

國立交通大學

電子工程學系 電子研究所碩士班

碩 士 論 文

穩健次臨界靜態隨機存取記憶體與超低功率先進先出



Robust Subthreshold SRAM and Ultra-Low Power FIFO

Memory Design

研 究 生：張牧天

指 導 教 授：黃 威 教 授

中 華 民 國 九 十 七 年 六 月

穩健次臨界靜態隨機存取記憶體與超低功率先進先出
記憶體設計

Robust Subthreshold SRAM and Ultra-Low Power FIFO
Memory Design

研究生：張牧天

Student : Mu-Tien Chang

指導教授：黃 威 教授

Advisor : Prof. Wei Hwang



A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics

College of Electrical Engineering and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Electronics Engineering

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

摘要

本論文提出高穩定性的超低功率靜態隨機存取記憶體及先進先出記憶體設計。超低電壓能夠有效地減少功率消耗，但靜態隨機存取記憶體也會因著電壓的下降而變得不穩健。本論文首先提出一個高穩定、全差分、單埠靜態隨機存取記憶體。利用自我補償及雜訊阻隔的機制，該記憶體能夠穩定地工作於超低電壓，而支援寫入的機制更可以提高於超低工作電壓時的寫入能力。利用UMC 90nm CMOS技術，模擬結果顯示，在0.2V工作電壓下，所提出的次臨界隨機存取記憶體與傳統的靜態隨機存取記憶體相比，有1.22倍的資料維持穩定性提升，2.09倍的讀取穩定性提升，及2.03倍的寫入能力提升。

除此之外，爲了進一步減少漏電流，本論文利用自我電壓控制以及電源阻斷技術來降低先進先出記憶體陣列的漏電流；同時提出了一個高穩定性的雙臨界電壓、雙埠記憶體來增加可靠度並且減少功率消耗。模擬結果顯示，所提出的先進先出記憶體與傳統先進先出模組相比，具有多達94%的功率削減。利用UMC 90nm CMOS技術，所提出的256行X16位元先進先出記憶體，於0.5V，5MHz讀取頻率，200kHz寫入頻率，有2.21uW的功率消耗。



Abstract

Subthreshold SRAM and ultra-low power FIFO memory are indispensable to energy-constrained SoC. The stability of SRAM cell, however, has always been a major challenge to subthreshold SRAM design. This thesis proposes a robust, fully-differential subthreshold 10-transistor SRAM cell with auto-compensation. With the auto-compensation mechanism, the proposed cell exhibits better hold static noise margin (SNM). The cell structure also prevents storage nodes from bitline noise interference, thus improving read SNM. Better write ability is achieved by applying the write-assist technique. Based on UMC 90nm CMOS technology, simulation results show that, at 0.2V supply voltage, the proposed cell has 1.22X hold SNM improvement, 2.09X read SNM improvement, and 2.03X write margin improvement when compared to the conventional 6T SRAM cell.

This thesis, in addition, also proposes a robust, ultra-low power asynchronous FIFO memory. With the self-adaptive power control and the complementary power gating technique, leakage power of the FIFO memory array is minimized. Further, the stability of the FIFO memory is improved under ultra-low supply voltage supply with the proposed dual- V_T 7T SRAM cell. Simulation results indicate that the proposed scheme has up to 94% power reduction over conventional designs. In this thesis, the proposed FIFO is implemented in UMC 90nm CMOS technology under 0.5V supply voltage, with 2.21uW power consumption at 5MHz reading frequency and 200kHz writing frequency.

Acknowledgements

My years in graduate school have been marked by inspiring teachers and encouraging friends. I thank everyone who has enriched my experience and contributed to my professional and personal growth. I sincerely thank my advisor, Prof. Wei Hwang, who has been supporting me in several ways. These include his wisdom and knowledge, both in depth and in breadth; the opportunities he provides for all his students; his enthusiasm for education and research development. Personally, he always encourages me to think and research independently on interesting topics, and tries his best to help.

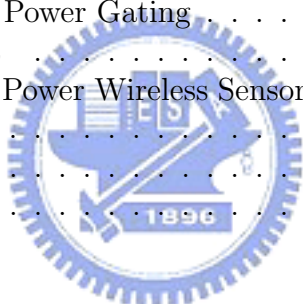
Next, I would like to thank Hong-Ren Liao, a manager of Taiwan Semiconductor Manufacturing Company (TSMC), who gave me the chance to work at TSMC. His solid suggestions helped me a lot in my graduate study. I would also like to thank my lab-mates in LPSOC Lab. These include Po-Tsang Huang, Hao-I Yang, Wei-Chih Hsieh, Ming-Hung Chang, Chi-Haoi Kan, Li-Pui Chuang, Wei-Li Fang, Ssu-Yun Lai, Tung-Hau Tsai, U-Chan Kuo. They provide me many inspiration and great companionship. In particular, Po-Tsang Huang has been fully supportive through the research progress. Moreover, I would like to thank Jui-Yuan Yu, Chien-Ying Yu, Shang-Bin Huang of SI2 group for fruitful discussions. Jui-Yuan Yu, in particular, has given me not only constructive suggestions, but also constant encouragement.

Finally, I would like to thank my loving family for their concern. With their support, I have the confidence to accomplish my goal.

Contents

1	Introduction	1
1.1	Background	1
1.2	Motivation	2
1.3	Thesis Organization	2
2	Overview of Ultra-Low Power Systems	3
2.1	Introduction	3
2.2	System Requirements	3
2.2.1	Battery Lifetimes	3
2.2.2	Energy Harvesting	3
2.3	Ultra-Low Power Systems	4
2.3.1	Radio Frequency Identification (RFID)	4
2.3.2	Digital Signal Processor (DSP)	5
2.3.3	Wireless Sensor Network	5
2.4	Summary	7
3	Overview of Low Power CMOS Circuit	8
3.1	Introduction	8
3.2	Power Dissipation	8
3.2.1	Dynamic Dissipation	8
3.2.2	Leakage Dissipation	9
3.2.3	Short Circuit Dissipation	12
3.2.4	Putting It All Together	13
3.3	Low Power Circuit Techniques	13
3.3.1	Supply Voltage Scaling	13
3.3.2	Transistor Stacking	16
3.3.3	Multiple Threshold Designs	22
3.4	Summary	28
4	Ultra-Low Voltage SRAM Design	29
4.1	Introduction	29
4.2	Overview of SRAM Operation	29
4.2.1	6T SRAM Cell	29
4.2.2	Dual-Port SRAM Cell	30
4.2.3	Column Circuitry	31
4.3	SRAM Cell Stability	32
4.3.1	Hold Stability	33
4.3.2	Read Stability	33

4.3.3	Write Ability	34
4.4	Single-Port Subthreshold SRAM Cell	35
4.4.1	Schmitt Trigger Based Subthreshold SRAM	36
4.4.2	Single-Port 10T Subthreshold SRAM	37
4.4.3	Proposed Subthreshold SRAM Cell with Auto-Compensation	37
4.4.4	Simulation Results	41
4.5	Dual-Port Subthreshold SRAM Cell	47
4.5.1	8T Subthreshold SRAM	47
4.5.2	Dual-Port 10T Subthreshold SRAM	48
4.5.3	Proposed Dual- V_T Subthreshold 7T SRAM cell	49
4.5.4	Simulation Results	49
4.6	Summary	54
5	A Robust Ultra-Low Power Asynchronous FIFO Memory	57
5.1	Introduction	57
5.2	Proposed Ultra-Low Power FIFO Memory	58
5.2.1	Logic Pointer	60
5.2.2	Read Operation	64
5.2.3	Write Operation	65
5.2.4	Self-Adaptive Power Control	67
5.2.5	Complementary Power Gating	68
5.2.6	Storage Element	69
5.3	Case Study: Ultra-Low Power Wireless Sensor Node for WBAN Application	70
5.4	Design Implementation	72
5.5	Simulation Results	73
5.6	Summary	75
6	Conclusions	77



List of Figures

1.1	Power density versus gate length.	1
2.1	Typical multi-hop wireless sensor network architecture.	6
3.1	A CMOS inverter.	9
3.2	Leakage current components in an NMOS transistor.	9
3.3	Components of tunneling current.	11
3.4	Gate leakage current versus gate oxide thickness.	11
3.5	Gate leakage current versus gate voltage.	11
3.6	A CMOS inverter chain.	13
3.7	Power versus supply voltage.	14
3.8	Time delay versus supply voltage.	14
3.9	PDP versus supply voltage.	15
3.10	Noise margin versus supply voltage.	15
3.11	Two-input NAND gate stacking effect illustration.	16
3.12	NMOS footer array power gating devices.	17
3.13	PMOS header array power gating devices.	17
3.14	Inverter chain with footer power gating.	18
3.15	Inverter chain with header power gating.	18
3.16	Standby power comparisons when applying footer power gating.	18
3.17	Standby power comparisons when applying header power gating.	19
3.18	Time delay comparisons when applying footer power gating.	19
3.19	Time delay comparisons when applying header power gating.	20
3.20	Active power comparisons when applying footer power gating.	20
3.21	Active power comparisons when applying header power gating.	21
3.22	Time delay comparisons between footer and header. One footer/header is applied on four inverters.	21
3.23	Standby power comparisons between footer and header. One footer/header is applied on four inverters.	22
3.24	Dual threshold CMOS circuit.	23
3.25	MVT CMOS scheme.	23
3.26	Footer insertion MTCMOS circuit.	24
3.27	Header insertion MTCMOS circuit.	24
3.28	MTCMOS inverter chain with footer power gating.	24
3.29	MTCMOS inverter chain with header power gating.	24
3.30	Standby power comparisons when applying footer insertion MTCMOS circuit.	25
3.31	Standby power comparisons when applying header insertion MTCMOS circuit.	25

3.32	Time delay comparisons when applying footer insertion MTCMOS circuit.	26
3.33	Time delay comparisons when applying header insertion MTCMOS circuit.	26
3.34	Active power comparisons when applying footer insertion MTCMOS circuit.	27
3.35	Active power comparisons when applying header insertion MTCMOS circuit.	27
4.1	SRAM organization.	30
4.2	Conventional 6T SRAM cell.	30
4.3	Read example of 6T SRAM cell.	31
4.4	Write example of 6T SRAM cell.	31
4.5	Conventional dual-port SRAM cell.	31
4.6	An SRAM column.	32
4.7	Standard setup for finding the Hold SNM.	33
4.8	Butterfly curve plots for representing SNM. The VTCs of the cross-coupled inverters are represented by the solid curves. The length of the side of the largest embedded square in the butterfly curve is the SNM. When the worst case static noise is applied (e.g., $V_N=SNM$), the bitcell is mono-stable, thus losing its data.	34
4.9	Standard setup for finding the Read SNM.	34
4.10	Example butterfly curve plots for hold SNM and read SNM.	35
4.11	Setup for finding WTP.	35
4.12	Write margin of a SRAM cell, determined by WTP.	36
4.13	Monte Carlo simulations indicating read/hold SNM failures and write margin failures of conventional 6T cell.	36
4.14	Schmitt trigger based subthreshold SRAM cell (ST cell).	37
4.15	Single-port 10T subthreshold SRAM cell (10T cell).	38
4.16	Proposed subthreshold SRAM cell with auto-compensation (AC cell).	39
4.17	Example of auto-compensation. The feedback system generated by AR1-RR holds node VL from being flipped.	39
4.18	Example of the read operation. AL1 and AR1 isolate storage nodes from bitlines. AR2-RR forms a read path from BR to GND.	40
4.19	Write operation of the AC cell.	40
4.20	Hold SNM comparisons under different supply voltage.	42
4.21	Distribution of hold SNM at 200mV.	42
4.22	Hold SNM comparisons at 200mV.	43
4.23	Read SNM comparisons under different supply voltage.	43
4.24	Distribution of read SNM at 200mV.	44
4.25	Read SNM comparisons at 200mV.	44
4.26	Write trip point comparisons under different supply voltage.	45
4.27	Distribution of write trip point at 200mV.	46
4.28	Write trip point comparisons at 200mV.	46
4.29	Leakage comparisons at 200mV. Leakage is normalized to the leakage of 6T cell.	47
4.30	8T Subthreshold SRAM cell.	47
4.31	Dual-port subthreshold 10T SRAM cell (10T_C).	48
4.32	Dual-port subthreshold 10T SRAM cell (10T_K).	48
4.33	Dual- V_T subthreshold 7T SRAM cell.	50
4.34	Hold SNM comparisons under different supply voltage.	51
4.35	Distribution of hold SNM at 300mV.	51

4.36	Hold SNM comparisons at 300mV.	52
4.37	Read SNM comparisons under different supply voltage.	52
4.38	Distribution of read SNM at 300mV.	53
4.39	Read SNM comparisons at 300mV.	53
4.40	Write margin comparisons under different supply voltage.	54
4.41	Distribution of write margin at 300mV.	55
4.42	Write margin comparisons at 200mV.	55
4.43	Cell leakage power comparison.	56
5.1	Block diagram of the wireless body network (WBAN) system wireless sensor node (WSN).	57
5.2	Block diagram of the proposed FIFO memory.	58
5.3	256x16 FIFO symbol.	59
5.4	Logic pointer composed by shift registers.	60
5.5	PowerPC master-slaver latch (PowerPC).	60
5.6	Modified C2MOS master-slaver latch (mC2MOS).	61
5.7	Hybrid-latch flip flop (HLFF).	61
5.8	Sense-amplifier-based flip-flop (SAFF).	61
5.9	Timing parameters of the flip-flop as a function of the supply voltage.	62
5.10	Energy dissipation as a function of the supply voltage for different switching activities.	63
5.11	EDP as a function of supply voltage and switching activities.	63
5.12	Read pointer.	64
5.13	Write pointer.	64
5.14	Read control circuit.	65
5.15	A FIFO memory Column.	66
5.16	Write control circuit.	66
5.17	FIFO operation example.	67
5.18	Adaptive power control circuit.	68
5.19	Waveform of the adaptive power control related signal (1st word).	68
5.20	Waveform of the adaptive power control related signal (256th word).	68
5.21	A FIFO memory word with complementary power gating.	69
5.22	Wireless body area network of intelligent sensors for patient monitoring.	71
5.23	Block diagram of the proposed WSN.	71
5.24	Layout view of the FIFO memory.	72
5.25	Layout view of the test chip.	73
5.26	Waveform of a complete data collection and data output.	74
5.27	Dual-port SRAM cells. (a) DP SRAM cell. (b) 8T SRAM cell. (c) 10T_C SRAM cell. (d) 10T_K SRAM cell.	75
5.28	Power consumption comparisons between conventional schemes and the proposed scheme. (a) Register based FIFO. (b) DP SRAM based FIFO. (c) 8T SRAM based FIFO; 10T_C SRAM based FIFO; 10T_K SRAM based FIFO.	76

List of Tables

2.1	Comparison of energy sources with fixed amount of energy storage. . . .	4
2.2	Comparison of ambient energy sources.	4
4.1	Summary of the AC cell operation.	41
5.1	Signal descriptions.	59
5.2	Command truth table.	59
5.3	Summary of FIFO word states and corresponding control signals.	69
5.4	Summary of the FIFO memory features.	72
5.5	Process corner simulation (@500mV; 25°C).	74
5.6	Voltage variation simulation (@TT corner; 25°C).	74
5.7	Temperature variation simulation (@TT corner; 500mV).	74



Chapter 1

Introduction

1.1 Background

Device miniaturization and the rapidly growing demand for mobile or power-aware systems have resulted in the urgent need for ultra-low power circuit design [1]. In modern CMOS technology, active power (dynamic power) and passive power (leakage power) are equally significant. This trend is shown in Figure 1.1 [2]. Therefore, to achieve ultra-low power operation, both active and passive power needs to be considered seriously.

In emerging system on chip (SoC) designs, an indispensable component is the on-chip memory module. As device density increases, a larger fraction of chip area is devoted to the memory block to enable more complex functionality and higher performance [3] [4]. As a result, power of memory blocks often dominates the total power consumption. Memory power consumption has thus been a major challenge and design consideration in future SoC.

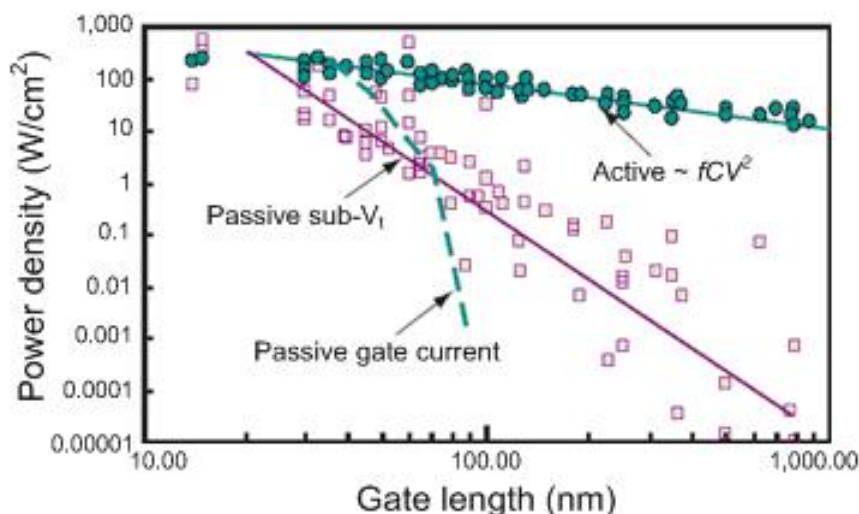


Figure 1.1: Power density versus gate length.

1.2 Motivation

In certain emerging applications, such as wireless sensor nodes, energy efficiency concerns supercede traditional emphasis on speed. These systems can be operated at much reduced performance levels in order to prolong their battery lifetimes. Many such low performance systems consume minimum energy in the subthreshold region, where the power supply voltage is below the device threshold voltage. This motivates the study of subthreshold circuits. In particular, the ever increasing demand for ultra-low power SRAM has motivates the design of subthreshold SRAM.

One example of the energy-constrained application is the emerging wireless body area network (WBAN), a generic concept of the short distance data transmission between personal wearing devices or implanted devices, which has been a breakthrough personal healthcare technology for body condition monitoring and diagnosis [5]. Due to limited power source, ultra-low power circuit is a key to WBAN to lengthen battery lifetime and enable energy harvesting. Meanwhile, first-in first-out (FIFO) memory used for data storage and data buffer is utilized in the sensor nodes. The FIFO memory often dominates the total die area and overall power consumption, thus motivates the design of robust, ultra-low power FIFO memory.

1.3 Thesis Organization

The rest of this thesis is organized as following. Chapter 2 presents the basic characteristics of ultra-low power systems, including system requirement and applications. Chapter 3 reviews CMOS circuit power sources and gives possible solutions to reduce power consumption. Ultra-low voltage SRAM designs are given in Chapter 4, including existing works, the proposed robust, fully-differential, single-port subthreshold SRAM with auto-compensation, and the proposed dual- V_T dual-port 7T SRAM cell with reduced bitline overhead and improved stability. In Chapter 5, a robust, ultra-low power FIFO memory design is proposed for WBAN application. Finally, Chapter 6 concludes this work.

Chapter 2

Overview of Ultra-Low Power Systems

2.1 Introduction

Emerging portable applications require low power operation due to limited energy source [6]. One class of portable applications operates in low activity rate or low speed, but is severely energy constrained. Therefore, for systems in this class, ultra-low power and energy conservation are the primary considerations.

This chapter first discusses the requirements for ultra-low power systems, which will be presented in Section 2.2. Examples of energy constrained applications, including *radio frequency identification (RFID)*, *digital signal processor (DSP)*, *wireless sensor network* are presented in Section 2.3. Summary of this chapter is presented in Section 2.4.

2.2 System Requirements

This section introduces requirements for ultra-low power and energy constrained systems, including battery lifetimes and energy harvesting mechanisms.

2.2.1 Battery Lifetimes

Modern portable micro-systems continue to integrate more functions into smaller devices. As a matter of fact, scaling down in size and cost of electronic circuits has far outpaced the scaling of energy density in batteries, which are by far the most common power sources currently used. Therefore, battery lifetimes is a key factor on the lifetime of portable devices.

Table 2.1 shows the comparison of power sources with a fixed amount of energy storage [7], indicating energy provided by batteries are limited, and ultra-low power operation is demanded for extended battery lifetimes.

2.2.2 Energy Harvesting

For some applications, changing batteries is impractical or impossible, and renewable energy source is required. Energy harvesting involves converting ambient energy from the environment into electrical energy to power circuits or to recharge batteries. Potential

ambient energy sources includes *photovoltaics (solar cells)*, *temperature gradients*, *human power*, *wind/air flow*, and *vibrations*. Comparison of ambient energy sources are summarized in Table 2.2 [8].

Due to unstable environment, the power available from these sources is impossible to maintain at a steady level. Thus, energy harvesting would be more effective to couple energy storage elements, which can theoretically extend system lifetimes indefinitely. Nevertheless, it is reasonable to keep the system's average power consumption in sub-mW range to enable energy harvesting.

Table 2.1: Comparison of energy sources with fixed amount of energy storage.

Power source	Power density	Power density
	($\mu W/cm^3$) one year lifetime	($\mu W/cm^3$) ten year lifetime
Batteries (non-rechargeable lithium)	45	3.5
Batteries (rechargeable lithium)	7	0
Hydrocarbon fuel (micro heat engine)	333	3
Fuel cells (methanol)	280	28

Table 2.2: Comparison of ambient energy sources.

Power source	Power density
Solar (outside)	15000 $\mu W/cm^2$
Solar (inside)	10 $\mu W/cm^2$
Temperature	40 $\mu W/cm^2$ (demonstrated from a $5^\circ C$ differential)
Human power	330 $\mu W/cm^3$
Air flow	380 $\mu W/cm^3$ (assumes air velocity of $5m/s$ and 5% conversion efficiency)
Vibrations	200 $\mu W/cm^3$

2.3 Ultra-Low Power Systems

In this section, three energy constrained ultra-low power applications are introduced, including RFID, DSP, and wireless sensor network. The emerging wireless body area network (WBAN), one specific application of wireless sensor network for personal health-care, will be presented in Chapter 5 as a case study, as well as the target application of the proposed robust ultra-low power asynchronous FIFO memory.

