

#### 4.2.4 Modified Design III

In this design, the width of the NMOS pass transistors is further reduced to 2.0  $\mu\text{m}$ . The device geometries for the Modified Design III are given in Table 4.3. The VTCs that give the RSNM and WSNM at 27°C and 120°C are shown in Figures 4.20 and 4.21, respectively. From Figure 4.20, the read static noise margin is 520 mV at 27°C and 420 mV at 120°C. There is an improvement of about 50 mV at 27°C and 120°C relative to Modified Design II. However, the write static noise margin has decreased by 130 mV at 27°C and 120°C relative to Modified Design II.

Similar to the simulation described in the previous section, the SRAM cell Modified Design III was simulated to determine the write and read access times. The simulations at 27°C and at 120°C are shown in Figures 4.22 and 4.23. The write and read access times were calculated to be 10.0 ns and 4.8 ns at 27°C, and 15.5 ns and 6.8 ns at 120°C. The write time increased by 0.6 ns and 1.7 ns at 27°C and at 120°C, respectively, and the read access time increased by 0.1 ns at 27°C and at 120°C, relative to the Modified Design II.

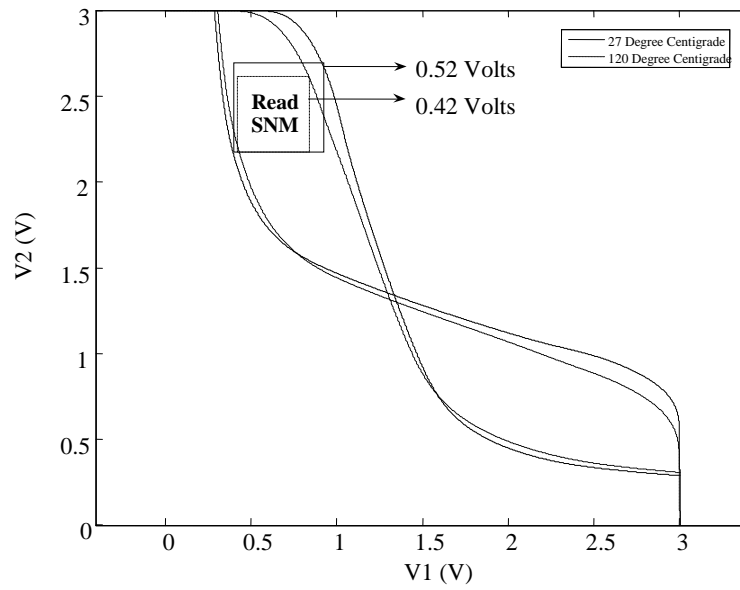


Figure 4.20 VTC showing read SNM for Modified Design III at 27°C and 120°C

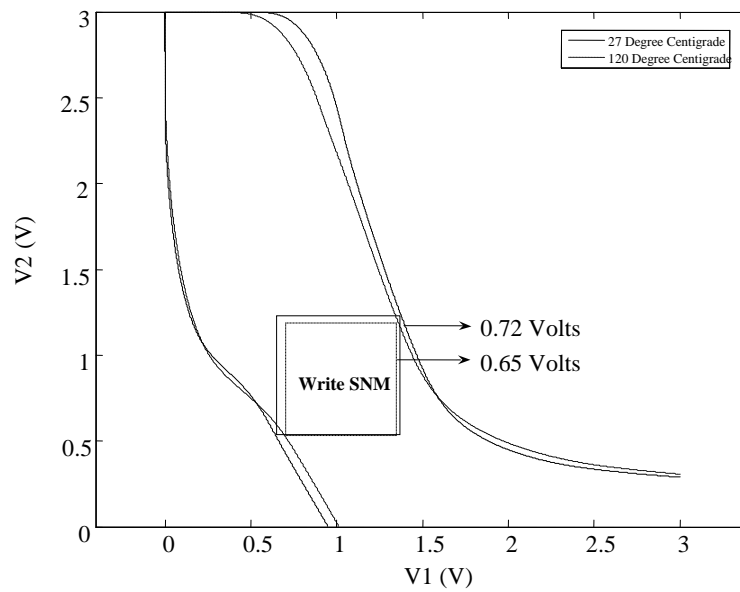


Figure 4.21 VTC showing write SNM for Modified Design III at 27°C and 120°C



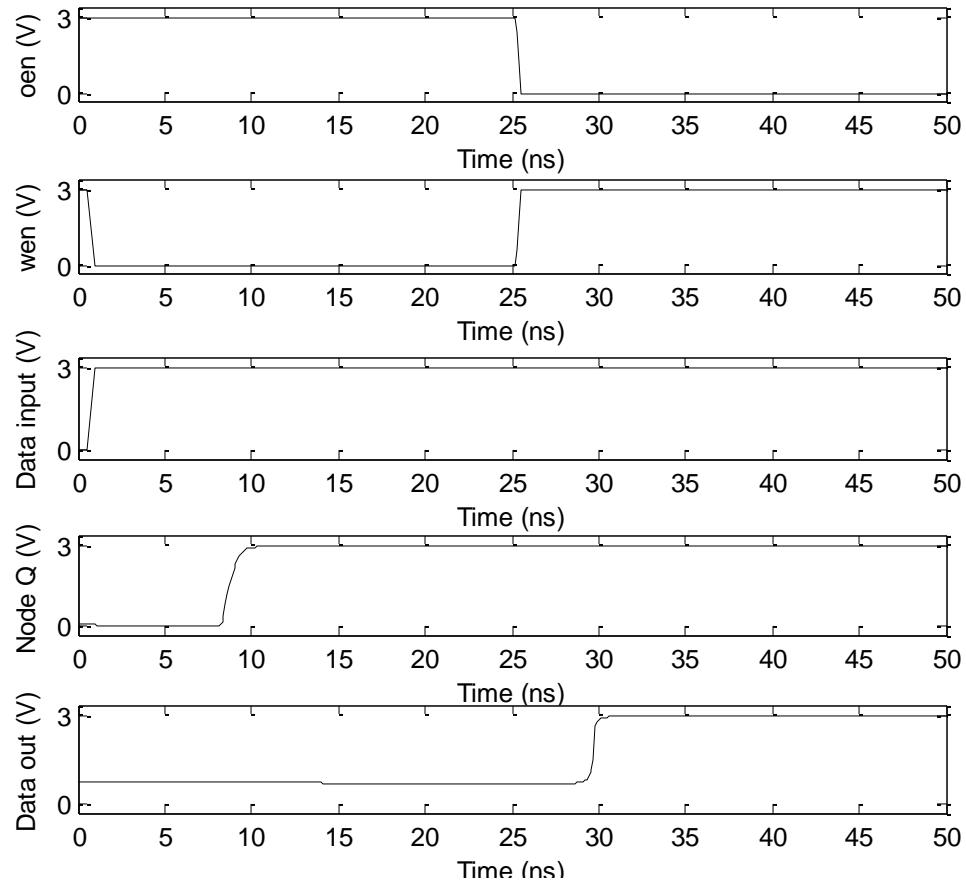


Figure 4.22 Simulation results showing write and read operations for the Modified Design III at 27°C

For the four designs presented, the decrease in the size of the NMOS pass transistors resulted in an increase in the RSNM at the cost of decreased WSNM and slightly increased write and read access times. Further reducing the size of the NMOS pass transistors would further improve the RSNM; however, later in this chapter it can be observed that the improvement in RSNM is at the cost of a steep increase in the write time at 120°C. The following section examines the trade-offs among these SRAM cell characteristics.

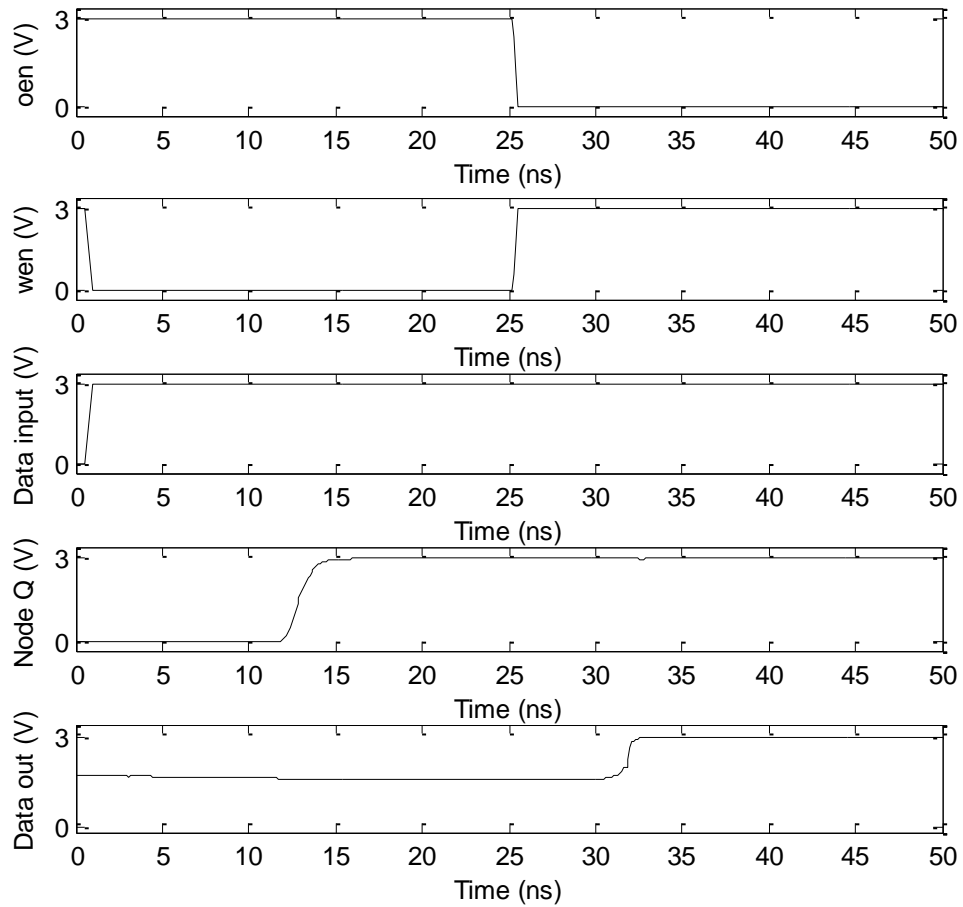


Figure 4.23 Simulation results showing write and read operations for the Modified Design III at 120°C

#### 4.3 Trade-offs among Stability Margins and Access Times

Four designs are presented: a Baseline Design and three modified designs, each with progressively smaller NMOS pass transistors. Decreasing the size of the NMOS pass transistors increases CR and PR. Increase in CR results in smaller load on the inverters during a read operation. Therefore, RSNM is improved. Increase in PR makes the write operation difficult which results in a decrease in WSNM. HSNM is not affected

Table 4.4 Static noise margins for the SRAM cell for the four designs at 27°C & 120°C

Design	Hold SNM (Volts)		Read SNM (Volts)		Write SNM (Volts)	
	27°C	120°C	27°C	120°C	27°C	120°C
Baseline Design CR= 1.25, PR= 0.47	0.88	0.78	0.18	0.09	1.08	0.98
Modified Design I CR= 2.00, PR= 0.75	0.88	0.78	0.38	0.28	0.94	0.88
Modified Design II CR= 2.50, PR= 0.93	0.88	0.78	0.47	0.37	0.85	0.78
Modified Design III CR= 3.00, PR= 1.12	0.88	0.78	0.52	0.42	0.72	0.65

Table 4.5 SRAM access times for the four designs at 27°C & 120°C

Design	Write Time (ns)		Read Access Time (ns)	
	27°C	120°C	27°C	120°C
Baseline Design	9.0	13.2	4.5	6.5
Modified Design I	9.1	13.3	4.6	6.6
Modified Design II	9.4	13.8	4.7	6.7
Modified Design III	10.0	15.5	4.8	6.8

because the sizes of the PMOS pull-up transistor and the NMOS pull down transistor are fixed to  $1.4 \mu\text{m} \times 0.5 \mu\text{m}$  and  $6 \mu\text{m} \times 0.8 \mu\text{m}$  respectively. Table 4.4 gives the HSNM, RSNM and WSNM values for the Baseline Design, Modified Design I, Modified Design II and Modified Design III at  $27^\circ\text{C}$  and  $120^\circ\text{C}$ .

Figure 4.24 shows the variation of RSNM with the cell ratio. It can be observed that RSNM increases with the cell ratio. As the cell ratio increases, the loading on the inverter during a read operation decreases, resulting in an increased RSNM.

The variation of WSNM with PR is shown in Figure 4.25. It is observed that as PR increases, WSNM decreases. This is because, as the size of the NMOS pass transistor is reduced, it gets more difficult to write the data.