2.3.1 Radio Frequency Identification (RFID)

Radio Frequency identification (RFID) is an automatic identification method, relying on storing and remotely retrieving data through RFID tags [9]. An RFID tag is an object that can be applied to or incorporated into a product, animal, or person for the purpose of identification using radiowaves. Current RFID applications include passports, transportation payments, product tracking, lap scoring, animal identification, inventory systems, supply chain management, human implants, libraries, schools and universities,

museums, and social retailing. There are many other potential uses of RFID not listed above.

Most of RFID tags contain at least two parts. One is an integrated circuit for storing and processing information, modulating and demodulating a RF signal, and other specialized functions. The second is an antenna for receiving and transmitting the signal. RFID tags are generally classified into two types: active and passive. An active tag contains its own power source, usually an on-board battery. A passive tag obtains power from the signal of an external reader.

Power reduction is a key for better RFID performance. If the digital processing power is reduced, the distance from the reader to the tag will be increased since less transmitted power has to reach the tag. Further, minimizing power consumption leads to longer lifetimes of active type RFID tags.

2.3.2 Digital Signal Processor (DSP)

Digital signal processor (DSP) is a specialized microprocessor designed specifically for digital signal processing, which in general, computes in real time. Most portable consumer electronics, such as mobile phones or mp3 players, require a low power DSP. As technology advances, mobile devices tend to have more and more functions, such as a modern mobile phone is integrated with digital camera, audio player, video player, and video games. This leads to an ever increasing loading for the battery, where battery lifetime is a major concern for the convenience led by mobility.

Take mobile phone as an example. Most of the time, standby takes place, and ultra-low power is required for extended usage time. In other circumstances, such as the user makes a call, high performance is required for high quality function. In applications with characteristics described above, DSP needs a wide dynamic range of power and performance, and tradeoffs should be considered seriously [10]. An effective method is the *dynamic voltage and frequency scaling* technique [11], where optimal balance between power and performance can be achieved.

2.3.3 Wireless Sensor Network

A wireless sensor network is a wireless network consisting of spatially distributed autonomous devices using sensors to cooperatively monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants, at different locations [12].

As shown in Figure 2.1, a sensor network normally constitutes a wireless ad-hoc network, meaning that each sensor supports a multi-hop routing algorithm (several nodes may forward data packets to the base station). In addition to one or more sensors, each node in a sensor network is typically equipped with a radio transceiver or other wireless communications device, a small microcontroller, and an energy source, usually a battery. The base stations are one or more distinguished components of the wireless sensor network with much more computational, energy and communication resources. They act as a gateway between sensor nodes and the end user. Unique characteristics of a wireless sensor network include:

- Limited power
- Ability to withstand harsh environmental conditions

- Ability to cope with node failures
- Mobility of nodes
- Dynamic network topology
- Communication failures
- Heterogeneity of nodes
- Large scale of deployment
- Unattended operation

Possible applications include:

- Military applications: monitoring friendly forces, equipment and ammunition; battlefield surveillance; reconnaissance of opposing forces and terrain; targeting; battle damage assessment; nuclear, biological and chemical attack detection and reconnaissance.
- Environmental applications: forest fire detection; biocomplexity mapping of the environment; flood detection; precision agriculture.
- Health applications: telemonitoring of human physiological data; tracking and monitoring doctors and patients inside a hospital; drug administration in hospitals.
- Home applications: home automation; smart environment.
- Other commercial applications: environmental control in office buildings; interactive museums; detecting and monitoring car thefts; managing inventory control; vehicle tracking and detection.

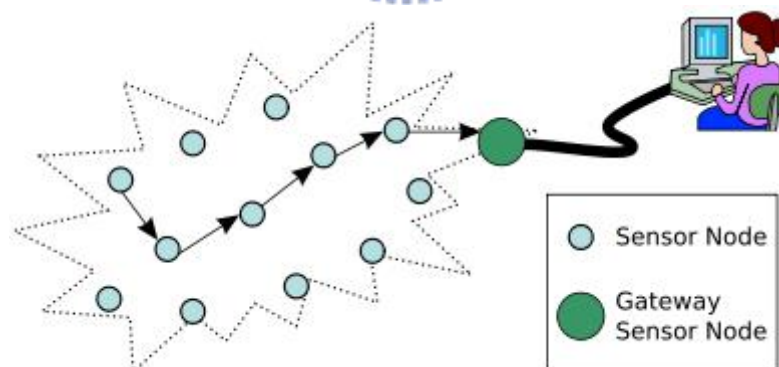


Figure 2.1: Typical multi-hop wireless sensor network architecture.

The wireless sensor node, being a micro-electronic device, can only be equipped with a limited power source. In some application scenarios, replacement of power resources might be impossible. Sensor node lifetime, therefore, shows a strong dependence on battery lifetime. In a multi-hop ad hoc sensor network, each node plays the dual role of data originator and data router. The disfunctioning of few nodes can cause significant topological changes and might require re-routing of packets and re-organization of the network. Hence, power conservation and power management take on additional importance, which directly influence the network efficiency and lifetime [13].

2.4 Summary

This chapter shows the motivation of ultra-low power operation by illustrating requirements for energy constrained applications. The effectiveness of ultra-low power operation not only extends battery lifetime, but also increase the possibility of enabling energy harvesting. In addition, three examples of ultra-low power system are given, including RFID, DSP, and wireless sensor network. Ultra-low power circuit design is no doubt a convincing solution for future portable SoC.



Chapter 3

Overview of Low Power CMOS Circuit

3.1 Introduction

This chapter begins with a study of power dissipation of CMOS circuit and circuit technique for power reduction. Power dissipation, including *dynamic dissipation*, *leakage dissipation*, and *short circuit dissipation*, is presented in Section 3.2. Low power circuit techniques, including *supply voltage scaling*, *transistor stacking*, *multiple threshold voltage design*, is presented in Section 3.3. Summary of this chapter is presented in Section 3.4.

3.2 Power Dissipation

3.2.1 Dynamic Dissipation

For a CMOS inverter, shown in Figure 3.1, the average dynamic power dissipation can be obtained by summing the average dynamic power in the NMOS transistor and the PMOS transistor. Assuming that the input V_{in} is a square wave having a period T and that the rise and fall times of the input are much less than the repetition period, the dynamic power is given by

$$P_D = \frac{1}{T} \int_0^{T/2} i_N(t)V_{out}dt + \frac{1}{T} \int_{T/2}^T i_P(t)(V_{DD} - V_{out})dt \quad (3.1)$$

Since $i_N(t) = C_L \frac{dV_{out}}{dt}$ and $i_P(t) = C_L \frac{d(V_{DD}-V_{out})}{dt}$,

$$P_D = \frac{C_L}{T} \int_0^{V_{DD}} V_{out}dV_{out} + \frac{C_L}{T} \int_{V_{DD}}^0 (V_{DD} - V_{out})d(V_{DD} - V_{out}) = \frac{C_L V_{DD}^2}{T} \quad (3.2)$$

Where C_L is the load capacitance, $\frac{1}{T} = f$, f is the operating frequency. Therefore

$$P_D = fC_L V_{DD}^2 \quad (3.3)$$

Moreover, power dissipation is data dependent, i.e. power dissipation depends on the switching probability α , thus, dynamic power can be expressed as

$$P_D = \alpha f C_L V_{DD}^2 \quad (3.4)$$

By (3.4), dynamic power dissipation of CMOS logic gate is proportional to switching frequency, load capacitance, square of the supply voltage, and operation frequency.

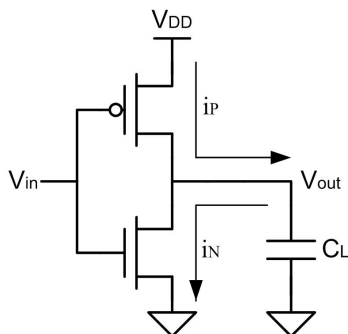


Figure 3.1: A CMOS inverter.

3.2.2 Leakage Dissipation

There are four main sources of leakage current in a CMOS transistor as illustrated in Figure 3.2 [14]–[16]. They are *reverse-biased junction leakage current* (I_{REV}), *gate induced drain leakage* (I_{GIDL}), *gate direct-tunneling leakage* (I_G), and *subthreshold leakage* (I_{SUB}). Each source of leakage current will be further described in the followings.

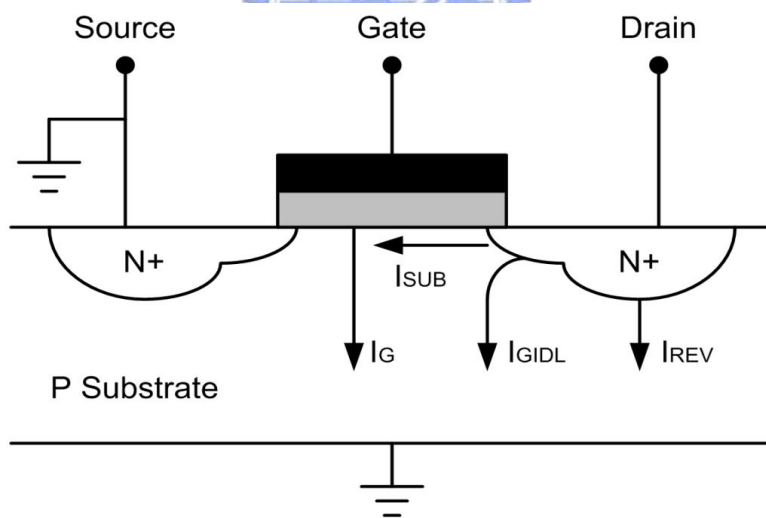


Figure 3.2: Leakage current components in an NMOS transistor.

Junction Leakage

The junction leakage occurs from the source/drain to the substrate through the reverse-biased diodes when the transistor is off, indicated as I_{REV} in Figure 3.2. A reverse-biased pn junction leakage has two major components: one is minority carrier diffusion/drift near the edge of the depletion region; the other is due to electron-hole pair generation in the depletion region of the reverse-biased junction.

Junction leakage current depends on the area of the drain diffusion and the leakage current density, which is in turn determined by the doping concentration. Junction leakage components from both the source-drain diodes and the well diodes are generally negligible with respect to the other three leakage components.

Gate-Induced Drain Leakage

Gate-induced drain leakage (GIDL), indicated as I_{GIDL} in Figure 3.2, arises in the high electric field under the gate/drain overlap region. GIDL occurs at large V_{DB} and generates carriers into the substrate and drain from surface traps or band-to-band tunneling. Thinner oxide, higher supply voltage, and lightly doped drain structures increase GIDL current.

Gate Direct Tunneling Leakage

Gate direct tunneling current is due to the tunneling of an electron/hole from the bulk silicon through the gate oxide potential barrier into the gate [17][18]. Reduction of gate oxide thickness results in the increase in the field across the oxide. The high electric field coupled with low oxide thickness results in tunneling of electrons from substrate to gate and also from gate to substrate through the gate oxide, resulting in the gate leakage. In nanometer-scale CMOS technologies, where ultra-thin gate oxide thickness takes place for effective gate control, gate leakage becomes appreciable and dominates the total leakage dissipation [19].

Figure 3.3 shows the components of tunneling current in a scaled NMOS transistor. They are classified in to three categories:

1. Edge direct tunneling (EDT) components between the gate and the source-drain extension (SDE) overlap region (I_{gso} and I_{gdo}).
2. Gate-to-channel current (I_{gc}), part of which goes to the source (I_{gcs}), and the rest goes to the drain (I_{gcd}).
3. Gate-to-substrate leakage current (I_{gb}).

Therefore, the gate leakage (I_G) can be divided into three major components:

1. Gate-to-source ($I_{gs} = I_{gso} + I_{gcs}$).
2. Gate-to-drain ($I_{gd} = I_{gdo} + I_{gcd}$).
3. Gate-to-substrate (I_{gb}).

The magnitude of the gate leakage current increases exponentially with the gate oxide thickness T_{OX} and the gate-to-source voltage V_{GS} , as shown in Figure 3.4 and Figure 3.5, respectively [20].

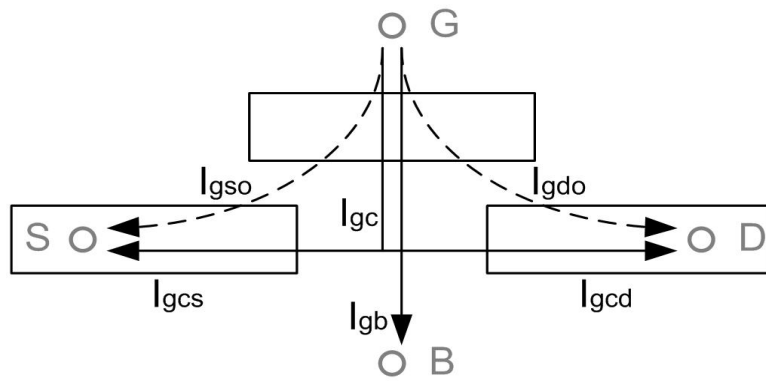


Figure 3.3: Components of tunneling current.

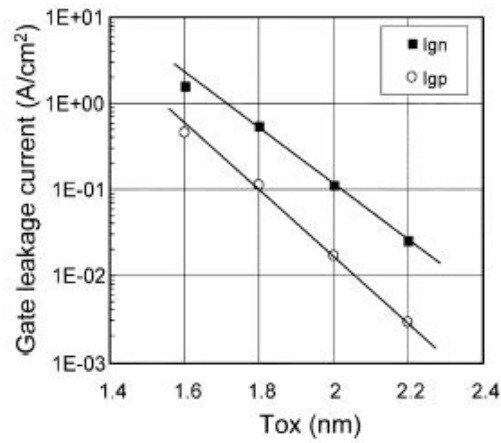


Figure 3.4: Gate leakage current versus gate oxide thickness.

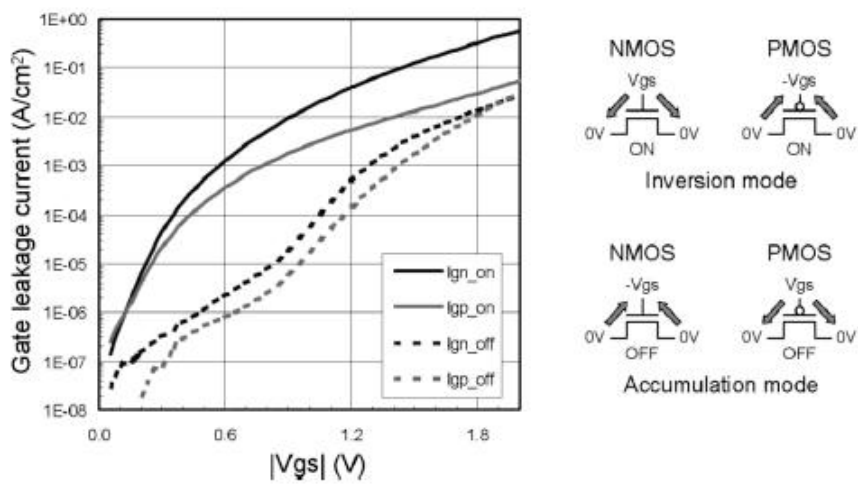


Figure 3.5: Gate leakage current versus gate voltage.

Subthreshold Leakage

Subthreshold or weak inversion conduction current between source and drain of an MOS transistor occurs when gate voltage is below the threshold voltage level. Unlike the strong inversion region in which the drift current dominates, the subthreshold conduction is due to the diffusion current of the minority carriers in the channel for a MOS device. For instance, in an inverter with a low input voltage and high output voltage, for the NMOS transistor, even V_{GS} is 0V, there is still a current flowing in the channel of the off NMOS transistor due to the V_{DD} potential of the V_{DS} .

Subthreshold leakage current (I_{SUB}) becomes apparent as CMOS technologies enter the submicron era [21]. I_{SUB} can be expressed based on the following:

$$I_{SUB} = \frac{W}{L} \mu \nu_{th}^2 C_{sth} e^{\frac{V_{GS} - V_T + \eta V_{DS}}{n \nu_{th}}} \left(1 - e^{\frac{-V_{DS}}{\nu_{th}}}\right) \quad (3.5)$$

where W and L denote the transistor width and length, μ denotes the carrier mobility, $\nu_{th} = kT/q$ denotes the thermal voltage at temperature T , $C_{sth} = C_{dep} + C_{it}$ denotes the summation of the depletion region capacitance and the interface trap capacitance both per unit area of the MOS gate, and η is the drain-induced barrier lowering (DIBL) coefficient. n is the slope shape factor and is calculated as:

$$n = 1 + \frac{C_{sth}}{C_{ox}} \quad (3.6)$$

where C_{ox} denotes the gate input capacitance per unit area of the MOS gate. Thus, the magnitude of the subthreshold leakage current is a function of the temperature, supply voltage, device size, and the process parameters out of which the threshold voltage plays a dominant role.

3.2.3 Short Circuit Dissipation

The short circuit power dissipation results due to a direct path current flowing from the power supply to the ground during the switching of a static CMOS gate. Short circuit dissipation can be expressed as:

$$P_{SC} = I_{mean} V_{DD} \quad (3.7)$$

where I_{mean} is the mean value of the short circuit current. Assuming a symmetrical inverter and using simple MOS formula, I_{mean} is modeled as [22]:

$$I_{mean} = \frac{1}{12} \frac{\beta}{V_{DD}} (V_{DD} - 2V_T)^3 \frac{\tau}{T} \quad (3.8)$$

where β is the gain factor of a MOS transistor, τ is the input rise/fall time.

From (3.7) and (3.8), short circuit dissipation of a CMOS inverter without load is derived as:

$$P_{SC} = \frac{\beta}{12} (V_{DD} - 2V_T)^3 \frac{\tau}{T} \quad (3.9)$$

Although this is a simplified model, it reveals the fact that short circuit dissipation is affected by supply voltage, threshold voltage, rise/fall time, and operation frequency. Therefore, it is effective to minimize short-circuit power by lowering supply voltage, increasing threshold voltage, and minimizing input rise/fall time.

3.2.4 Putting It All Together

The total power consumption of a digital CMOS circuit can be expressed as the sum of its three components:

$$P_{Total} = P_D + P_{Leak} + P_{SC} = \alpha f C_L V_{DD}^2 + I_{Leak} V_{DD} + I_{SC} V_{DD} \quad (3.10)$$

Clearly, supply voltage has a major dominance over power consumption. In the next section, several circuit techniques for power control and reduction are presented, including *supply voltage scaling*, *transistor stacking*, and *multiple threshold voltage design*. Both Active and standby power reduction are considered.

3.3 Low Power Circuit Techniques

3.3.1 Supply Voltage Scaling

In a given technology, supply voltage reduction is the key to low power operation [23][24]. When lowering the supply voltage, there are two issues that must be considered:

1. Impact on delay: Since both capacitance and threshold voltage are constant, the speed of the basic gates will also decrease with the voltage scaling, where the relation between time delay T_d and supply voltage V_{DD} can be modeled by using a quadratic model:

$$T_d = k \frac{C_L V_{DD}}{(V_{DD} - V_T)^2} \quad (3.11)$$

2. Impact on stability: Low supply voltage circuits are very sensitive to both manufacturing variations and operating point changes, which leads to less stable and less robust operation.

Following is an example of supply voltage scaling. Figure 3.6 shows an inverter chain composed of four inverters. Figure 3.7 shows the relation between power and supply voltage; Figure 3.8 shows the relation between time delay and supply voltage. It is revealed that as supply voltage drops, power consumption is reduced, but the time delay is increased. A common vector for finding the optimal supply voltage is the *power delay product (PDP)*, which is the product of power and time delay, as shown in Figure 3.9. Another strategy is to find the worst case critical time delay and choose the minimum supply voltage that is capable of performing the expected operation speed.

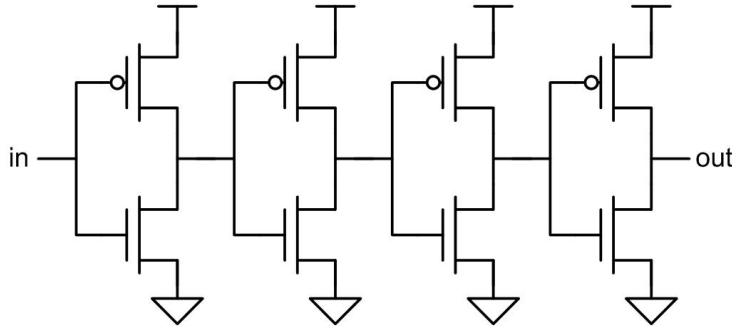


Figure 3.6: A CMOS inverter chain.

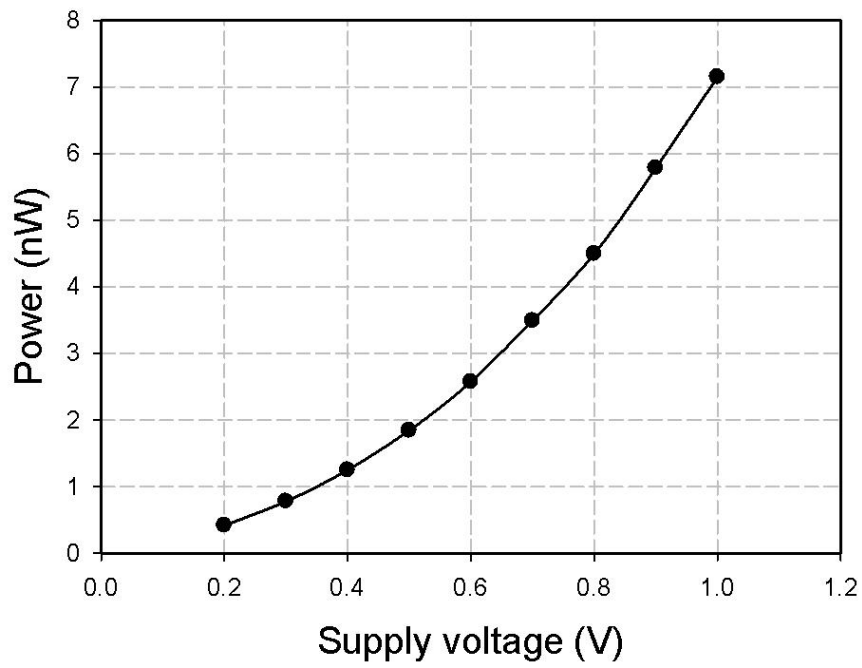


Figure 3.7: Power versus supply voltage.

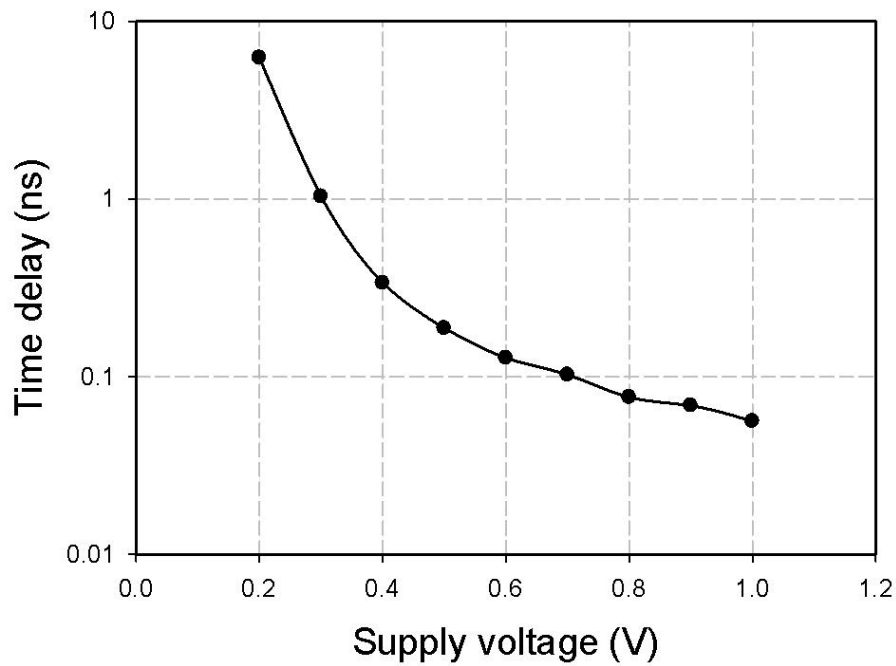
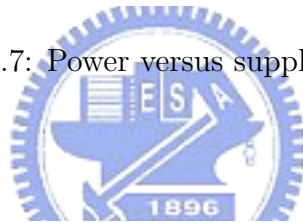


Figure 3.8: Time delay versus supply voltage.

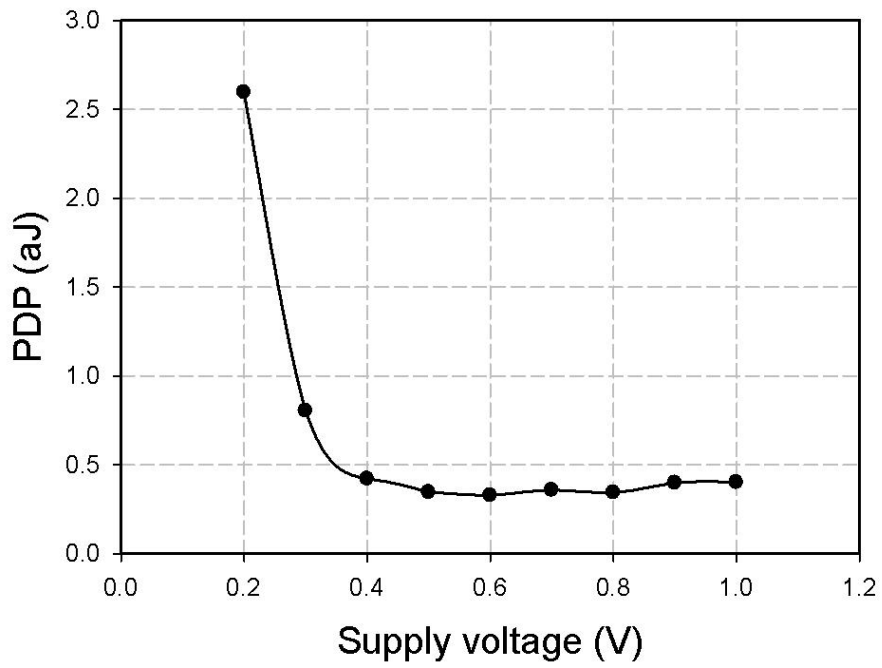


Figure 3.9: PDP versus supply voltage.

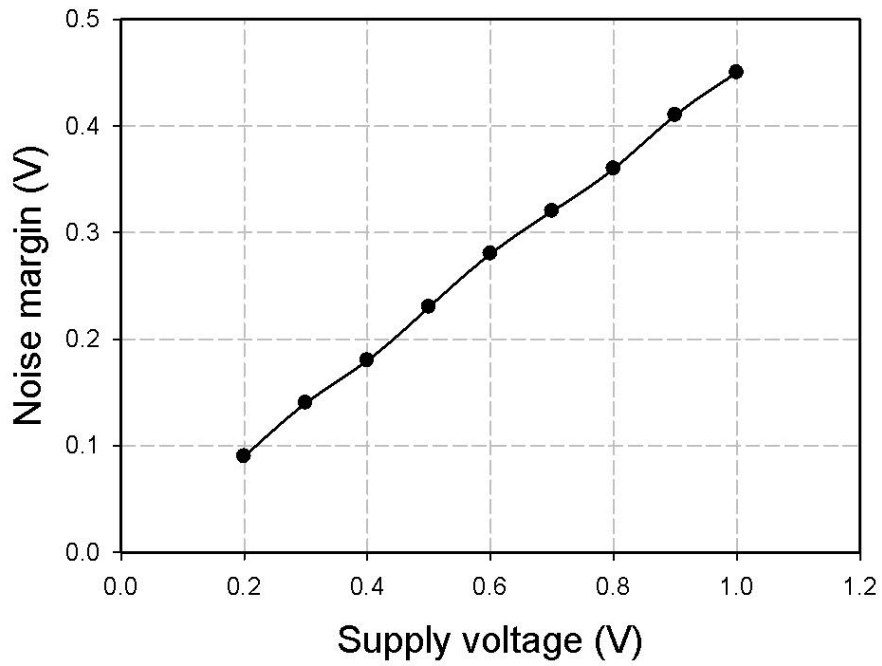


Figure 3.10: Noise margin versus supply voltage.

Relation between noise margin [25] and supply voltage is shown in Figure 3.10. As shown, noise margin decreases as supply voltage drops. Noise margin issue is especially important in ultra-low voltage and subthreshold circuit designs [26].

3.3.2 Transistor Stacking

Transistor stacking is an effective technique to reduce subthreshold and gate leakage current [27][28]. Leakage current flowing through a stack of series-connected transistors reduces if more than one transistor in the stack is off, which is known as the stacking effect. The stacking effect can be understood by considering a two-input NAND gate, as shown in Figure 3.11. When both MN1 and MN2 are off, the voltage at the intermediate node (V_M) raises to a positive value due to a small drain current. Positive potential at the intermediate node leads to three effects:

1. Gate-to-source voltage of MN1 becomes negative.
2. Negative body-to-source potential of MN1 causes more body effect. The body effect describes how the potential difference between source and body affects the threshold voltage, which can be modeled as:

$$V_T = V_{T0} + \gamma(\sqrt{\phi_s + V_{SB}} - \sqrt{\phi_s}) \quad (3.12)$$

where V_{T0} is the threshold voltage when the source is at the body potential; ϕ_s is the surface potential at threshold, and γ is the body effect coefficient.

3. Drain-to-source potential of MN1 decreases, resulting in less drain-induced barrier lowering.

As a result, negative gate-to-source voltage, higher threshold voltage due to the body effect, and less drain-induced barrier lowering due to the reduction of drain-to-source voltage, leakage current is reduced.

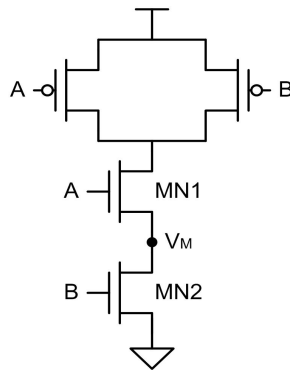


Figure 3.11: Two-input NAND gate stacking effect illustration.

Transistor stacking for low power can be referred to power gating. Power gating devices can be classified into two main categories: footer and header devices. Footer is by inserting NMOS sleep transistors between real GND and virtual GND, while header is by inserting PMOS sleep transistors between real V_{DD} and virtual V_{DD} , as shown in Figure 3.12 and Figure 3.13, respectively.

Figure 3.14 and Figure 3.15 are testing examples of footer and header. The effectiveness of standby power saving by footer and header are shown in Figure 3.16 and Figure 3.17. Time delay comparisons are shown in Figure 3.18 and Figure 3.19. As shown, by sacrificing operation speed, a circuit with power gating devices has significant standby power (leakage power) reduction. Trade off between power and speed is also illustrated. For a circuit with power gating, the less power gating are inserted, the more power is saved, and the more power gating are inserted, the less time delay it performs. Adding power gating devices usually contributes very slight active power overhead, which is revealed in Figure 3.20 and Figure 3.21. Another interesting thing worth notice is the comparison between footer and header, which is demonstrated in Figure 3.22 and Figure 3.23. NMOS has stronger driving ability than PMOS, resulting in smaller time delay when applying footer power gating. On the other hand, as shown in Figure 3.5, PMOS has smaller leakage current than NMOS, resulting in smaller power consumption when applying header power gating.

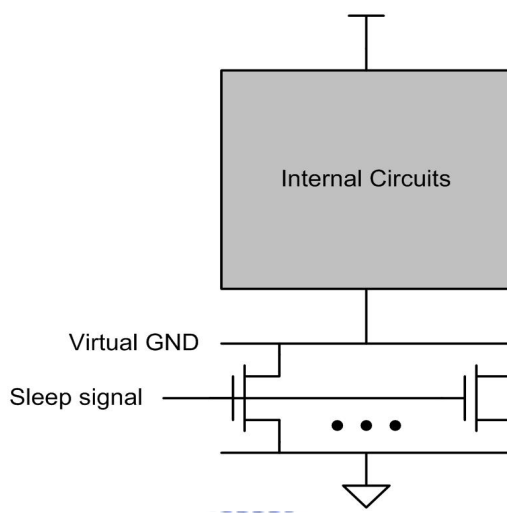


Figure 3.12: NMOS footer array power gating devices.

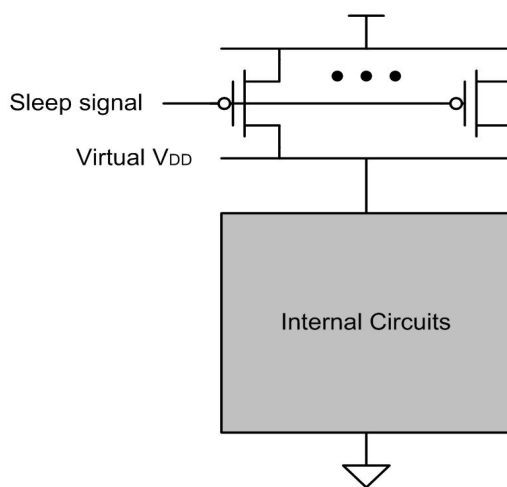


Figure 3.13: PMOS header array power gating devices.

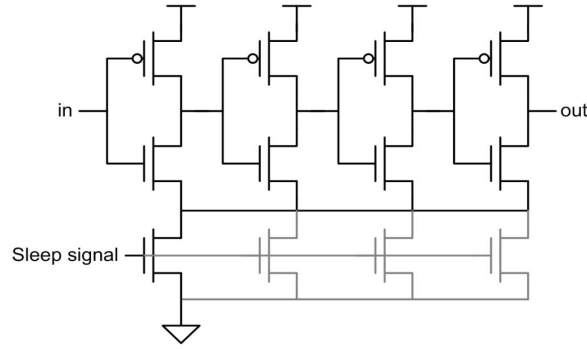


Figure 3.14: Inverter chain with footer power gating.

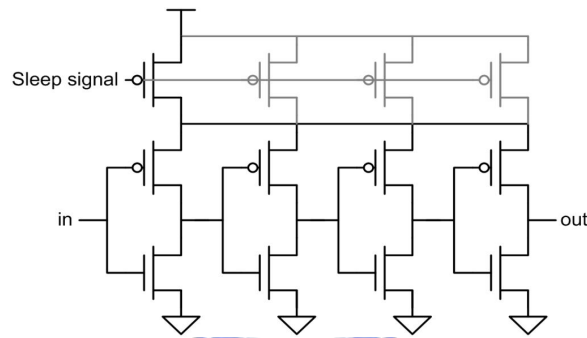


Figure 3.15: Inverter chain with header power gating.

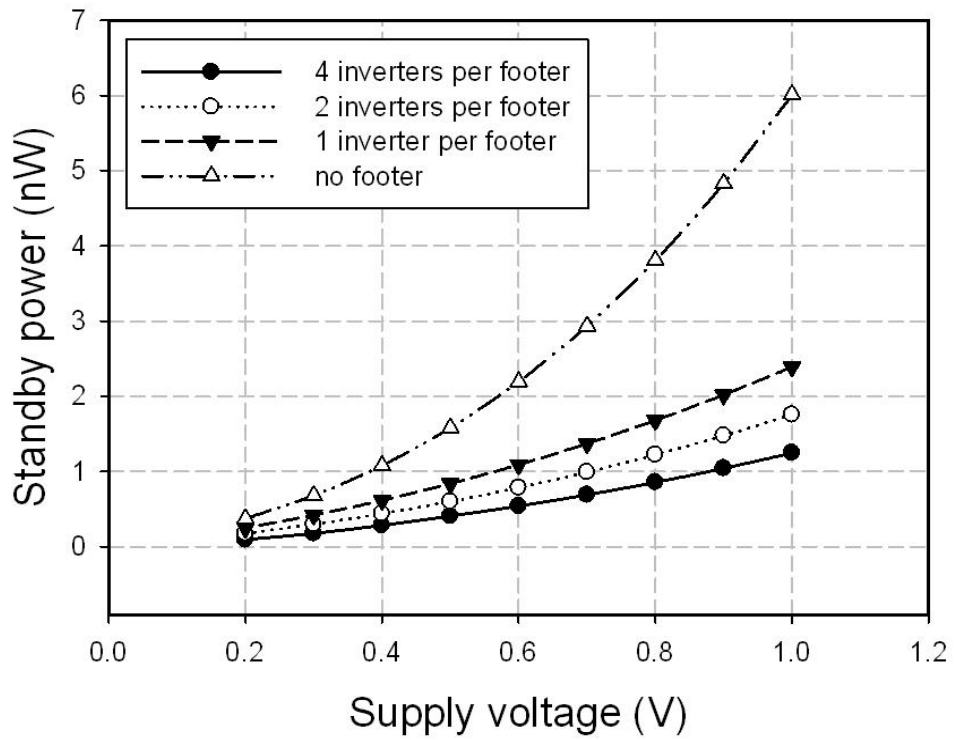


Figure 3.16: Standby power comparisons when applying footer power gating.

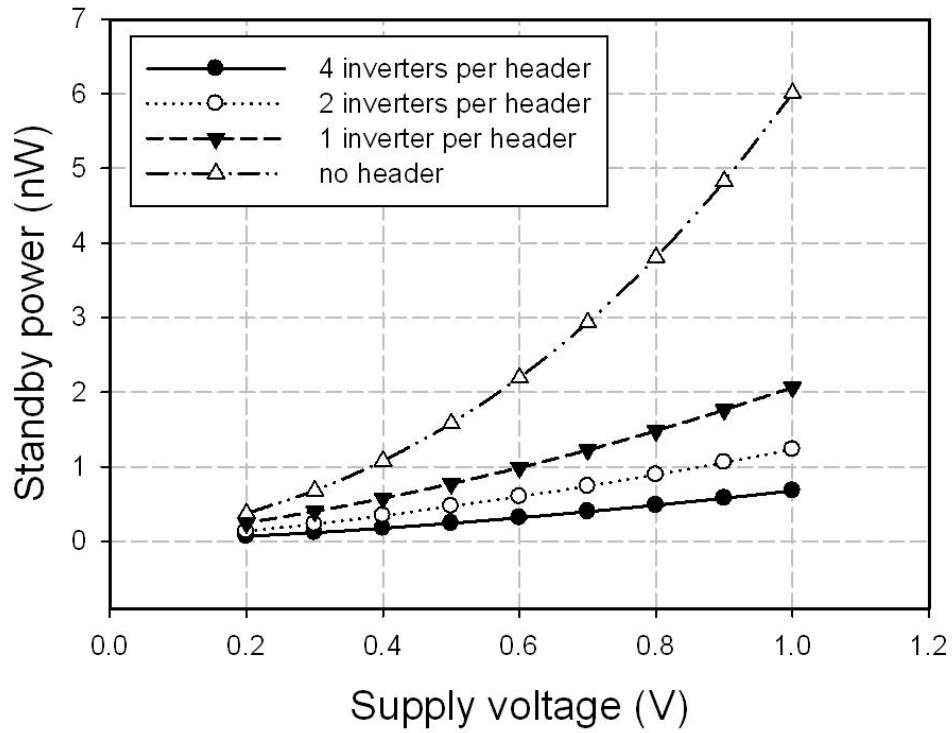


Figure 3.17: Standby power comparisons when applying header power gating.

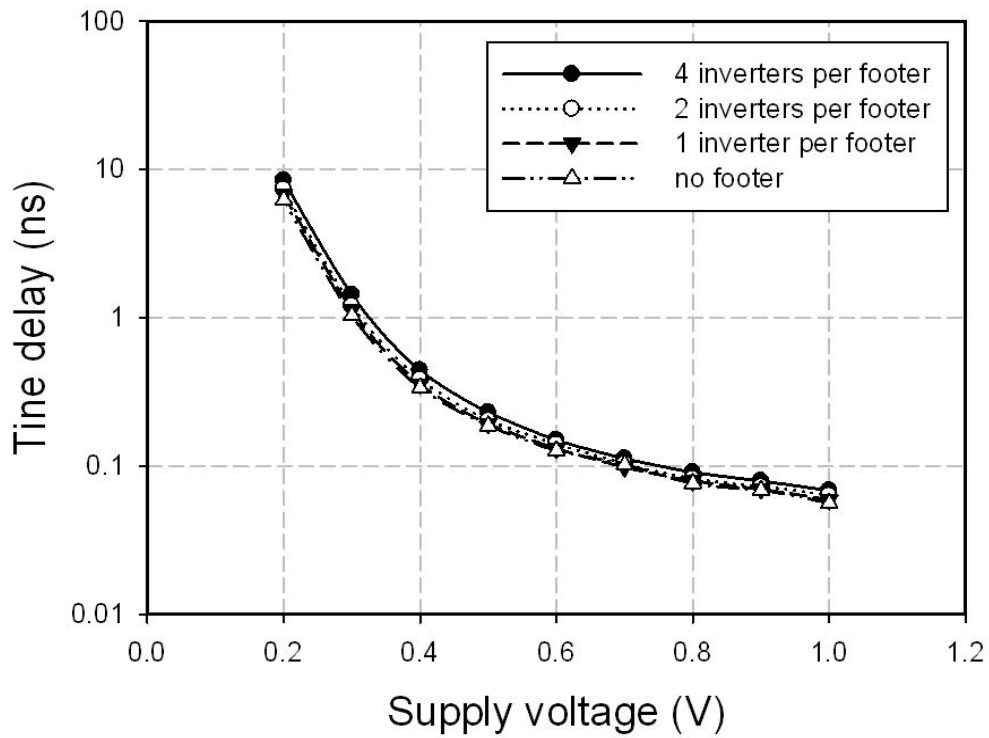


Figure 3.18: Time delay comparisons when applying footer power gating.

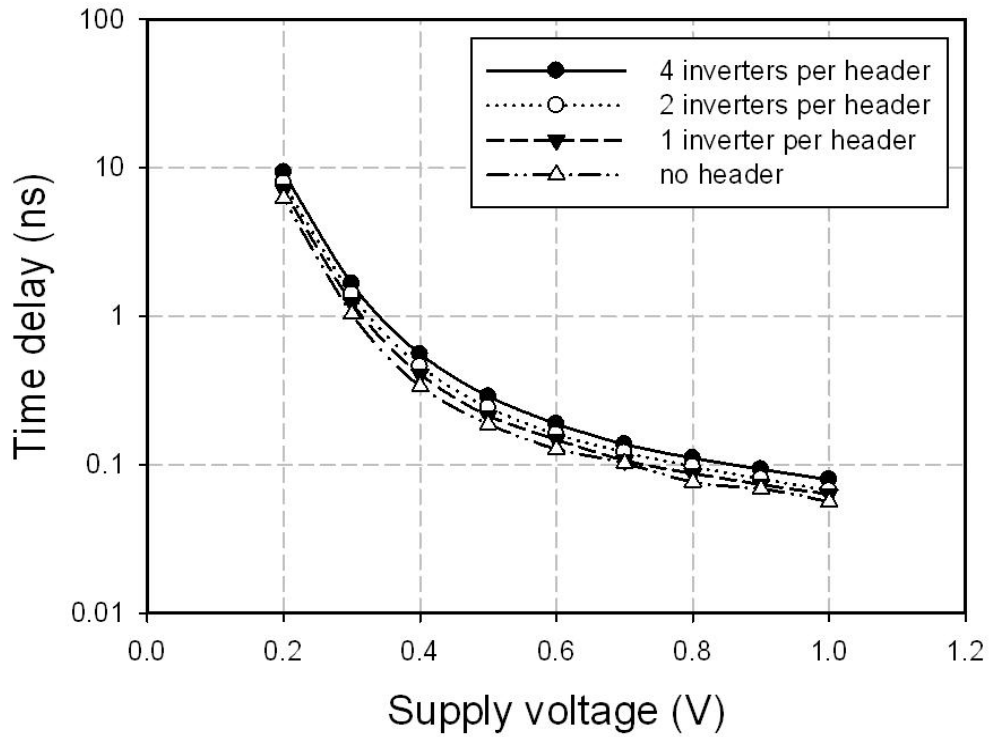


Figure 3.19: Time delay comparisons when applying header power gating.