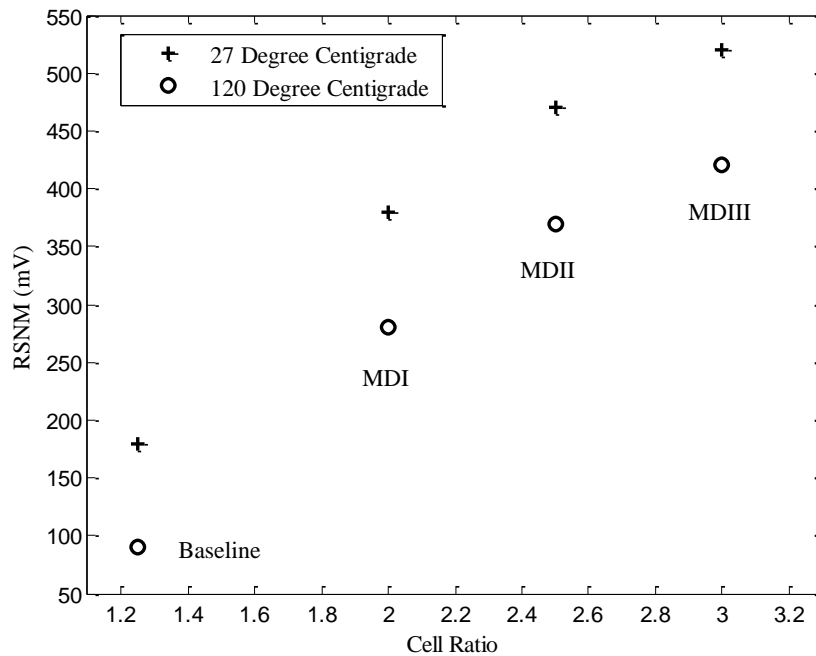


Figure 4.24 Variation of RSNM with cell ratio at  $27^\circ\text{C}$  and  $120^\circ\text{C}$