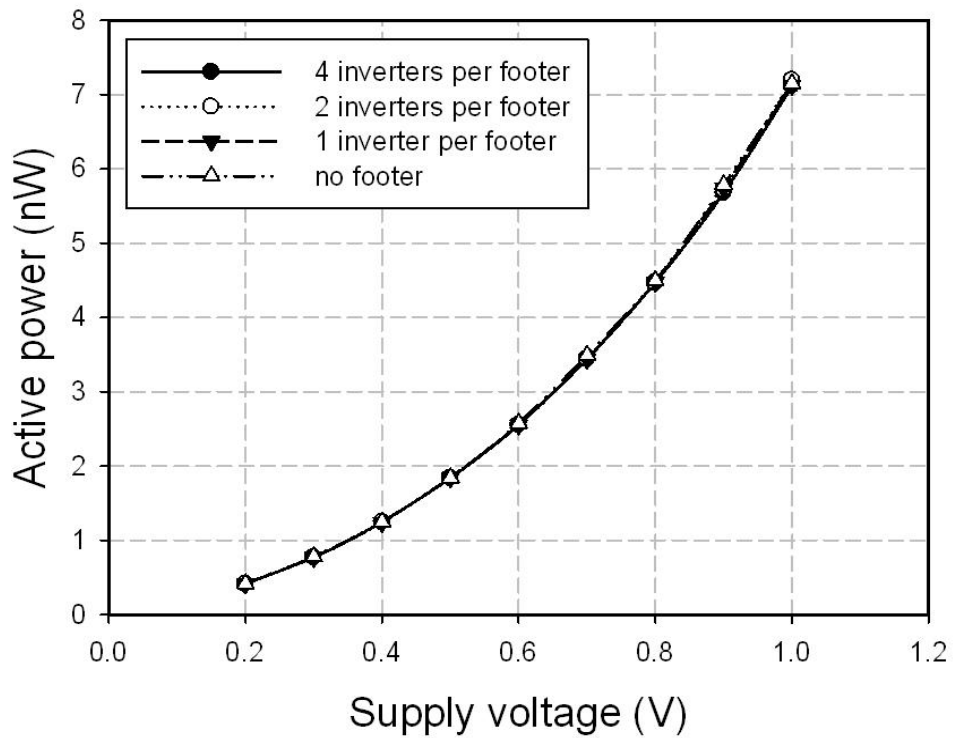


Figure 3.20: Active power comparisons when applying footer power gating.

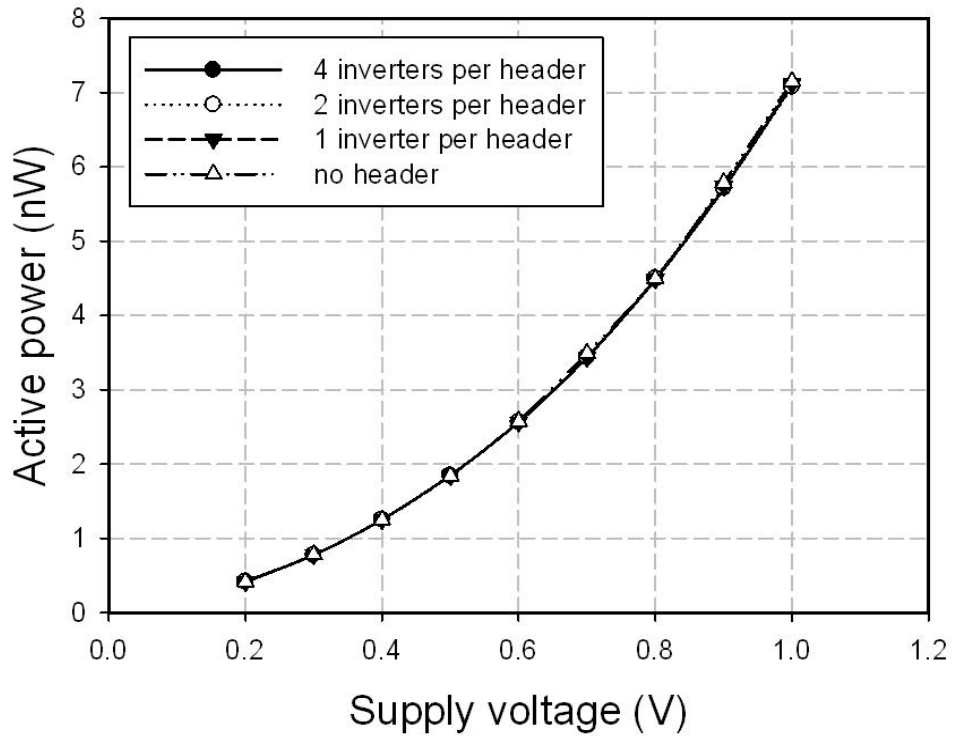


Figure 3.21: Active power comparisons when applying header power gating.

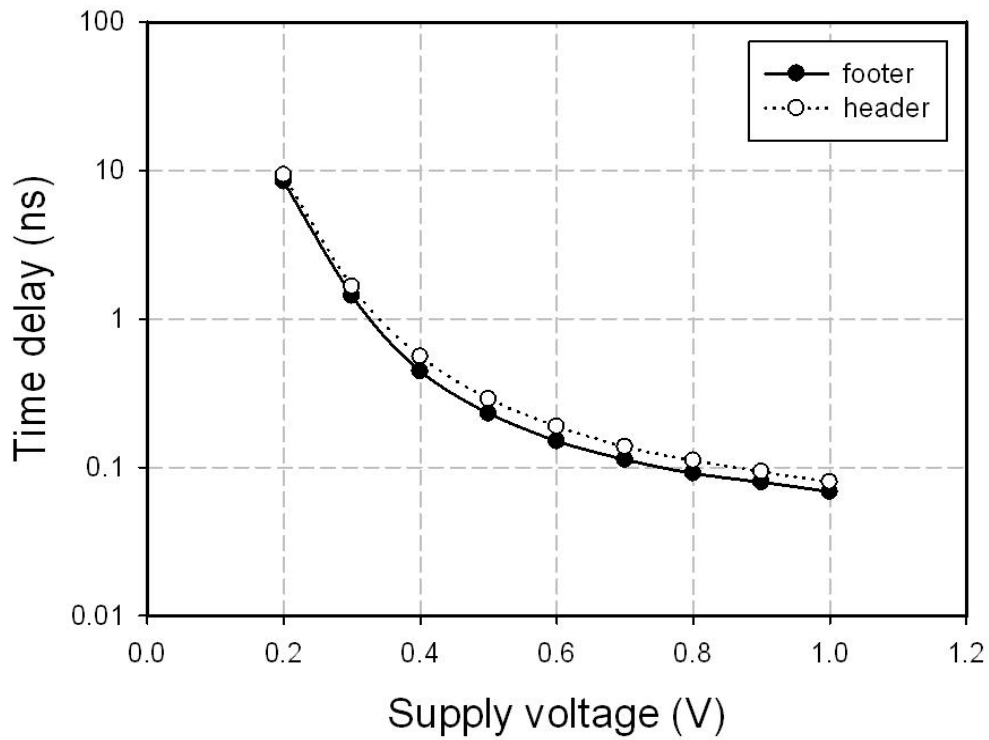


Figure 3.22: Time delay comparisons between footer and header. One footer/header is applied on four inverters.

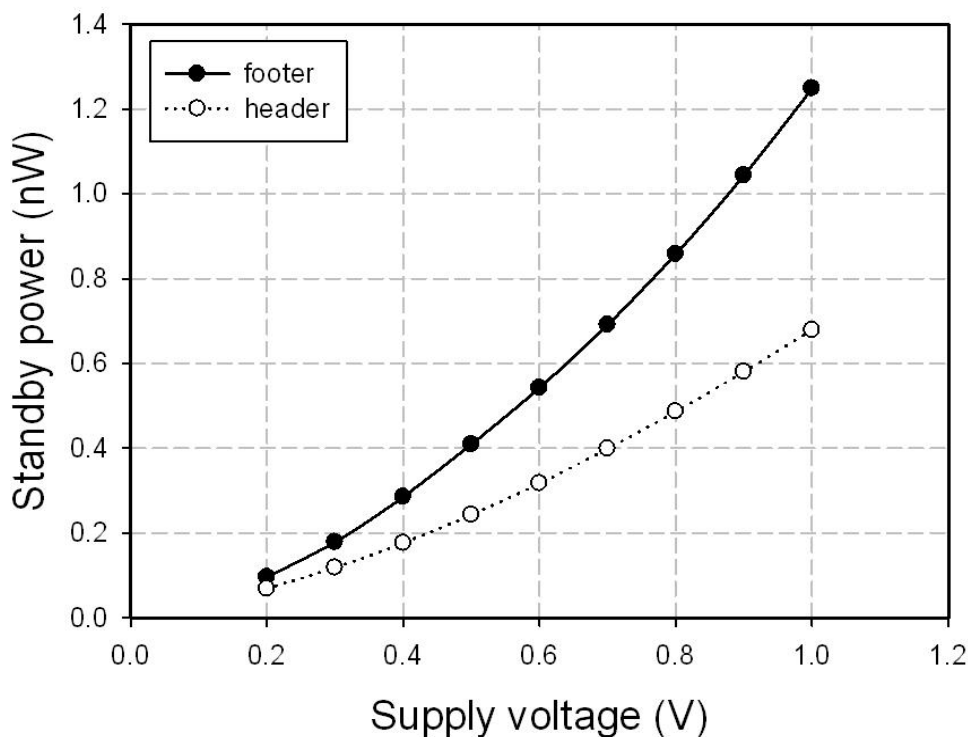


Figure 3.23: Standby power comparisons between footer and header. One footer/header is applied on four inverters.

3.3.3 Multiple Threshold Designs

Multiple threshold CMOS (MTCMOS) circuit has transistors with different threshold voltage. In general, there are regular threshold (regular- V_T) transistors, low threshold (low- V_T) transistors, and high threshold (high- V_T) transistors. Low- V_T transistors has larger driving ability, and can be used to achieve high performance, but it has the largest leakage current among the three types of transistors. High- V_T transistors has the least leakage current, but its performance is the slowest among the three types of transistors. The performance of regular- V_T transistors is in between low- V_T and high- V_T transistors. Following are three multiple threshold technologies:

1. Dual threshold CMOS: In a logic circuit, if a logic gate is in the critical path, the gate is implemented by low- V_T transistors to maintain performance; if a logic gate is in a non-critical path, the gate is implemented by high- V_T transistors for leakage power reduction [29]. This technique is demonstrated in Figure 3.24.
2. Mixed- V_T (MVT) CMOS technique: Unlike dual threshold CMOS technique, MVT CMOS design technique allows different thresholds within a logic gate, placing high- V_T transistors in non-critical paths to reduce leakage power, and placing low- V_T transistors in critical path(s) to maintain performance [30][31]. Figure 3.25 is an example of MVT CMOS logic gate. Suppose that the transistors in squares are the transistors in the critical paths, thus, assigning low- V_T . For the other transistors, high- V_T are assigned for leakage power reduction without degrading performance. Both dual threshold CMOS and MVT CMOS technique can achieve power reduction

without delay and area overhead.

3. Multithreshold-voltage CMOS: Multithreshold-voltage CMOS (MTCMOS) technique is based on transistor stacking technique, but utilizes low- V_T transistors for logic gates and apply high- V_T transistors to power gating [32]. Examples are shown in Figure 3.26 and Figure 3.27. Assigning high- V_T to power gating devices can further improve leakage cut off efficiency, while the delay overhead can be compensated by low- V_T logic gates. Figure 3.28 and Figure 3.29 are testing examples of MTCMOS circuit. Figure 3.30 and Figure 3.31 show the standby power comparison between inverter chain with and without MTCMOS technique. It is obvious MTCMOS technique significantly reduces standby power. Figure 3.32 and Figure 3.33 show the time delay comparison between inverter chain with and without MTCMOS technique. High- V_T has smaller driving current, thus resulting delay overhead. Delay overhead can be reduced by replacing regular- V_T transistors with low- V_T transistors. Figure 3.34 and Figure 3.35 show the active power comparison between inverter chain with and without MTCMOS technique. Active power reduction by MTCMOS is not apparent in this case, since the gate count under simulation is very limited.

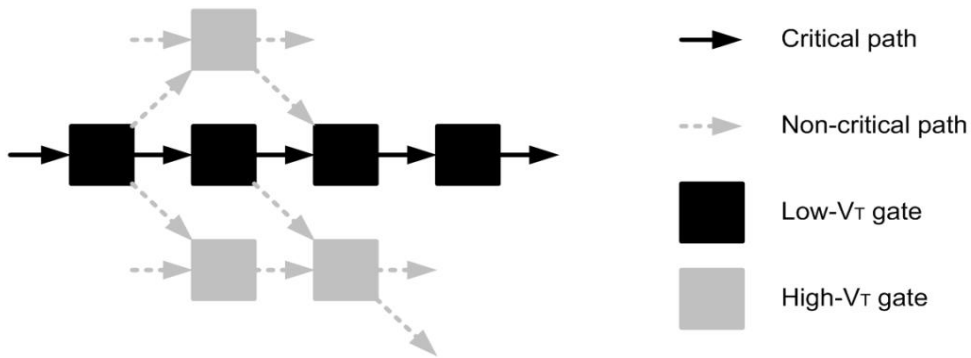


Figure 3.24: Dual threshold CMOS circuit.

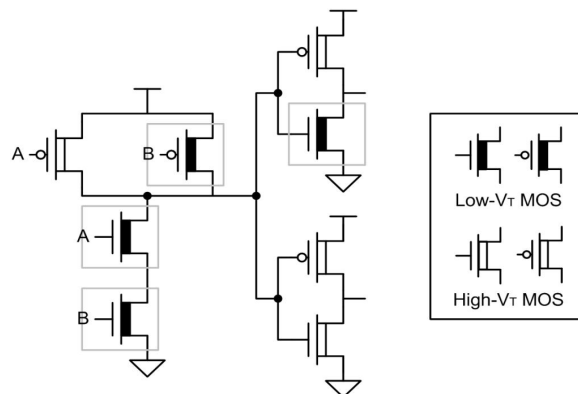


Figure 3.25: MVT CMOS scheme.

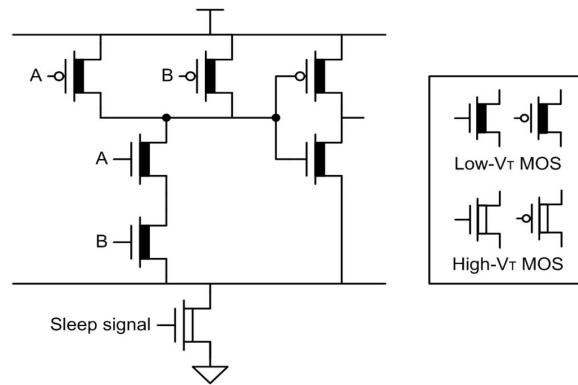


Figure 3.26: Footer insertion MTCMOS circuit.

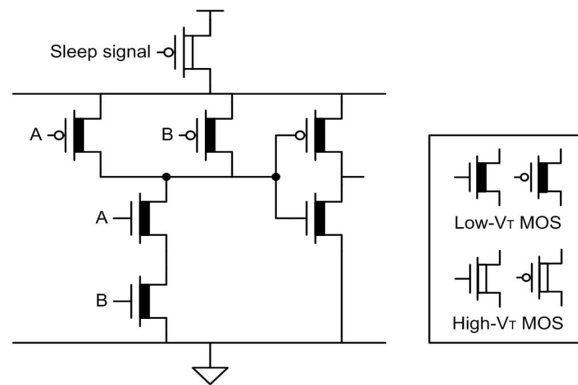


Figure 3.27: Header insertion MTCMOS circuit.

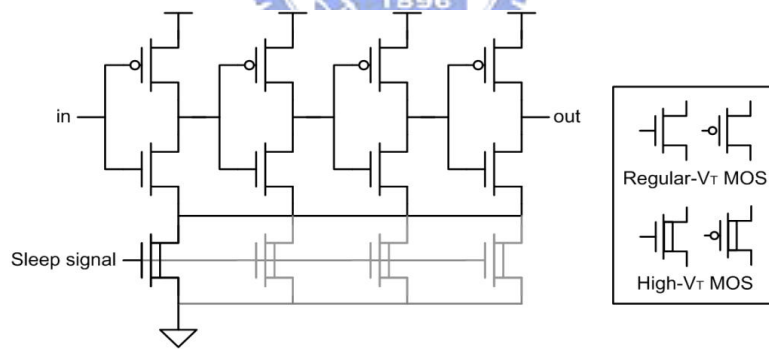


Figure 3.28: MTCMOS inverter chain with footer power gating.

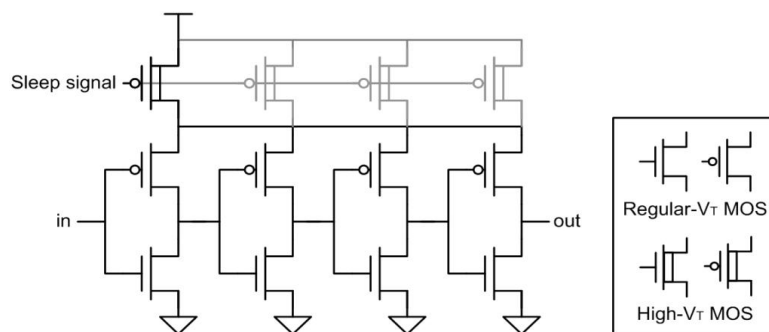


Figure 3.29: MTCMOS inverter chain with header power gating.

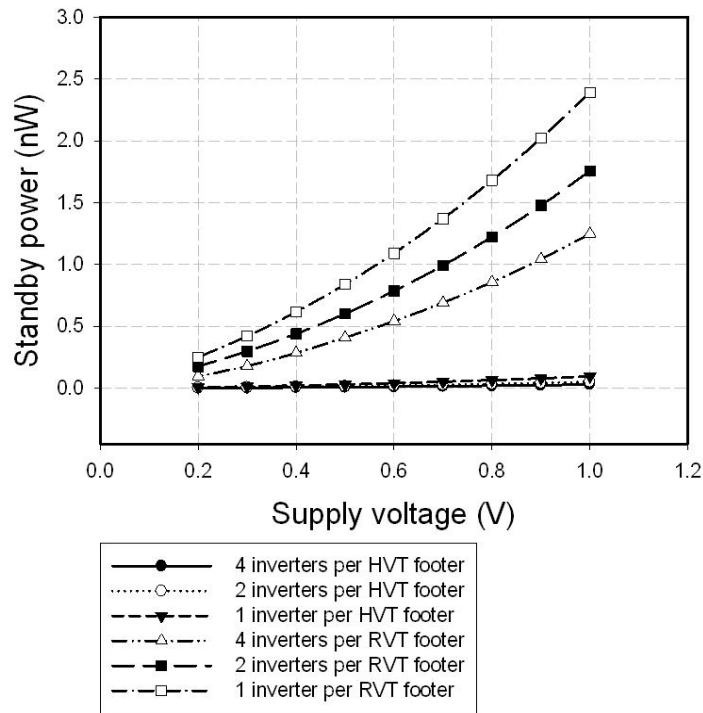


Figure 3.30: Standby power comparisons when applying footer insertion MTCMOS circuit.

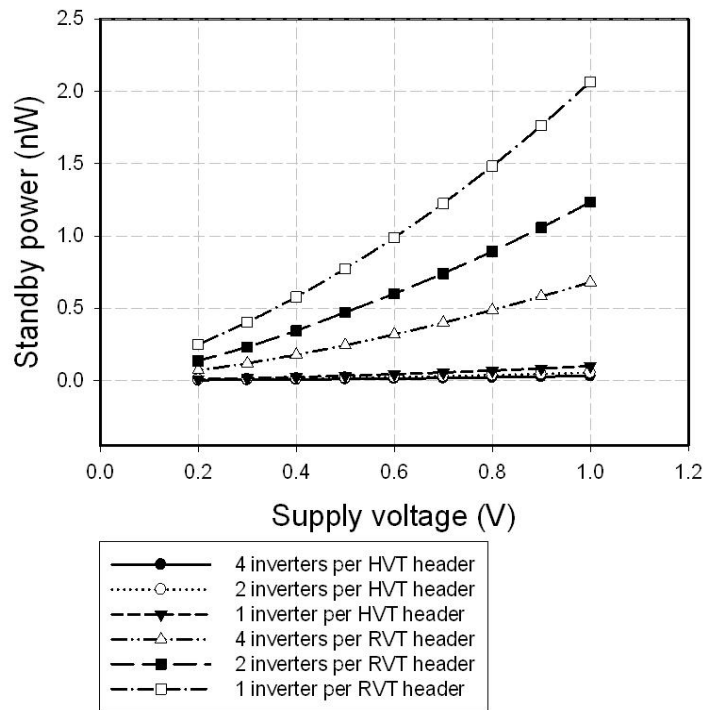


Figure 3.31: Standby power comparisons when applying header insertion MTCMOS circuit.

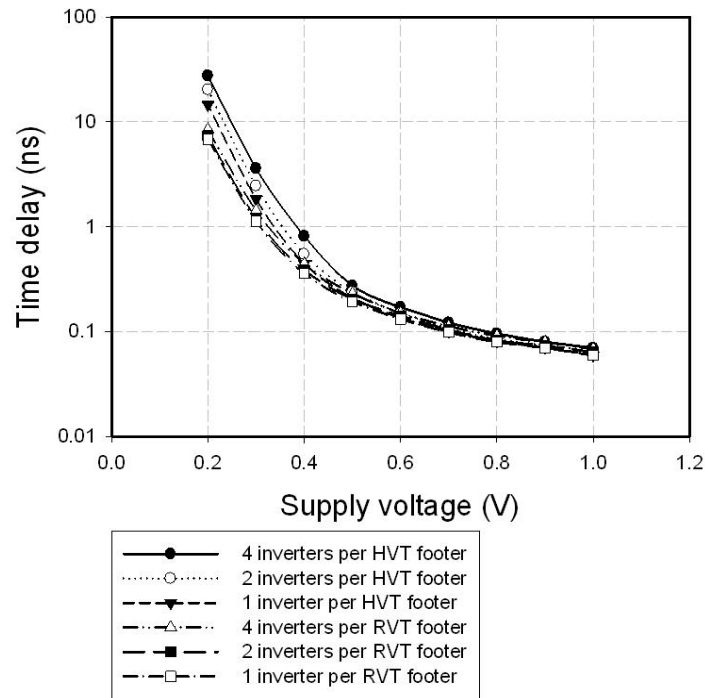


Figure 3.32: Time delay comparisons when applying footer insertion MTCMOS circuit.

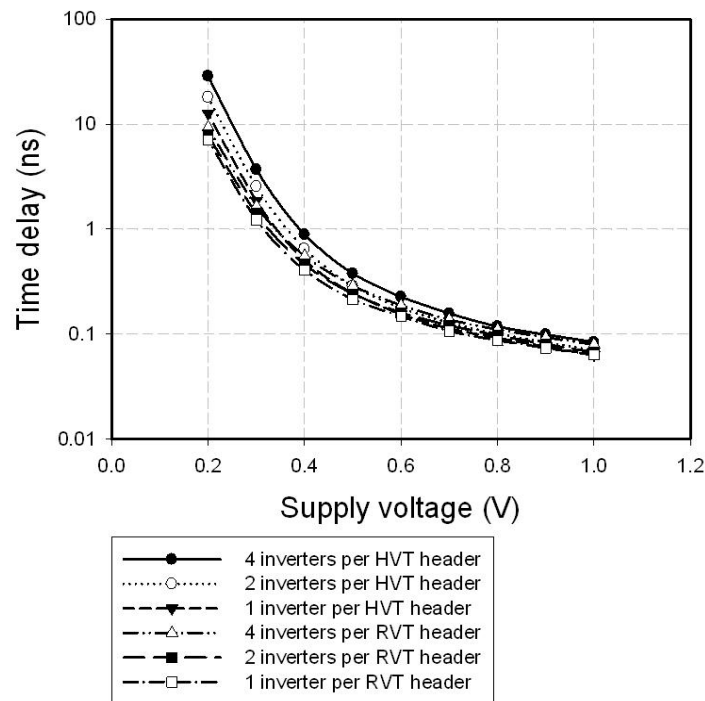


Figure 3.33: Time delay comparisons when applying header insertion MTCMOS circuit.

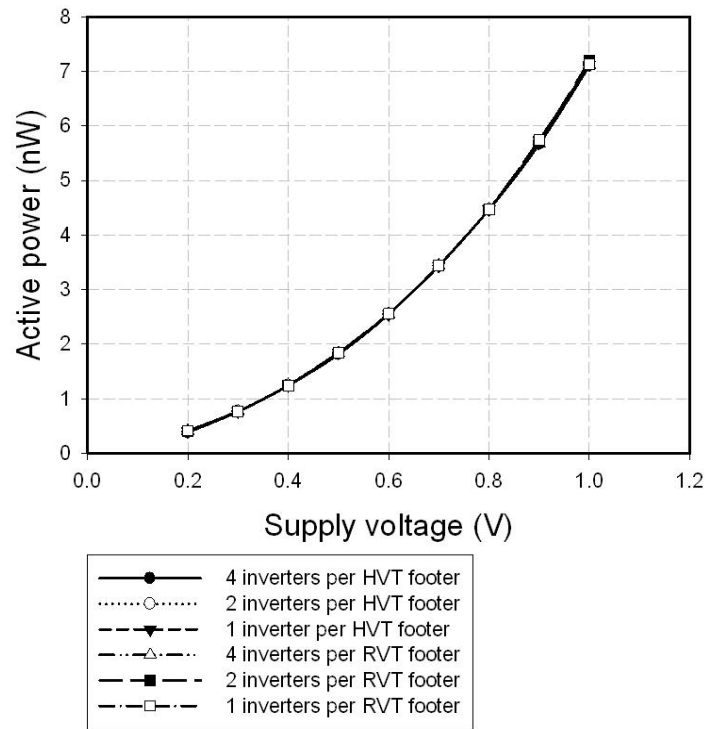


Figure 3.34: Active power comparisons when applying footer insertion MTCMOS circuit.

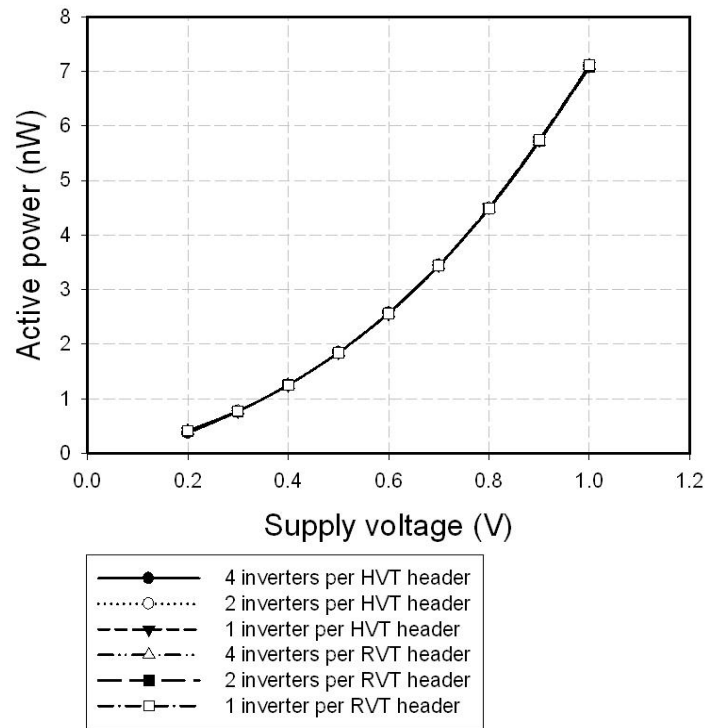


Figure 3.35: Active power comparisons when applying header insertion MTCMOS circuit.

3.4 Summary

In this chapter, power dissipation is first reviewed, including dynamic dissipation, leakage dissipation, and short circuit dissipation. After analyzing power dissipation sources, some useful low power techniques are presented, including supply voltage scaling, transistor stacking, and multiple threshold design. Testing examples and simulation results are demonstrated, which shows the effectiveness of applying these low power techniques. All simulations done in this chapter is based on UMC 90nm CMOS technology.



Chapter 4

Ultra-Low Voltage SRAM Design

4.1 Introduction

Embedded memory typically occupies the largest portion of SoC die area, and has the largest influence on cost, power, performance, and reliability. It is predicted that over 90% of the future chip area is occupied by memory circuits [33]. Thus, robust ultra-low power memory design is a key for ultra-low power systems.

The most widely used form of embedded memory is the *static random access memory (SRAM)*. This chapter begins with the overview of SRAM operation, which will be given in Section 4.2. In Section 4.3, stability issues of SRAM cells, including *hold stability*, *read stability*, and *write ability* will be stated. In Section 4.4, *single-port subthreshold SRAM* will be presented, including prior arts and the proposed subthreshold 10T SRAM with auto-compensation. *Dual-port subthreshold SRAM* will be further presented in Section 4.5, including existing works and the proposed 7T SRAM cell suitable for ultra-low voltage operation and long term activation. The 7T SRAM cell will be the basic storage cell of the FIFO memory proposed in Chapter 5. Finally, summary will be given in Section 4.6.

4.2 Overview of SRAM Operation

Figure 4.1 [34] is a typical SRAM organization. It includes storage cells, row and column decoder for appropriate word selection, sense amplifiers to amplify bitline swing, read/write circuitry for proper read/write control and data buffer.

4.2.1 6T SRAM Cell

Figure 4.2 shows the schematic of the 6T SRAM cell commonly used in practice. The cell uses a single wordline and both true and complementary bitlines. The cell contains a pair of cross-coupled inverters for data storage and an access transistor for each bitline.

For read operation, bitlines are first precharged to high. The wordline is then activated, and one of the bitlines will be pulled down by the cell. For example, in Figure 4.3, $Q = 0$ and $Qb = 1$, BL will therefore be pulled down by transistors MAL-MNL, while BLb stays high. A differential signal is generated on the bitline pair, and the sense amplifier at the read output end will detect this small signal and transforms it into full swing voltage.

For write operation, one bitline is driven high and the other low. The wordline is

then turned on, and data on bitlines will overpower the cell content with the new value. For example, in Figure 4.4, $Q = 0$, $Qb = 1$, $BL = 1$, and $BLb = 0$, Qb will be forced to low, and Q will rise high.

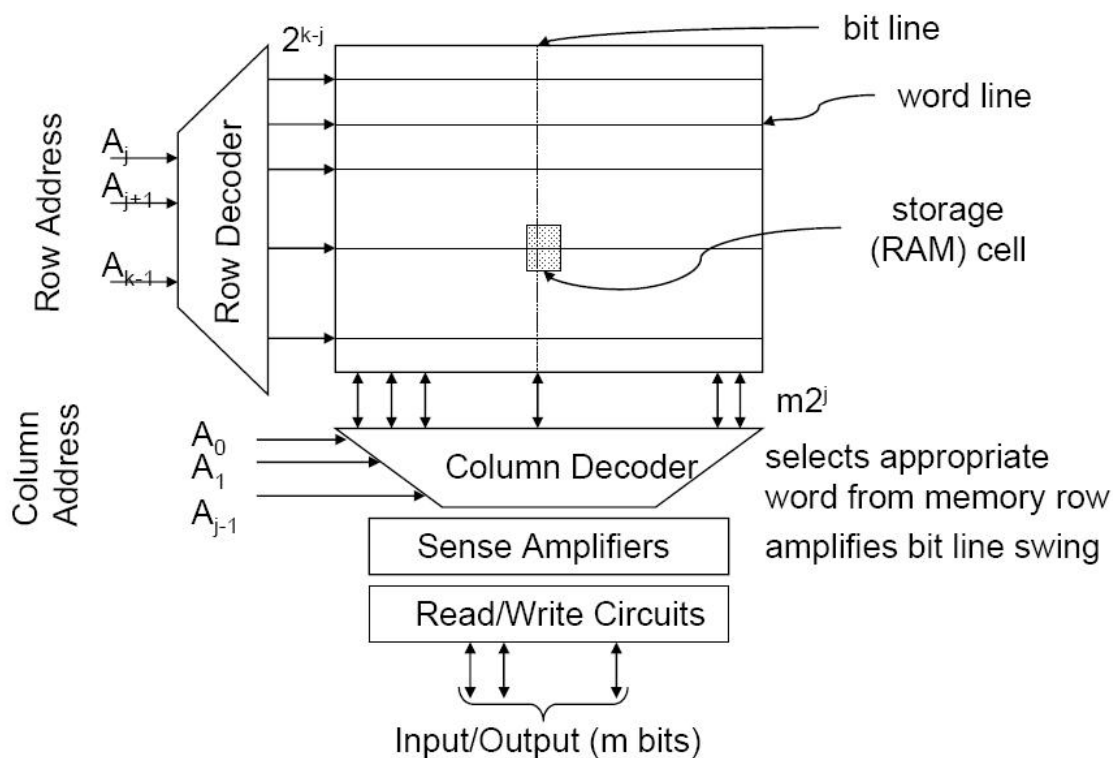


Figure 4.1: SRAM organization.

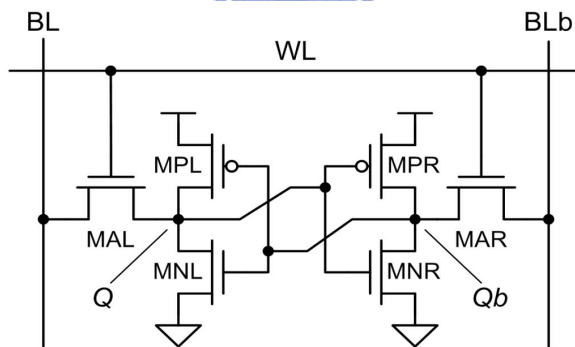


Figure 4.2: Conventional 6T SRAM cell.

4.2.2 Dual-Port SRAM Cell

The difference between single port RAM and dual port RAM is that single port RAM can be accessed at one address at one time, thus only one memory word can be read/write during each clock cycle. Dual port RAM has the ability to simultaneously read and write different memory cells at different addresses.

Figure 4.5 shows the conventional dual-port SRAM cell (DP cell). It is similar to the 6T SRAM cell just described, except it has two more access transistors for an additional

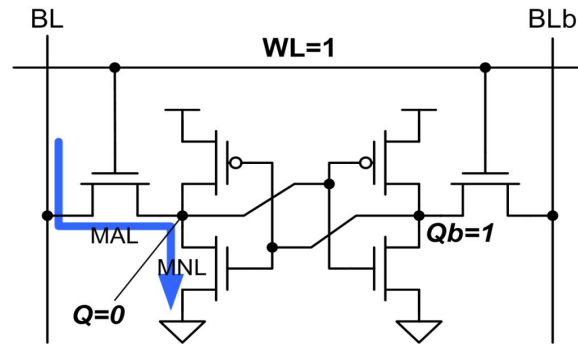


Figure 4.3: Read example of 6T SRAM cell.

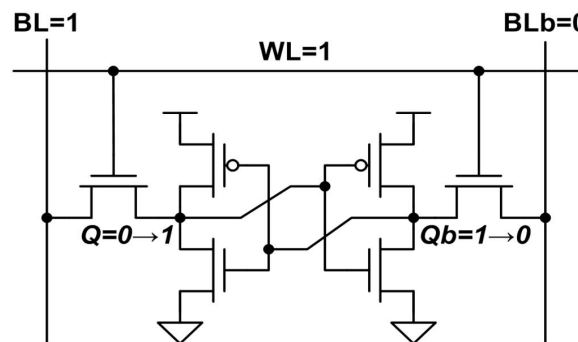


Figure 4.4: Write example of 6T SRAM cell.

port. Read and write operation of DP cell is the same as 6T cell, but extra peripheral circuitry is needed to support the dual-port structure. Notice that data control, such as when to write and when to read, is extremely important for dual-port SRAM since improper data scheduling can result in data conflict or incorrect SRAM function.

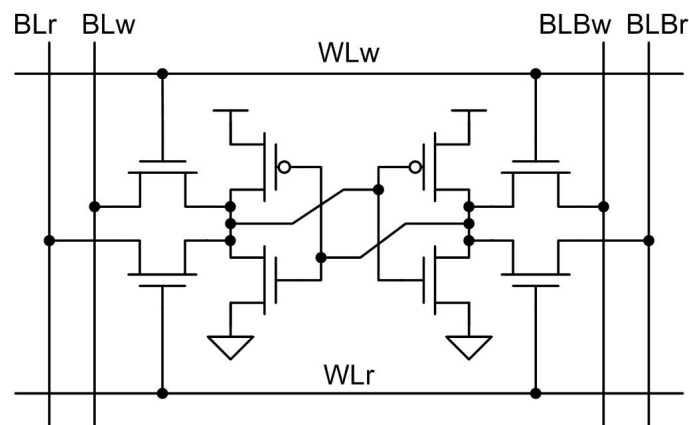


Figure 4.5: Conventional dual-port SRAM cell.

4.2.3 Column Circuitry

Figure 4.6 shows a SRAM column configuration. The precharge circuit is used to precharge the bitlines high and equalize bitline pair before operation. Each column must also contain write drivers and read sensing circuits. Write drivers pull the bitline or its complement

low during write operation. The sense amplifier shown is a commonly used latch type sense amplifier. When the sense amplifier is activated, the cross-coupled inverter pair pulls one output low and the other high through regenerative feedback.

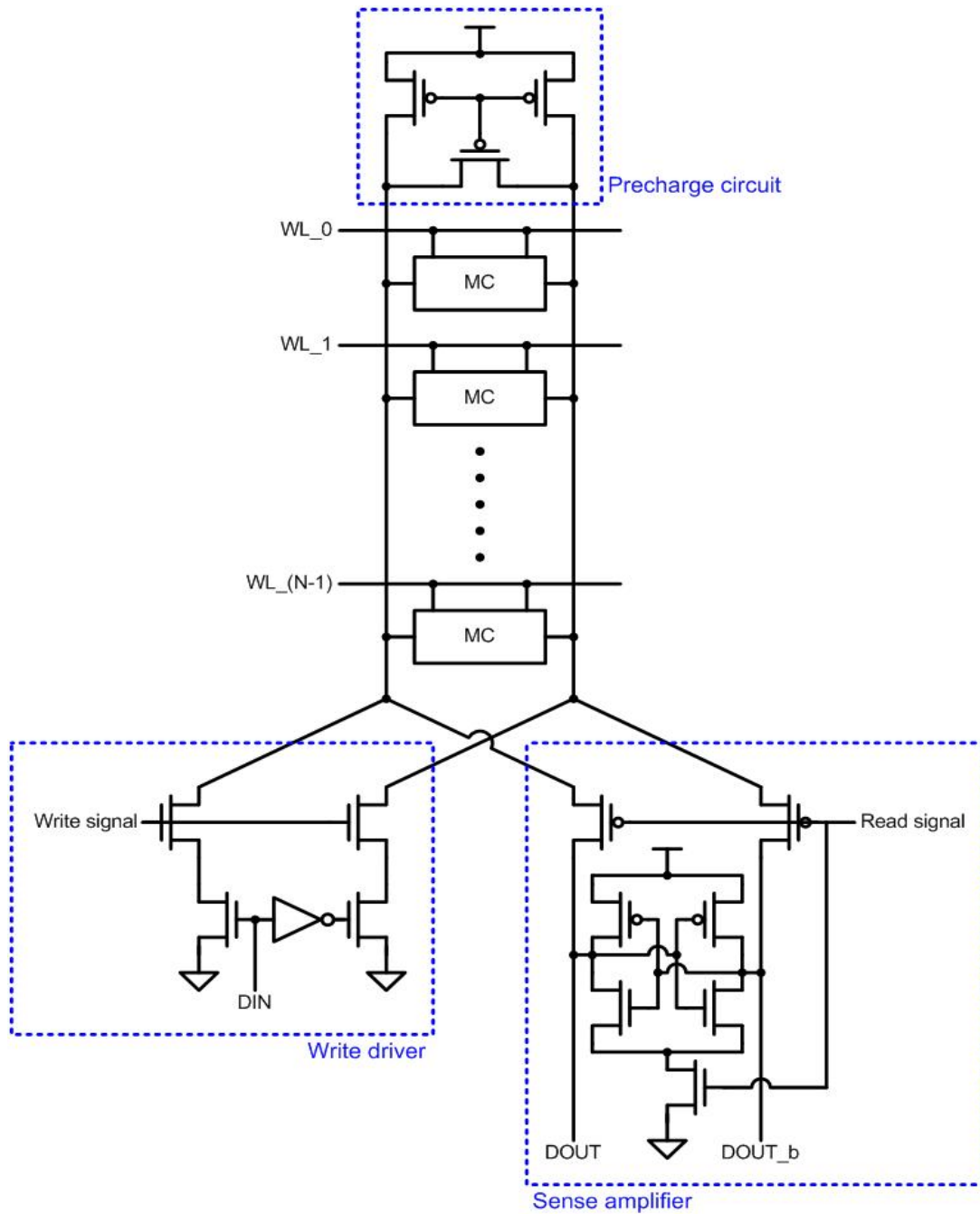


Figure 4.6: An SRAM column.

4.3 SRAM Cell Stability

Reliability has always been a major consideration for SRAM memory cells. As technology scales down, process, voltage, temperature (PVT) variations are becoming an ever increasing concern [35]. Furthermore, in ultra-low supply voltage operation or subthresh-

old operation, SRAM cells are much more sensitive to noise [36], thus the study of SRAM cell stability must be taken seriously. The following of this section will state the most widely adopted SRAM cell stability definition. Analysis of SRAM cell stability in later sections will be based on definitions described here.

4.3.1 Hold Stability

Figure 4.2 is a conventional 6T SRAM bitcell. When the bitcell is holding data, the wordline (WL) is low so that NMOS access transistors (MAL and MAR) are off. The cross-coupled inverters must maintain bi-stable operating points in order to properly hold data. The best measure of the ability of the cross-coupled inverters to maintain their state is the static noise margin (SNM) [37]. The Hold SNM is defined as the maximum value of DC voltage noise that can be tolerated by the SRAM cell without changing the stored bit when the access transistors are off. Figure 4.7 shows the standard setup for modeling Hold SNM. DC noise sources V_N are introduced at each of the internal nodes in the bitcell. Cell stability changes as V_N increases. Figure 4.8 [36], known as the butterfly curve, is the most common way of representing the SNM graphically. The butterfly curve plots the voltage transfer characteristic (VTC) of Inverter R and the inverse VTC of Inverter L. Inverter R and Inverter L are shown in Figure 4.7. The SNM is defined as the length of the side of the largest square that can be embedded inside the lobes of the butterfly curve. When the value of V_N increases, the VTCs move horizontally and/or vertically. When the value of V_N is equal to the value of SNM, the VTCs meet at only two points. Further noise flips the cell content.