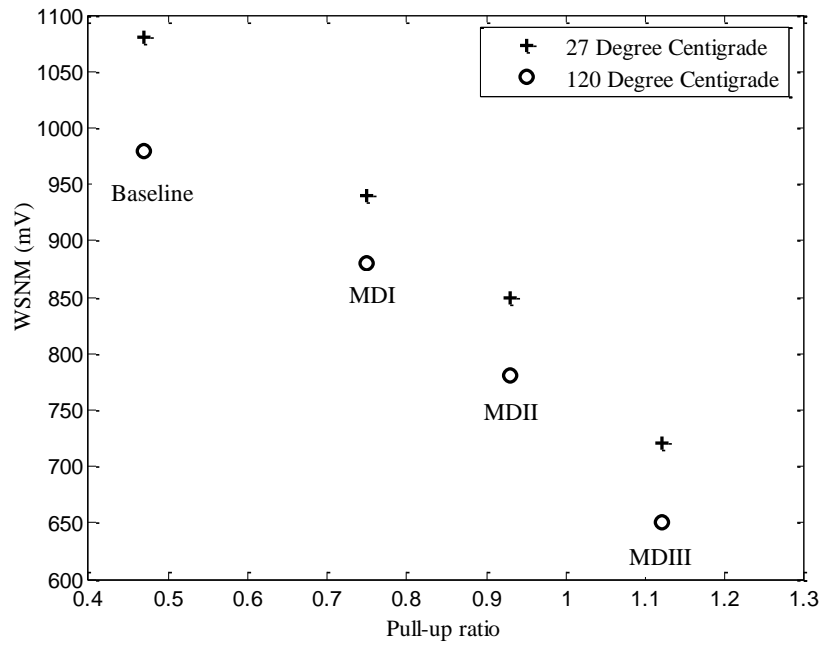


Figure 4.25 Variation of WSNM with pull-up ratio at 27°C and 120°C

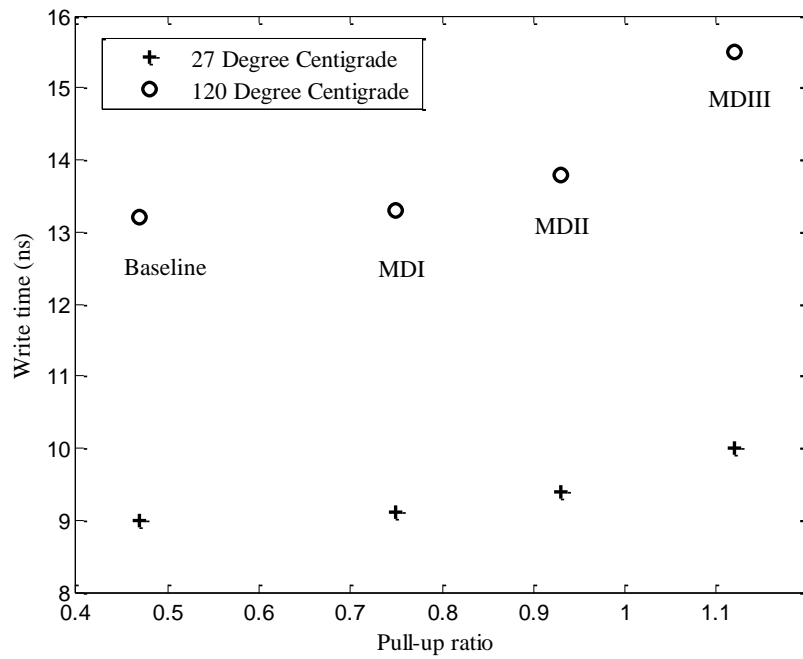


Figure 4.26 Variation of write time with pull-up ratio at 27°C and 120°C

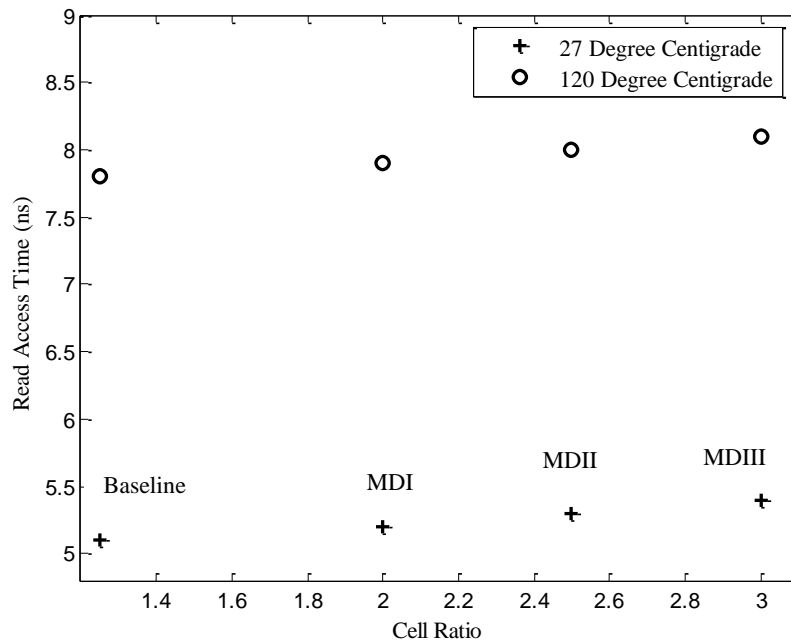


Figure 4.27 Variation of read access time with cell ratio at 27°C and 120°C

Table 4.5 gives the read access time and the write time for the four designs at 27°C and 120°C. The effect of the change in PR and CR on the write time and the read access time is observed. The variation of the write time with the PR is shown in Figure 4.26. It is observed that as the pull-up ratio is increased, the write time also increases. This is because, with a decrease in the size of the NMOS pass transistor, it takes longer for the node that stores a “1” to discharge resulting in an increased write time.

Figure 4.27 shows the variation of the read access time with the cell ratio. As the size of the NMOS pass transistors is reduced, it takes longer for the bit line to discharge during a read operation, resulting in an increased read access time. It can be observed from Figure 4.26 and Figure 4.27 that the variation in the read access time is less than the variation in the write time. This is because read operation is a low-swing event, which

means that even a small change on the bit lines is sufficient to get a valid output. By contrast, the write is a full-swing operation; therefore, larger effect is observed in the write time.

The main goal of the design modifications presented has been to improve the RSNM of the SRAM cell. To illustrate the trade-offs involved in these modifications, the SRAM characteristics are now plotted as functions of the RSNM. The variation of read access time with RSNM is shown in Figure 4.28. It is observed that RSNM is improved with only a slight increase in read access time. For the designs considered, this trade-off does not seem to be significant.

The variation of the write time with RSNM is shown in Figure 4.29. It is observed that as RSNM increases the write time also increases. For the designs considered, this