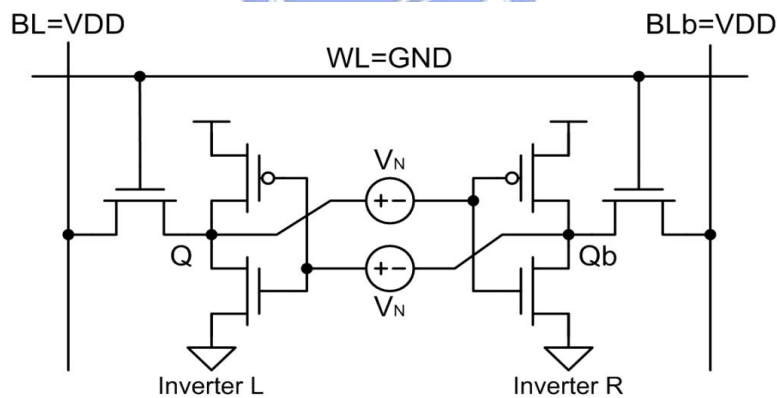


Figure 4.7: Standard setup for finding the Hold SNM.

4.3.2 Read Stability

The most common method to measure read stability is the Read SNM. SNM is defined in the previous subsection, but the setup for Read SNM is different from Hold SNM. Figure 4.9 shows the standard setup for modeling Read SNM. WL is on for read access; BL and BLb are set to V_{DD} to indicate the initial value of bitlines are precharged to high.

In a conventional 6T cell, Read SNM is worse than Hold SNM. During read, the cell begins with the wordline being turned on, with the bitlines initially high. This causes the low node within the cell to rise due to the voltage dividing effect across the access transistors and the pull down transistors. If this node voltage becomes close to the

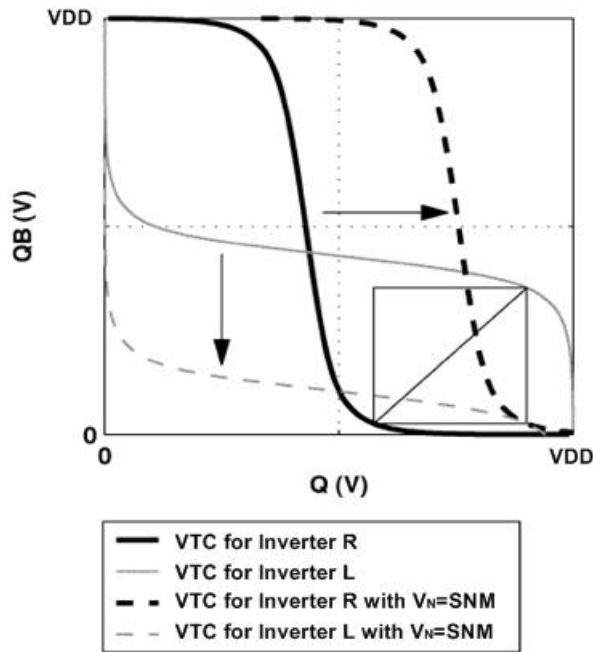


Figure 4.8: Butterfly curve plots for representing SNM. The VTCs of the cross-coupled inverters are represented by the solid curves. The length of the side of the largest embedded square in the butterfly curve is the SNM. When the worst case static noise is applied (e.g., $V_N=SNM$), the bitcell is mono-stable, thus losing its data.

threshold of the pull down devices, process variations combined with noise coupling may flip the state of the cell. Figure 4.10 [36] shows example of butterfly curves during hold and read, revealing the degradation in SNM during read.

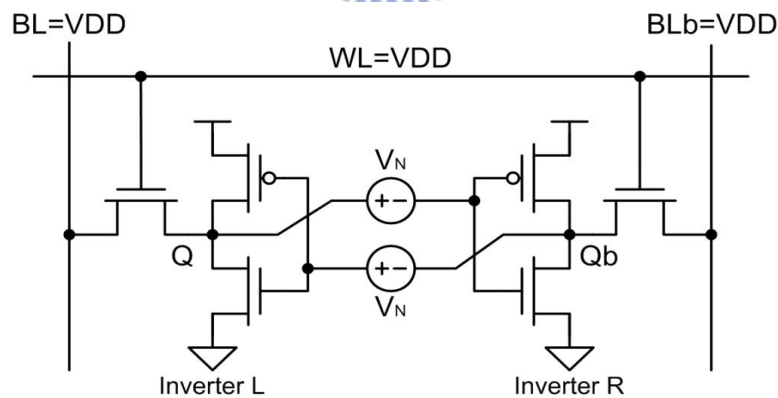


Figure 4.9: Standard setup for finding the Read SNM.

4.3.3 Write Ability

A common way to characterize write ability is the write margin (WM) or write trip point (WTP) [38][39]. WTP defines the maximum voltage on the bitline needed to flip the cell content. Figure 4.11 shows the conceptual setup to measure WTP of 6T SRAM cell. Figure 4.12 [35] shows a corresponding example of finding WTP. As the bitline voltage is

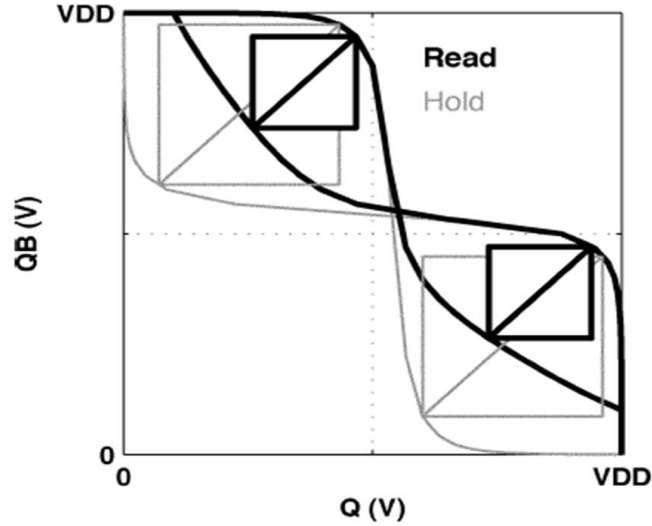


Figure 4.10: Example butterfly curve plots for hold SNM and read SNM.

lowered to a certain level, the cell content is flipped, indicating a successful write. Larger WTP means smaller voltage must be lowered below V_{DD} for successful write, indicating it is easier to write into the cell. If the WTP becomes negative, it means that it is not possible to write into the cell. To sum up, a higher WTP represents better write ability.

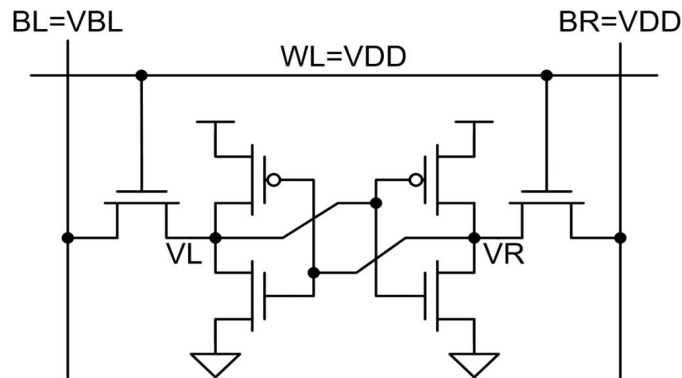


Figure 4.11: Setup for finding WTP.

4.4 Single-Port Subthreshold SRAM Cell

The conventional 6T cell fails to operate in the subthreshold regime because of reduced signal levels and increased variation. For instance, the read SNM requires that the pull down devices be stronger than the access devices, and as shown in Figure 4.13 [40], at low voltages it vanishes and becomes negative. Similarly, the write margin characterizes the ability of the access devices to overpower the pull up devices, and once again, in Figure 4.13, it vanishes at low voltages, indicating write failure.

To overcome the stability degradation of conventional 6T SRAM under subthreshold operation, several robust SRAM cells were presented. In particular, a Schmitt trigger based subthreshold SRAM cell [41] and a 10T subthreshold SRAM cell [42] were proposed to enhance cell stability and inherit differential read scheme for more reliable read function

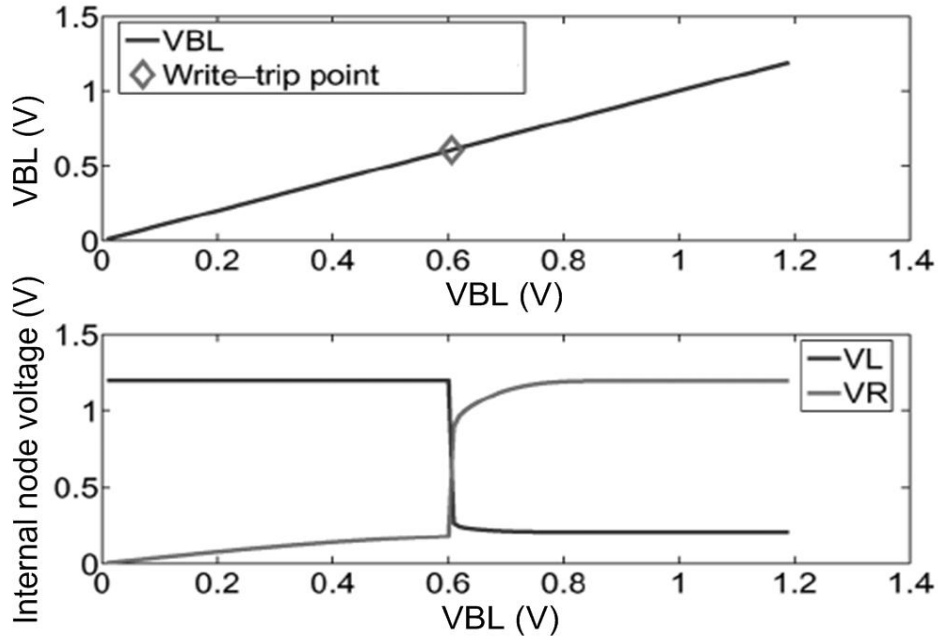


Figure 4.12: Write margin of a SRAM cell, determined by WTP.

at the sensing end. The Schmitt trigger based subthreshold SRAM cell and the 10T subthreshold SRAM cell will be referred to ST cell and 10T cell here after.

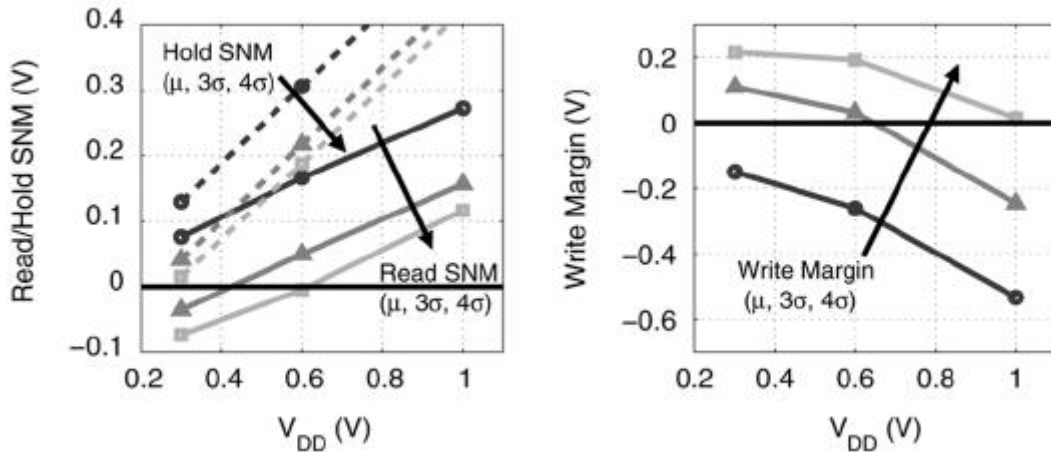


Figure 4.13: Monte Carlo simulations indicating read/hold SNM failures and write margin failures of conventional 6T cell.

4.4.1 Schmitt Trigger Based Subthreshold SRAM

Figure 4.14 shows the schematic of the ST cell. Transistors AXL and AXR are the access transistors. Transistors PL, NL1, NL2, and NFL form one Schmitt trigger (ST) inverter, while PR, NR1, NR2, and NFR form another ST inverter. A ST inverter increases or decreases the switching threshold depending on the direction of the input transition, and therefore, the ST cell is able to adaptively change the flipping point for better data

preservation ability.

Meanwhile, due to the series connected NMOS transistors NL1-NL2 and NR1-NR2, pull down strength is reduced. Reduced pull down strength along with the absence of feedback during 1 to 0 input transition enables ST cell to achieve better write ability.

Although ST cell has better hold SNM and write ability, read SNM improvement is rather limited. This is because the storage nodes VL, VR are not isolated from bitlines, and data inside the cell can be affected by bitline noise during read. Read SNM limitation could be eliminated by isolating storage nodes from bitlines.

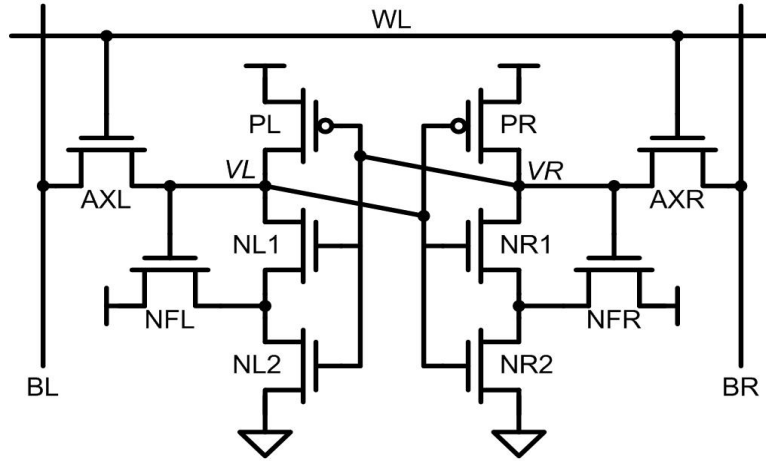


Figure 4.14: Schmitt trigger based subthreshold SRAM cell (ST cell).

4.4.2 Single-Port 10T Subthreshold SRAM

Schematic of the single-port fully differential 10T cell is shown in Figure 4.15. During read, WL1 is off, WL2 is on, and VGND is low to activate the read path formed by transistors AL2-NL or AR2-NR. Differential signals on bitline pair are therefore generated, and a differential sense amplifier at the read output end amplifies the signal to full swing voltage. This read scheme isolates storage nodes from bitlines, and read SNM is improved. During hold, wordlines are turned off, and data retention solely depends on the conventional inverter pair latch. The 10T cell thus has similar hold SNM value compared to conventional 6T cell. During write, both WL1 and WL2 are on to form a write path from bitlines to storage nodes. In order to compensate weak write ability, VGND is high during write, and wordlines should be boosted higher than the original logic-high. This mechanism improves write margin with the expense of additional charge pump circuit or additional voltage source. Furthermore, the 10T SRAM structure can be used for bit-interleaving to reduce the multiple-bit soft errors.

4.4.3 Proposed Subthreshold SRAM Cell with Auto-Compensation

The proposed robust subthreshold SRAM cell is shown in Figure 4.16. It is composed of ten transistors, including a cross-coupled inverter pair (PL, PR, NL, and NR), access transistors (AL1, AL2, AR1, and AR2), and read/auto-compensation transistors (RL and RR).

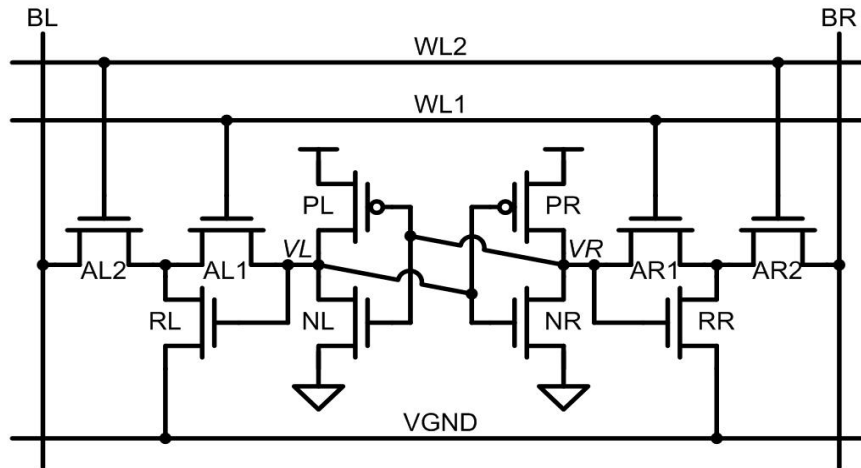


Figure 4.15: Single-port 10T subthreshold SRAM cell (10T cell).

Auto-Compensation

The auto-compensation mechanism is used for better data preservation in hold mode. During hold, WL1 is on, WL2 is off, VGND is low, and a current path to GND is formed by AL1-RL or AR1-RR to supply the node storing data-0. Figure 4.17 is a conceptual illustration showing how auto-compensation works. Transistors with red forbidden signs mean that they are turned off. The blue line with arrow represents the feedback path for auto-compensation. If storage nodes are disturbed, the feedback system will automatically compensate the imposed noise, and hold the storage nodes to their proper value. As a result, better hold SNM is achieved.

Read Operation

During read, WL1 is off, WL2 is on, and VGND is low. With AL1 and AR1 turned off, storage nodes are isolated from bitlines to prevent bitline interference, where AL2-RL and AR2-RR act as read buffers. An example of read operation is demonstrated in Figure 4.18. If VR stores data-1 and VL stores data-0, AR2-RR will form a read path from BR to GND, where BL stays at the precharged state. Differential sense amplifier will detect the generated differential signal on bitline pair and amplify the signal into full swing voltage. The proposed cell structure is able to inherit fully differential read scheme for more reliable read operation, and enable better SRAM cell read SNM.

Write Operation

To improve write ability under subthreshold operation, write assist technique was proposed [40]. The basic idea of write assist is to destroy the cell content during write, so that data on bitlines are easier to write over the original cell content. A modified write assist is applied by controlling VGND. During write, both WL1 and WL2 are on, and VGND is high, as illustrated in Figure 4.19. With VGND at the level of VDD, the cell loses data retention ability. Furthermore, due to imperfect data-1 transformation characteristic of NMOS transistors, current through RL and RR will have weaker strength to interfere the normal write operation. As a result, write operation can be easily achieved. In the proposed scheme, charge pump or additional voltage source are not needed to gain write ability improvement.

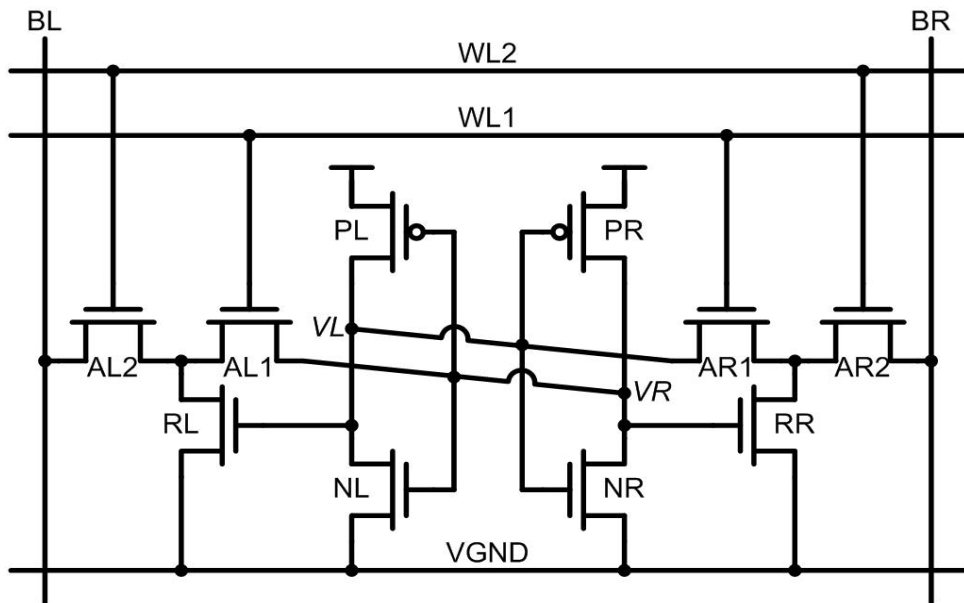


Figure 4.16: Proposed subthreshold SRAM cell with auto-compensation (AC cell).

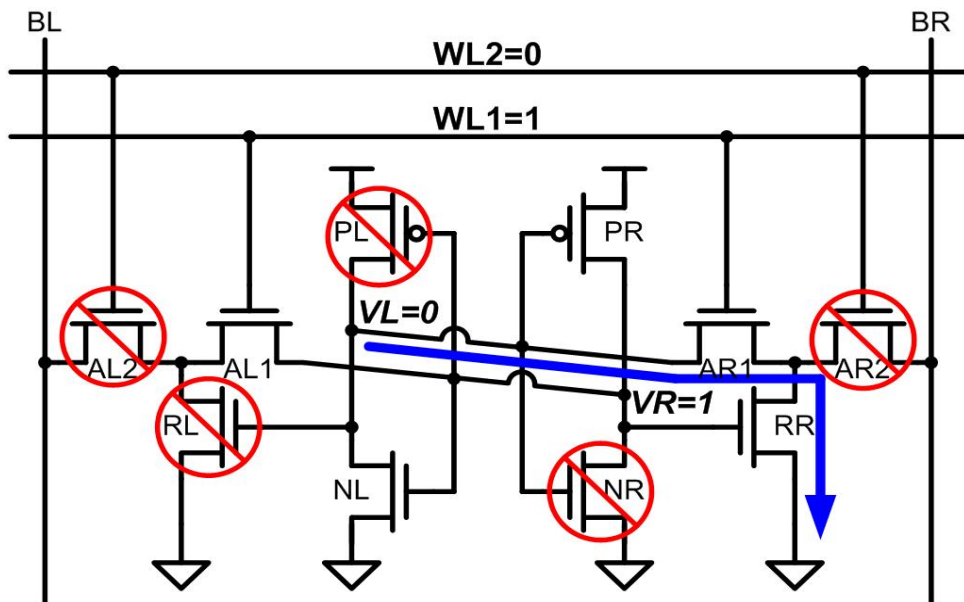


Figure 4.17: Example of auto-compensation. The feedback system generated by AR1-RR holds node VL from being flipped.

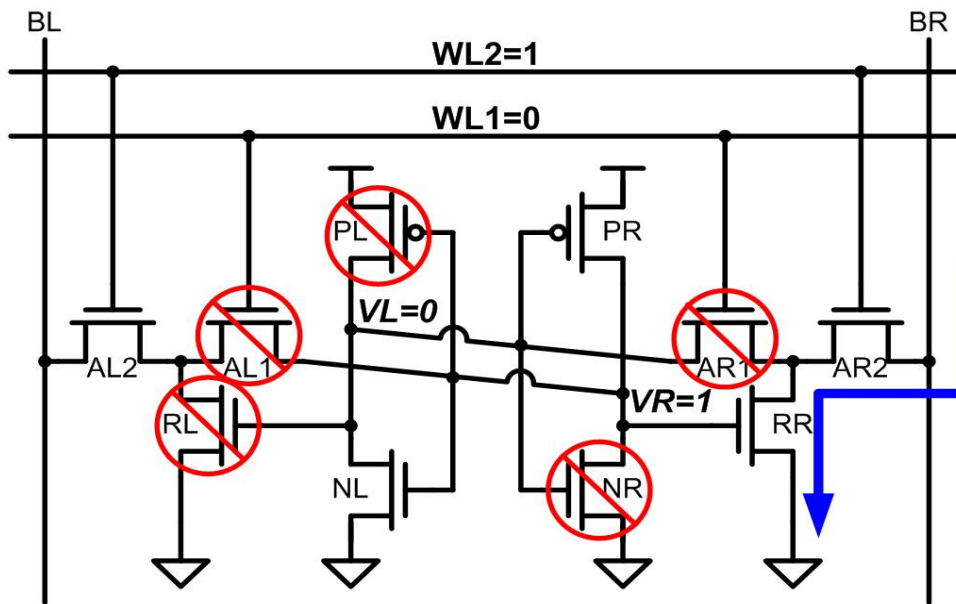


Figure 4.18: Example of the read operation. AL1 and AR1 isolate storage nodes from bitlines. AR2-RR forms a read path from BR to GND.

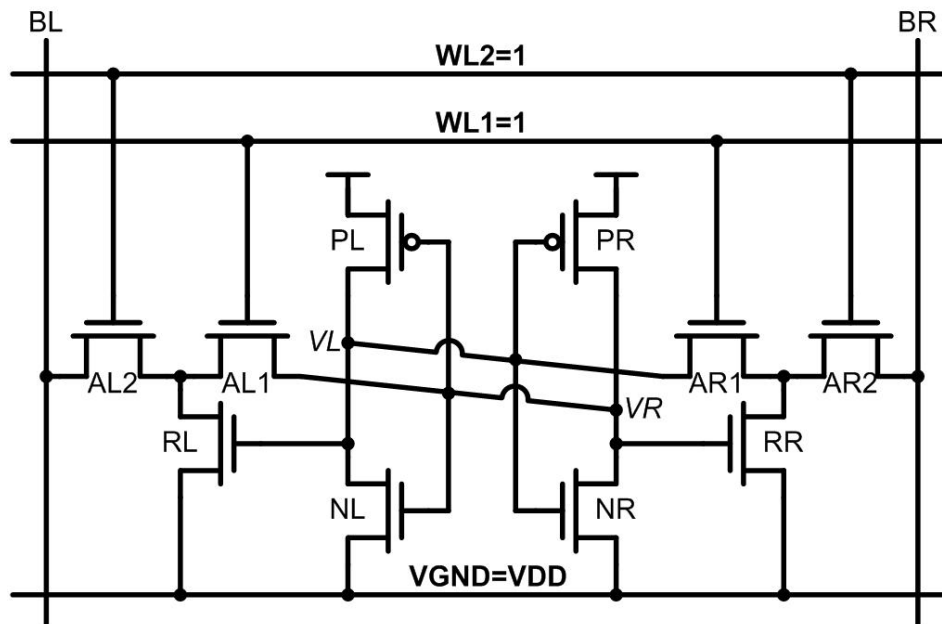


Figure 4.19: Write operation of the AC cell.

Cell Operation Summary

Table 4.1 shows the summary of the cell operation in hold, read, and write mode.

Table 4.1: Summary of the AC cell operation.

	WL1	WL2	VGND
Hold	High	Low	GND
Read	Low	High	GND
Write	High	High	VDD

4.4.4 Simulation Results

In this section, hold stability, read stability, write ability, and leakage power are compared between the conventional 6T cell, the ST cell, the 10T cell, and the proposed AC cell. All simulation results are based on UMC 90nm CMOS technology using HSPICE.

Hold Stability

Static noise margin (SNM) is the most common way to measure hold stability and read stability. SNM defines the largest noise that can be imposed to the storage nodes before the cell content is flipped. Figure 4.20 shows the hold SNM versus supply voltage of various cell structures. The hold SNM of 6T and 10T cell is almost the same due to similar data preservation structure. ST cell has better hold SNM and has the best hold SNM performance in super threshold regime. The AC cell also demonstrate better hold SNM compared to 6T and 10T cell, and has the best hold SNM performance in subthreshold regime.

Figure 4.21 shows the Monte Carlo simulations of hold SNM comparisons under 200mV supply voltage. It is observed that the AC cell gives higher mean hold SNM, with 1.22X, 1.09X, and 1.21X improvement compared to 6T, ST, and 10T cell, respectively. The AC cell also gives better 3σ hold SNM, with 1.15X, 1.12X, and 1.15X improvement compared to 6T, ST, and 10T cell, respectively. The above mentioned observations are demonstrated in Figure 4.22, which confirms the effectiveness of auto-compensation.

Read Stability

Read stability is characterized by the read SNM. Figure 4.23 shows the read SNM comparisons under different supply voltage. As mentioned in the beginning of Section 4.4, 6T cell exhibits serious read SNM degradation. It is also mentioned in Section 4.4.1 that the ST cell has limited read SNM improvement.

Figure 4.24 shows the Monte Carlo simulations of read SNM under 200mV supply voltage. Figure 4.25 shows the mean read SNM and 3σ read SNM comparisons extracted from Fig. 8. Due to storage nodes isolation and the assistance of read buffers (AL2-RL; AR2-RR), the AC cell has 2.09X, 1.55X improved mean read SNM, and 2.54X, 1.65X improved 3σ read SNM compared to 6T and ST cell, respectively. AC cell and 10T cell have nearly equal mean read SNM and 3σ read SNM due to similar read structure.

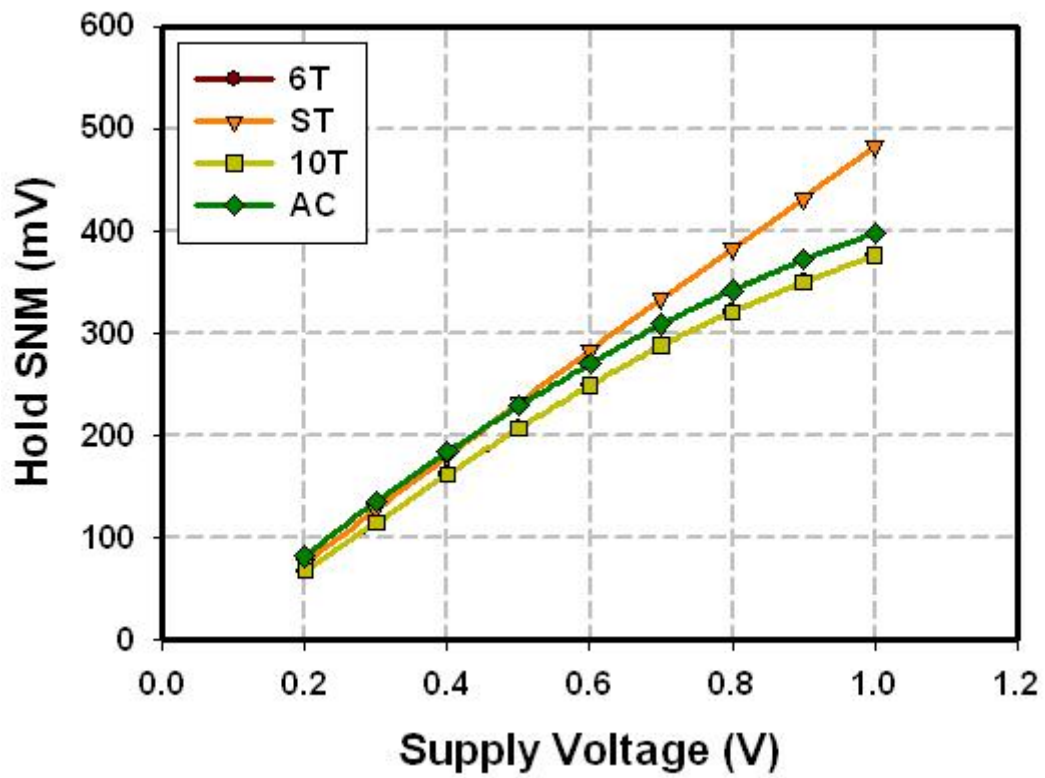


Figure 4.20: Hold SNM comparisons under different supply voltage.

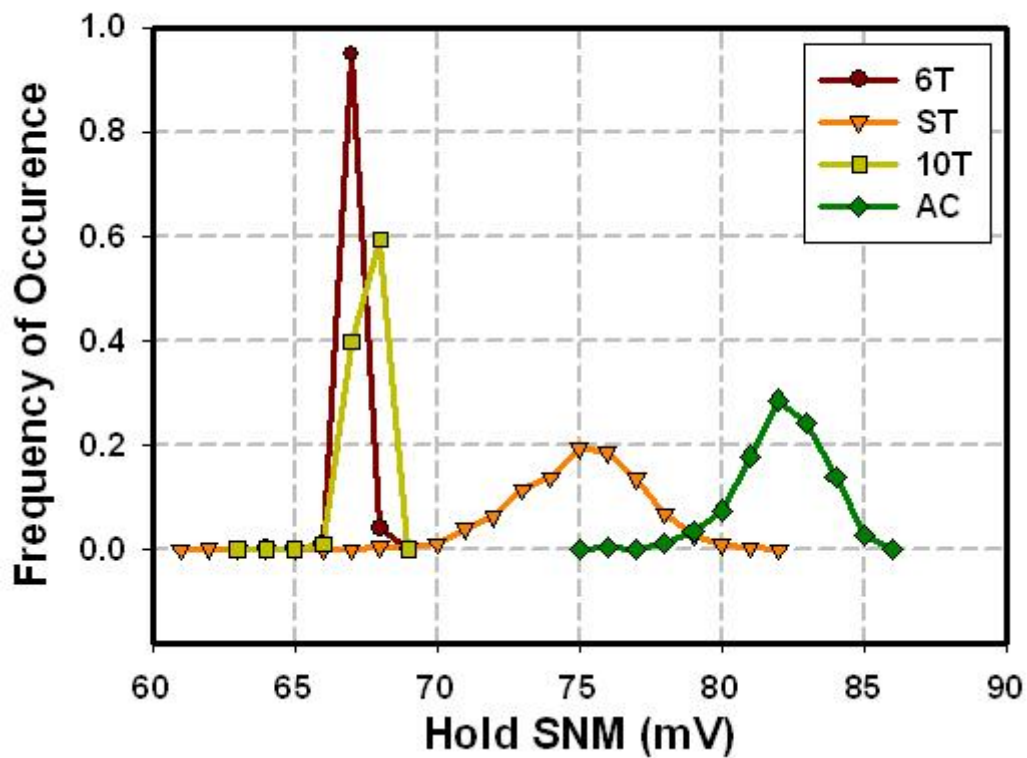


Figure 4.21: Distribution of hold SNM at 200mV.

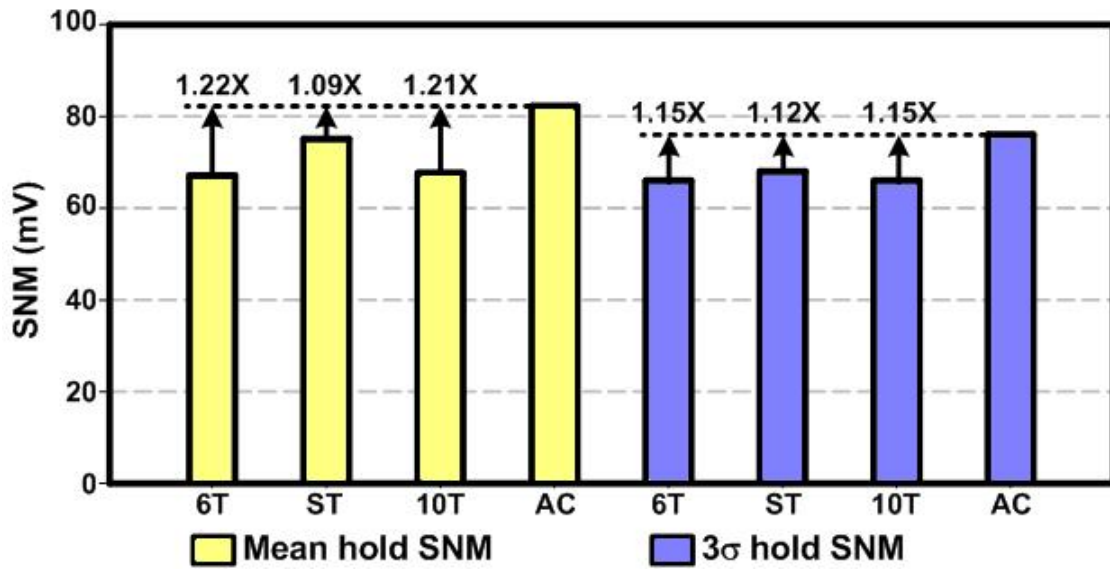


Figure 4.22: Hold SNM comparisons at 200mV.

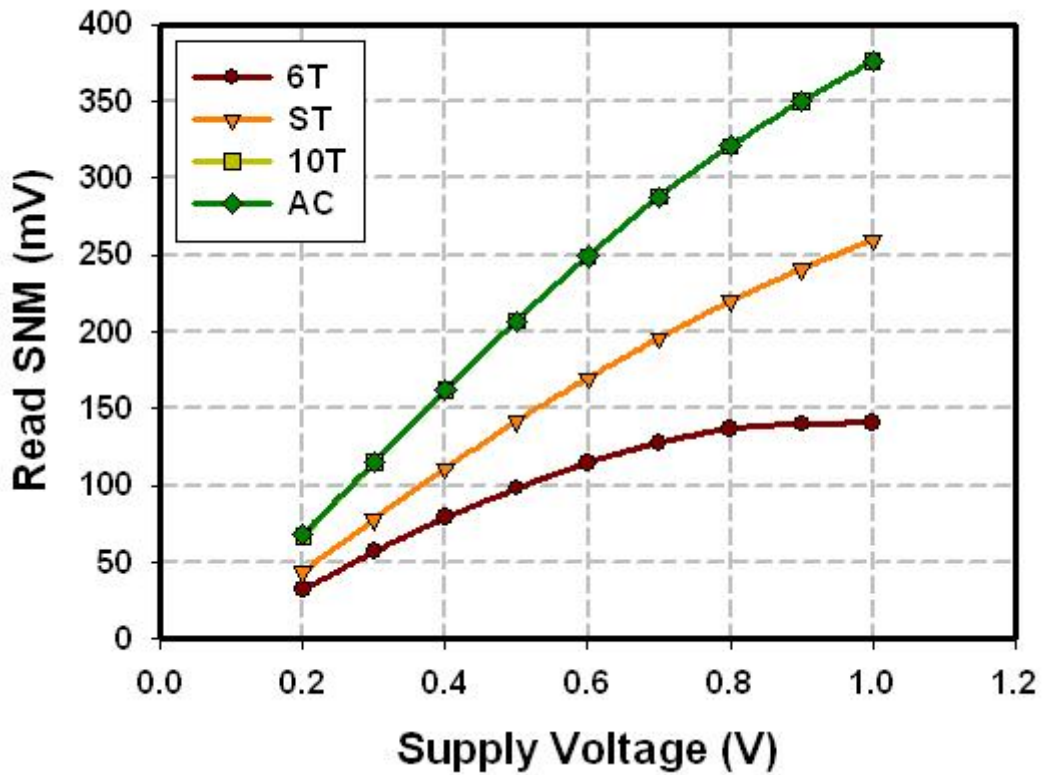


Figure 4.23: Read SNM comparisons under different supply voltage.

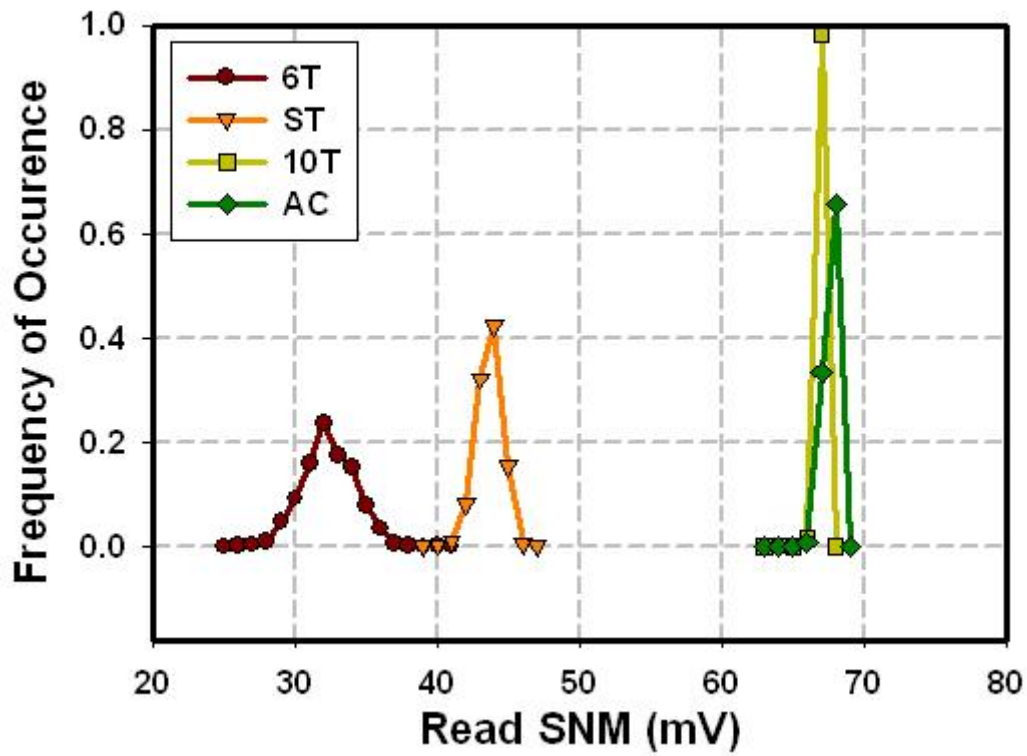


Figure 4.24: Distribution of read SNM at 200mV.

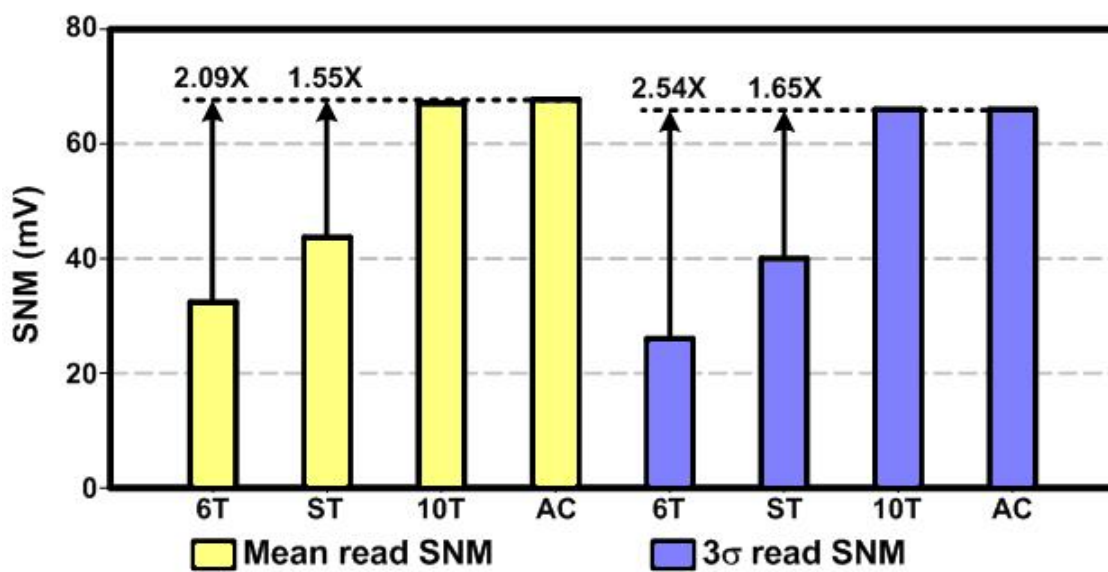


Figure 4.25: Read SNM comparisons at 200mV.