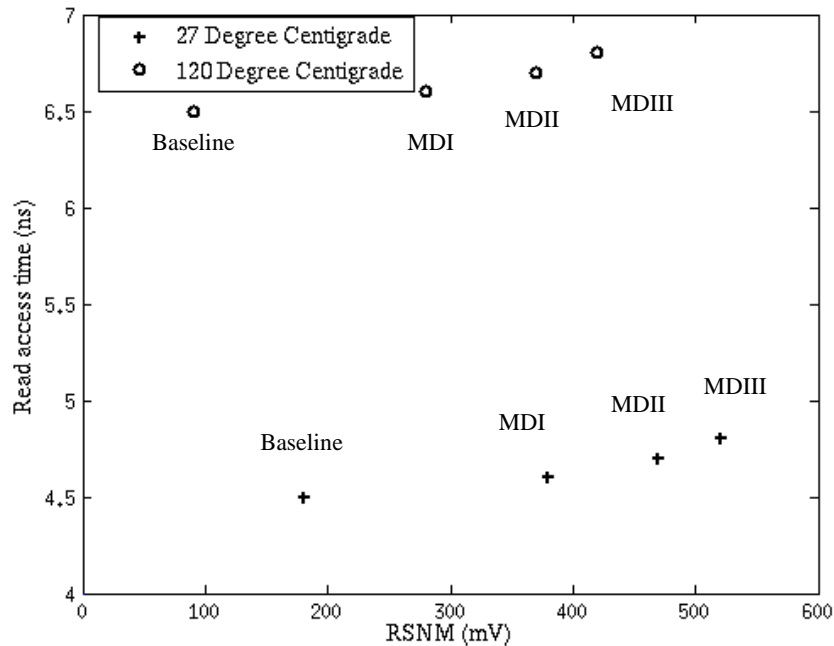


Figure 4.28 Variation of read access time with RSNM at 27°C and 120°C

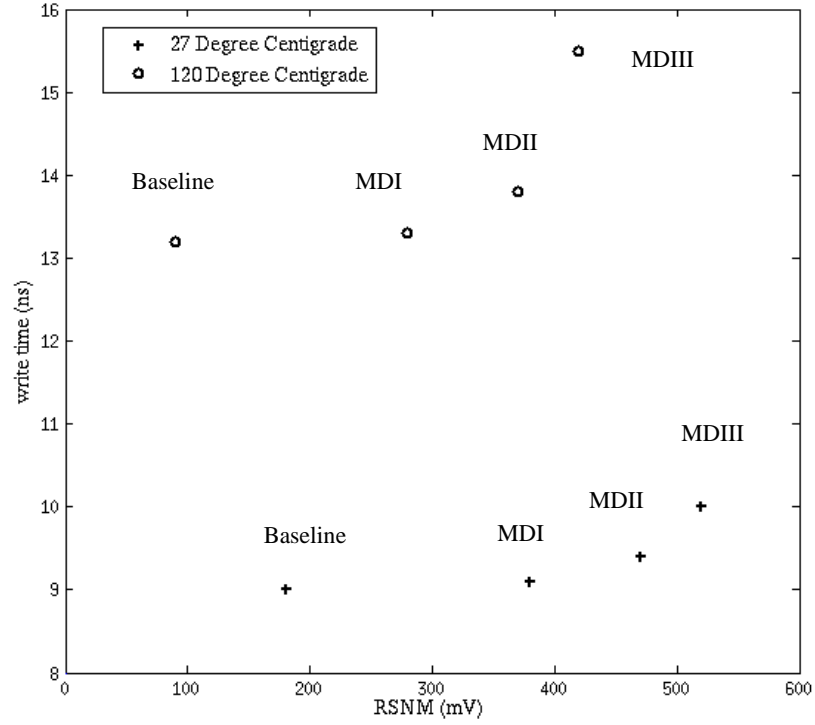


Figure 4.29 Variation of write time with RSNM at 27°C and 120°C

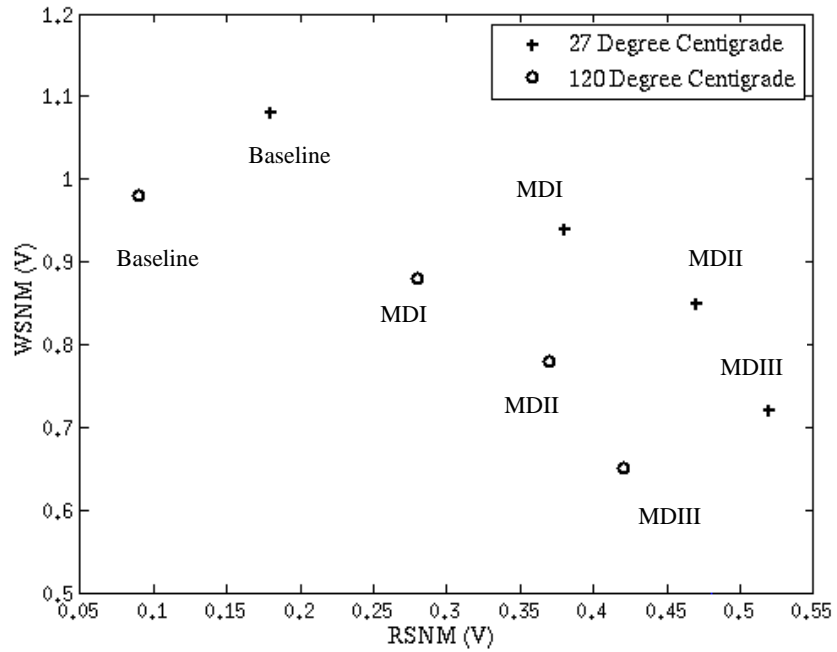


Figure 4.30 Variation of WSNM with RSNM at 27°C and 120°C



trade-off does not seem to be significant at 27°C but the write time becomes more sensitive to even a small improvement in RSNM at 120°C past Modified Design II. It may not be worth improving the RSNM after a certain point and so from Figure 4.29 Modified Design II could be considered over Modified Design III.

Figure 4.30 shows the variation of WSNM with RSNM. With the decrease in the width of the NMOS pass transistor with each design, the RSNM improved at the cost of decreased WSNM. At 27°C there is a 0.34 V increase in RSNM and 0.36 V decrease in the WSNM from Baseline Design to Modified Design III. At 120°C there is a 0.33 V increase in RSNM and 0.33 V decrease in WSNM from Baseline Design to the Modified Design III. At 27°C the increase in the RSNM is less than the decrease in the WSNM and at 120°C the increase in RSNM and the decrease in WSNM are almost proportional.

This chapter covers the stability analysis of the SRAM cell. Four different designs have been presented that show the improvement in RSNM with every design. It is also observed that the WSNM, the read access time and the write time are all degraded when the RSNM is improved. There is a direct trade-off between RSNM and WSNM, both at 27°C and at 120°C, throughout the range of designs considered. The trade-off between RSNM and write time is only significant at high temperature, and only for the highest values of RSNM. The trade-off between RSNM and read access time does not seem significant over the range of designs and temperatures considered.

## CHAPTER V

### CONCLUSIONS AND FUTURE WORK

This work concentrated on the design of a 1Kword SRAM in SOI CMOS that is capable of stable operation over an extended temperature range. The design was simulated for temperatures up to 120°C, which is the highest temperature for which the transistor models have been validated. The main contribution of the thesis is a study of the trade-offs in the design of a high-temperature asynchronous SRAM in an SOI CMOS process suitable for a mixed-signal application-specific integrated circuit. The design could be used in a sensor interface with a digital back end. It is observed that the WSNM, the read access time and the write times are all degraded when the RSNM is improved. There is a direct trade-off between RSNM and WSNM, both at 27°C and at 120°C, throughout the range of designs considered. The trade-off between RSNM and write time is only significant at high temperatures, and only for the highest values of RSNM. The trade-off between RSNM and read access time does not seem significant over the range of designs and temperatures considered.

A standard six-transistor SRAM cell is used to hold each bit of data. The stability margins for the SRAM cell, that is, the HSNM, the RSNM and the WSNM, are obtained for a Baseline Design at room temperature and at high temperatures. The Baseline Design

was shown to be suitable at room temperature, but analysis showed reduced stability margins at high temperatures. Therefore, three redesigns were done, the Modified Design I, the Modified Design II and the Modified Design III. With every design, improved RSNM was achieved by reducing the size of the NMOS pass transistors in the SRAM cell. The most stable design has RSNM of 520 mV at 27°C and 420 mV at 120°C, and WSNM of 720 mV at 27°C and 650 mV at 120°C respectively. The SRAM was designed for acceptable RSNM over temperature at the expense of slightly increased write and read access time with respect to the Baseline Design.

This work can be extended to design a low-power SRAM using the strongest design here, the Modified Design III, as a starting point. To achieve a low-power design, one could use high threshold voltage devices in the SRAM cell; this would help in reducing the leakage currents, thereby lowering the power consumption [2]. As presented in [2] different six-transistor SRAM cell configurations corresponding to different threshold voltages can be used in the SRAM array to reduce the power consumption without altering the architecture.

Process parameters such as process transconductance and the threshold voltages of the transistors can vary, with different values for different fabrication runs. These process variations can affect timing of circuitry; they can also affect the stability of the six-transistor SRAM cell. Models provided by the fabrication facility allow for simulating the design over process variation to understand better how the fabricated integrated circuits will operate. As a next step, simulations over process variations should be done, to further verify the quality of the design before it is fabricated.

In this thesis, parasitic resistance and capacitance of the read word line, bit lines, column select lines, pre-charge signal, sense enable signal were not considered during the analysis. Once the layout is done, capacitance and resistance extraction tools could be used to get accurate R and C values that could be included in the simulations to obtain more accurate read access time and write times, and to ensure the correct timing of the control signals.

The interconnect wire width must be carefully designed so as not to affect the overall reliability of design [7]. If a wire is too narrow for the current it is carrying, the high current density might cause internal failure or a void in the wire or an open circuit. To avoid this, the design could be tested to observe if there are any interconnect wires that might be affected by electromigration over time. Once a layout is done, available tools [1] could be used to check the SRAM design for susceptibility to electromigration.

## REFERENCES

- [1] S. M. Alam, G. C. Lip, C. V. Thompson, D. E. Troxel, "Circuit Level Reliability Analysis of Cu Interconnects," *Fifth International Symposium on Quality Electronic Design*, pp. 238-243, San Jose, CA, August 2004.
- [2] B. Amelifard, F. Fallah, M. Pedram, "Leakage Minimization of SRAM Cells in a Dual- $V_t$  and Dual- $T_{ox}$  Technology," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Volume 16, Issue 7, July 2008.
- [3] S. V. Avadhanam, "Design of a Flexible High Temperature SRAM with Reduced Design Time," M.S. Thesis, Oklahoma State University, 2007.
- [4] P. G. Drennan, M. L. Kniffin, D. R. Locascio, "Implications of Proximity Effects for Analog Design," *Proceedings of the IEEE Custom Integrated Circuits Conference (CICC)*, pp. 169-176, San Jose, CA, September 2006.
- [5] B. Erbađci, "High Temperature Processor and Memory Implementation," M.S. Thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2011.
- [6] J. R. Hauser, "Noise Margin Criteria for Digital Logic Circuits," *IEEE Transactions on Education*, Volume 36, Issue 4, pp. 363-368, 2002.
- [7] J. P. Heliot, L. Tullos, "Call for Improved Electromigration Simulation Tool," *International SEMATECH*, Technology Transfer #01084148A-XFR, August 2001.
- [8] J. A. Henfling, R. A. Normann, "High Temperature Downhole Reservoir Monitoring System," *Twenty-Ninth Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, California, January 26-28, 2004.
- [9] A. Marshall, S. Natarajan, *SOI Design: Analog, Memory and Digital Techniques*, Springer Netherlands, 2001.
- [10] D. Mukherjee, H. R. Mondal, B. V. R. Reddy, "Static Noise Margin Analysis of SRAM Cell for High Speed Application," *International Journal of Computer Science Issues*, Volume 7, Issue 5, September 2010.

- [11] J. M. O'Connor, J. C. Tsang, J. B. McKitterick, "225 C High Temperature Silicon-on-Insulator (SOI) ASICs for Harsh Environments," *IEEE International Workshop on Integrated Power Packaging*, pp. 2-5, Chicago, IL, September 1998.
- [12] J. M. Rabaey, F. De Bernardinis, A. M. Niknejad, B. Nikolic, A. Sangiovanni-Vincentelli, "Embedding Mixed-Signal Design in Systems-on-Chip," *Proceedings of the IEEE*, Volume 94, pp. 1070-1088, 2006.
- [13] J. M. Rabaey, B. Nikolic, A. P. Chandrakasan, *Digital Integrated Circuits: A Design Perspective*, 2<sup>nd</sup> Edition, Prentice-Hall, 2002.
- [14] E. Seevinck, F. J. List, J. Lohstroh, "Static-Noise Margin Analysis of MOS SRAM cells," *IEEE Journal of Solid-State Circuits*, Volume 22, pp. 748-754, 1987.
- [15] H. Singh, "Analysis of SRAM Reliability under Combined Effect of NBTI, Process and Temperature Variations in Nano-Scale CMOS," *IEEE Fifth International Conference on Future Information Technology (Future Tech)*, pp. 1-4, Busan, May 2010.
- [16] J. Wang, S. Nalam, B. H. Calhoun, "Analyzing Static and Dynamic Write Margin for Nanometer SRAMs," *ACM / IEEE International Symposium on Low Power Electronics and Design (ISLPED)*, pp. 129-134, Bangalore, August 2008.
- [17] N. Weste, D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*, 3<sup>rd</sup> Edition, Addison Wesley, 2004.