Write Ability

Write ability indicates how easy or difficult it is to write to a cell, which can be determined by the write trip point. Write trip point is defined as the maximum voltage on the bitline needed to flip the cell content. Higher bitline voltage means that the cell can be easier written. If the write trip point is negative, successful write operation is unable to achieve. Figure 4.26 shows the write trip point comparisons under different supply voltage, indicating the write assist technique improves write ability significantly. Figure 4.27 shows the Monte Carlo simulation of write trip point comparisons under 200mV supply voltage. Mean write trip point and 3σ write trip point comparisons are further shown in Figure 4.28. The proposed scheme achieves 2.03X, 1.82X improved mean write trip point, and 2.91X, 2.31X improved 3σ write trip point compared to 6T and ST cell, respectively. For the 10T cell, it is very difficult to write to the cell without leveling up the wordline voltage. Monte Carlo simulation shows that about 33% write trip point occurrence are negative, indicating a high write failure rate for 10T cell without boosting the wordline voltage. The proposed scheme has 41.57X mean write trip point improvement compared to the 10T cell.

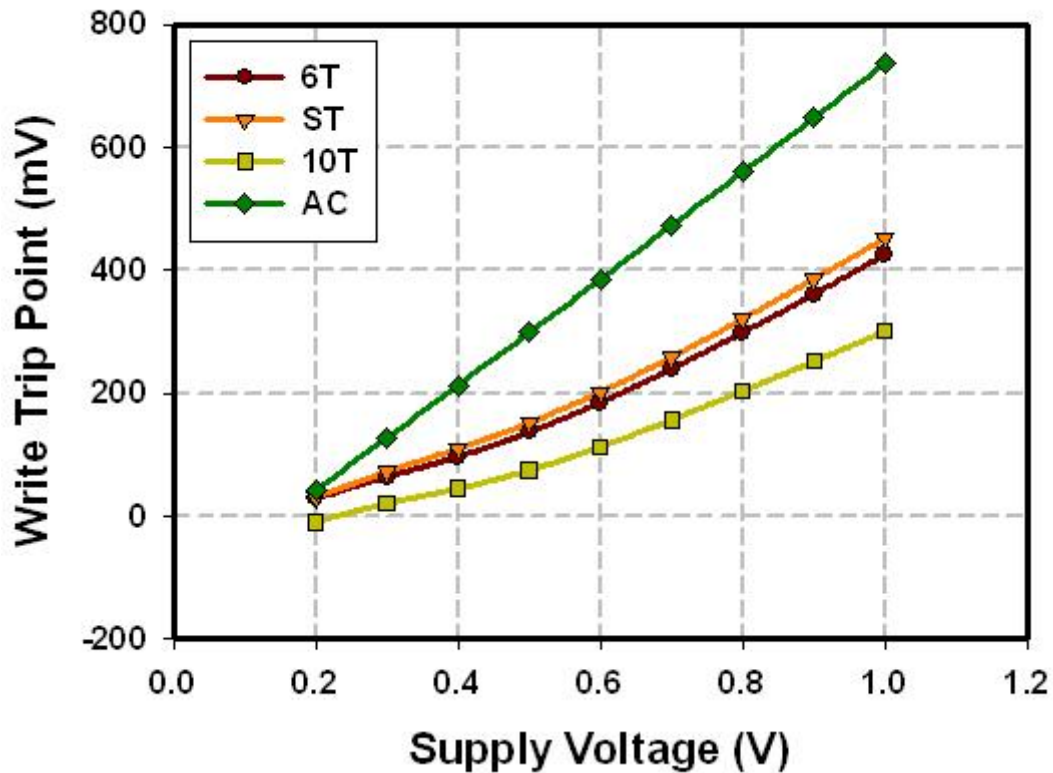


Figure 4.26: Write trip point comparisons under different supply voltage.

Leakage Power

AC cell consumes slightly larger leakage power during standby, as shown in Figure 4.29. This is due to the on transistors AL1 and AR1 for auto-compensation (Figure 4.17). Nevertheless, the leakage power overhead imposed by auto-compensation is an acceptable

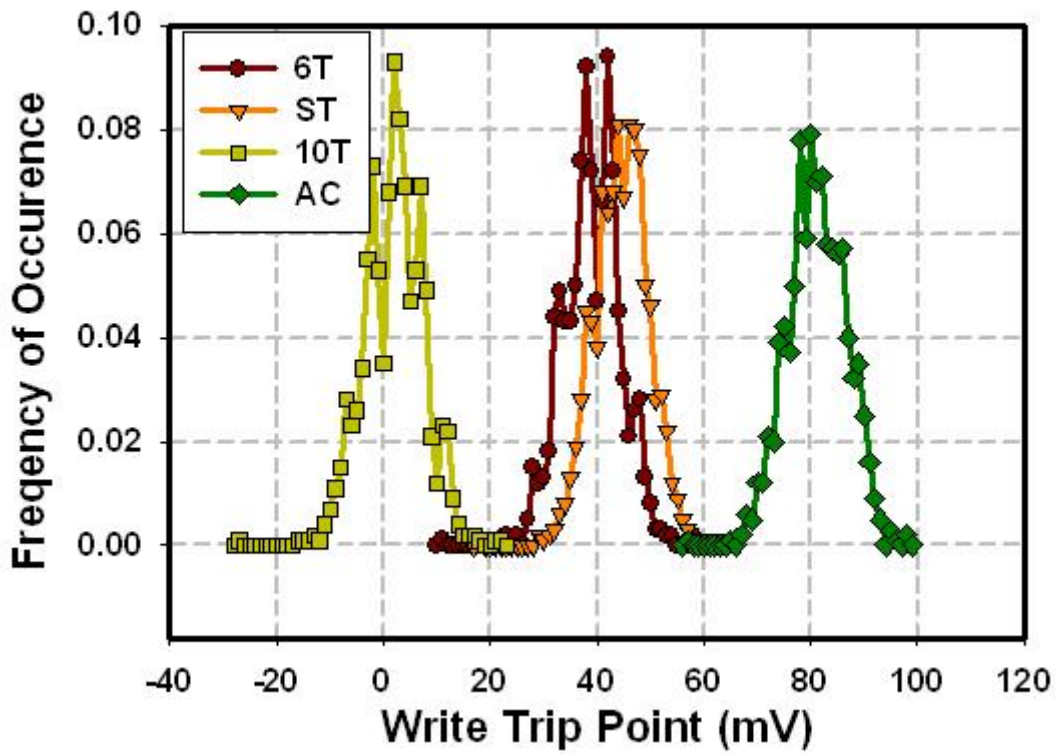


Figure 4.27: Distribution of write trip point at 200mV.

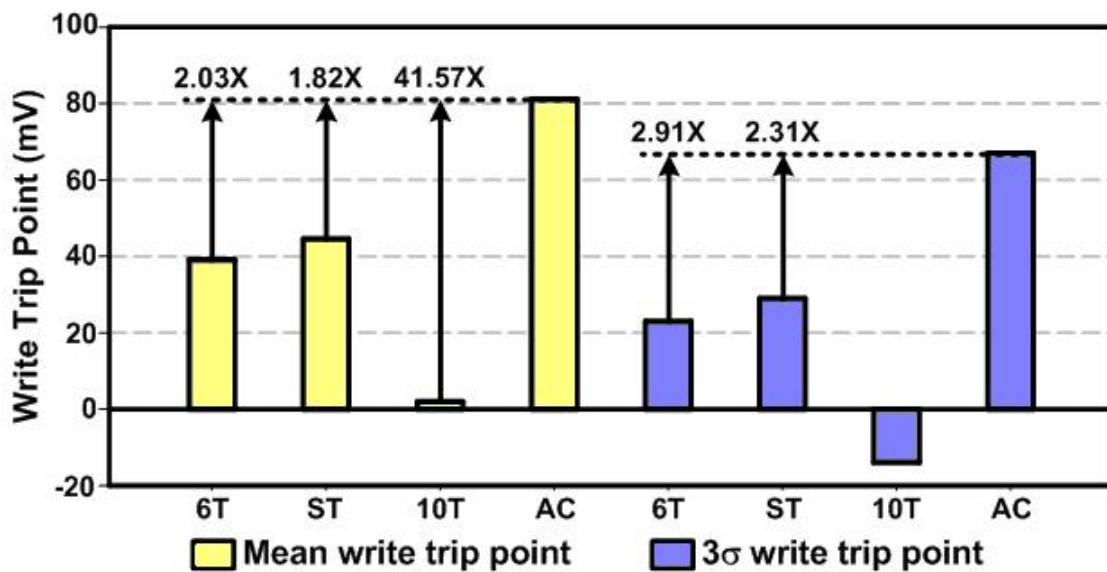


Figure 4.28: Write trip point comparisons at 200mV.

tradeoff, since better hold stability is ensured. As a result, the proposed AC cell has higher potential for ultra-low voltage data retention in future nanoscaled technologies.

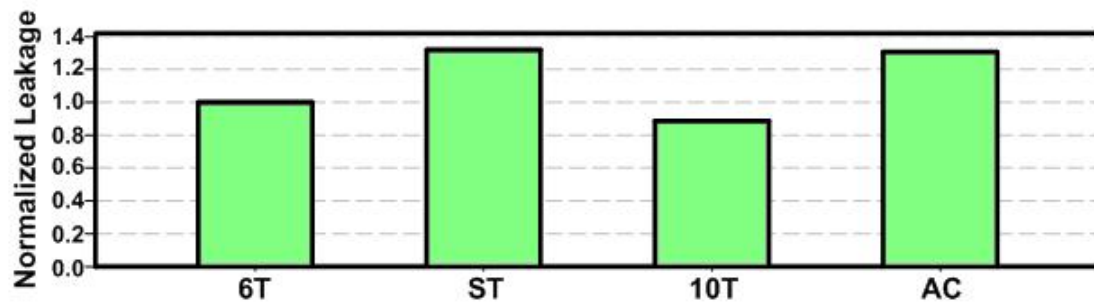


Figure 4.29: Leakage comparisons at 200mV. Leakage is normalized to the leakage of 6T cell.

4.5 Dual-Port Subthreshold SRAM Cell

For some applications, such as FIFO memories, dual-port SRAM cell is essential. Conventional DP SRAM cell described in Section 4.2.2, like conventional single-port 6T SRAM cell, faces stability problem under subthreshold operation. Several subthreshold dual-port SRAM designs were proposed for more robust operation. They will be presented in the following subsections.

4.5.1 8T Subthreshold SRAM

The fundamental stability problem in conventional SRAM cell is that in read condition, an access transistor pulls the data-0 storage node up to a non-zero value. As shown in Figure 4.30, adding two transistors (MR1 and MR2) to a 6T cell that serves as read buffer provides a read mechanism that does not disturb the internal nodes of the cell, thereby eliminating the worst-case stability condition [43]. Without read disturbs, the 8T cell provides significantly larger SNM than 6T cell. Furthermore, this 8T topology requires separate read and write wordlines and can accommodate dual-port operation with separate read and write bitlines.

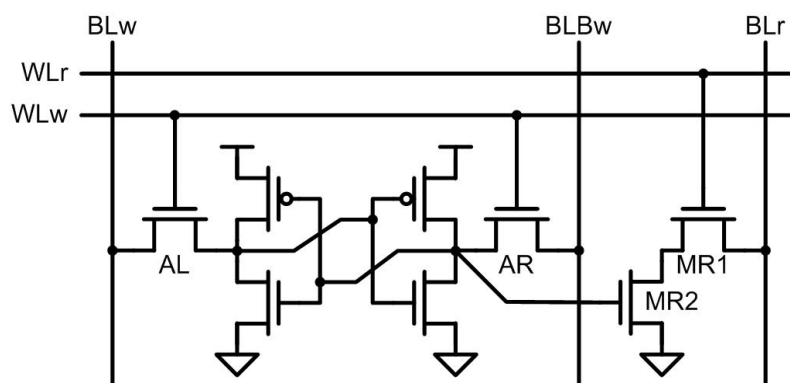


Figure 4.30: 8T Subthreshold SRAM cell.

The 8T cell operation is described in the following. Write access of the 8T cell is similar to conventional 6T cell write operation, which occurs through the write access transistors AL and AR from the write bitlines (BLw and BLBw). Read access is single-ended and occurs on a separate read bitline (BLr), which is precharged to V_{DD} prior to read access.

4.5.2 Dual-Port 10T Subthreshold SRAM

The key idea of eliminating the read SNM limitation is to isolate storage nodes during read access. The 8T cell just described features the read SNM elimination, but leakage imposed by the read buffer causes extra power consumption. Moreover, leakage imposed by the read buffer will increase the read failure rate, because bitline leakage from the unaccessed cells can rival the read current of the accessed cell making it hard to distinguish between the bitline high and low levels. To reduce the read failure rate, one method is to limit the number of cells on each bitline; another approach is by reducing the bitline leakage. This can be done by modifying the cell structure. Therefore, dual-port 10T subthreshold SRAM designs were proposed, as shown in Figure 4.31 [44] and Figure 4.32 [45]. The cell shown in Figure 4.31 will be referred to 10T_C (C represents Chandrakasan, the author of [44]), and the cell shown in Figure 4.32 will be referred to 10T_K (K represents Kim, the author of [45]).

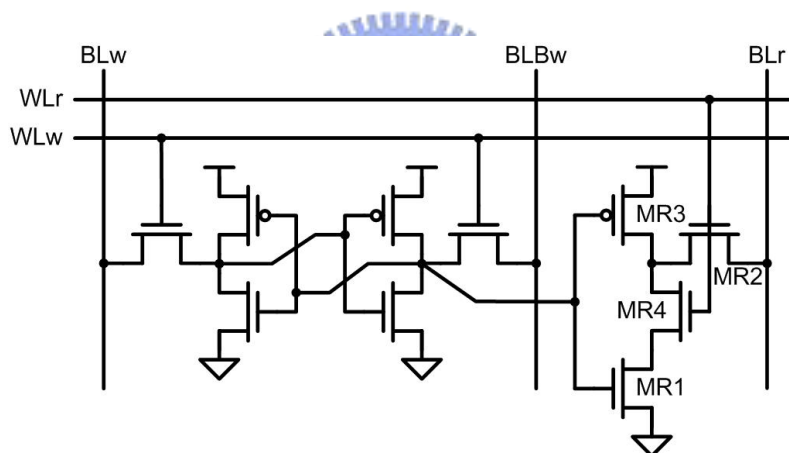


Figure 4.31: Dual-port subthreshold 10T SRAM cell (10T_C).

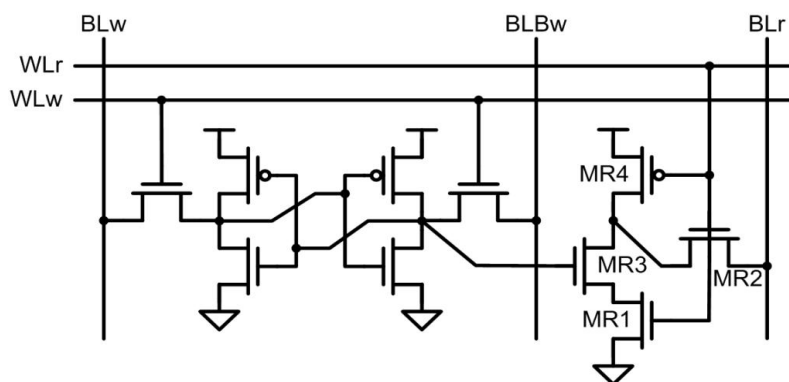


Figure 4.32: Dual-port subthreshold 10T SRAM cell (10T_K).

For the 10T_C cell, write access operates like the 8T cell. Transistors MR1, MR2, MR3, and MR4 implement a buffer used for reading. Read access is single-ended and occurs on the read bitline (BLr), which is precharged prior to read access. Transistors MR1, MR2, MR3, and MR4 remove the problem of read SNM by buffering the stored data during read. Moreover, due to the stack effect, this buffer style reduces leakage current and allows more cells on a bitline during read.

Write operation of the 10T_K cell is also like the 8T cell. Transistors MR1, MR2, MR3, and MR4 form a read buffer, which eliminates the read SNM. Read access is single ended and occurs on a read bitline (BLr), which is precharged prior to read access. Unlike the 10T_C cell, when read is disabled, the leakage path of 10T_K cell's read buffer always flow from V_{DD} to BLr, regardless of the data stored in the SRAM cell. This characteristic makes sense amplifier more easy to distinguish the whether data-0 or data-1 should be read, and decrease the read failure probability. As a result, more cells can be placed on a bitline, enabling high-density SRAM macro.

4.5.3 Proposed Dual- V_T Subthreshold 7T SRAM cell

Voltage swing on bitlines is a crucial active power dissipation source in memory architectures. The proposed dual- V_T 7T SRAM cell, shown in Figure 4.33 features separated single-ended read/write port, which eliminates bitline overhead. However, in ultra-low voltage SRAM design, write ability degrades, where in single-ended write port scheme, the degradation becomes more severe. To compensate the degraded write ability, dual- V_T transistors are applied. Due to voltage divider created by M1 and M3, it is more difficult to write data-1, thus, assigning low- V_T M1 and high- V_T M3 assuages the voltage divider effect and improves write ability.

Write assist technique, as stated in [40], can be further applied for more write ability improvement. This technique will be implemented in the ultra-low power FIFO design proposed in the next chapter.

Hold stability and read stability are also important for SRAM that operates under ultra-low voltage. Hold stability can be improved by assigning high- V_T transistors to the cross coupled inverters M2, M3, M4, and M5, since high threshold increases the flipping voltage level, which makes storage nodes more immune to noise. Read stability can be improved by isolating storage nodes from read bitlines during read operation, which is done by inserting read buffer M6 and M7. Note that although applying high- V_T M6 and M7 results in smaller pull down current, thereby increasing read delay, smaller leakage is achieved, which reduces power consumption and decreases read failure rate at the read output end.

The read failure issue caused by the read bitline leakage can be solved by pulling the feet of all the unaccessed read buffers (M6 and M7 in the 7T cell case) to high [40]. Because the read bitline is precharged to V_{DD} , cross voltage of the read buffer will be minimized when the feet of the read buffer is high, thus reducing leakage current. This implementation will be shown in the next chapter.

4.5.4 Simulation Results

In this section, hold stability, read stability, write ability, and leakage power are compared between the conventional scheme and the proposed scheme. All simulations results are based on UMC 90nm CMOS technology using HSPICE.

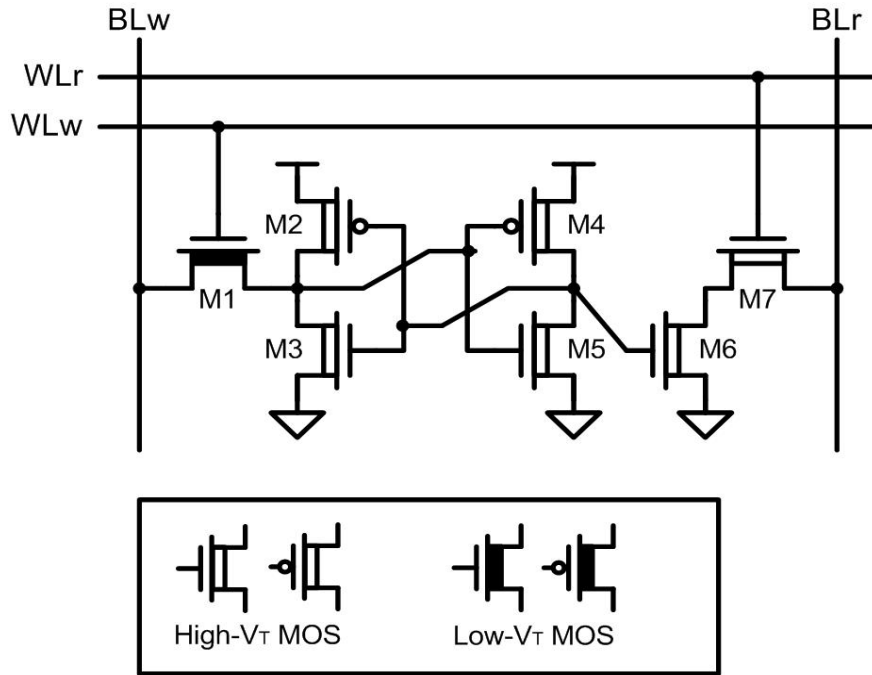


Figure 4.33: Dual- V_T subthreshold 7T SRAM cell.

Hold Stability

Hold stability of the proposed dual- V_T 7T SRAM cell is compared to conventional dual-port SRAM cell implemented with regular- V_T (Figure 4.5). Figure 4.34 shows the hold SNM comparisons versus different supply voltage. Figure 4.35 shows the Monte Carlo simulations of hold SNM comparisons under 300mV supply voltage. Mean hold SNM and 3σ hold SNM are further shown in Figure 4.36. It is obvious that the proposed scheme exhibits better hold stability, with 1.07X mean hold SNM improvement and 1.13X 3σ hold SNM improvement, confirming the effectiveness of utilizing high- V_T inverter latch. It is also observed that the proposed scheme has smaller deviation, therefore promising better process variation immunity.

Read Stability

Read stability of the proposed dual- V_T 7T SRAM cell is compared to conventional dual-port SRAM cell implemented with regular- V_T (Figure 4.5). Figure 4.37 shows the read SNM comparisons versus different supply voltage. Figure 4.38 shows the Monte Carlo simulations of read SNM comparisons under 300mV supply voltage. Mean read SNM and 3σ read SNM are further shown in Figure 4.39. It is obvious that the proposed scheme exhibits better read stability, with 1.69X mean read SNM improvement and 2.22X 3σ read SNM improvement, revealing the effectiveness of isolating storage nodes from read bitline. It is also observed that the proposed scheme has much smaller deviation, therefore promising better process variation immunity.

Write Ability

Write ability of the proposed dual- V_T 7T SRAM cell is compared to conventional dual-port SRAM cell implemented with regular- V_T (Figure 4.5). Figure 4.40 shows the write

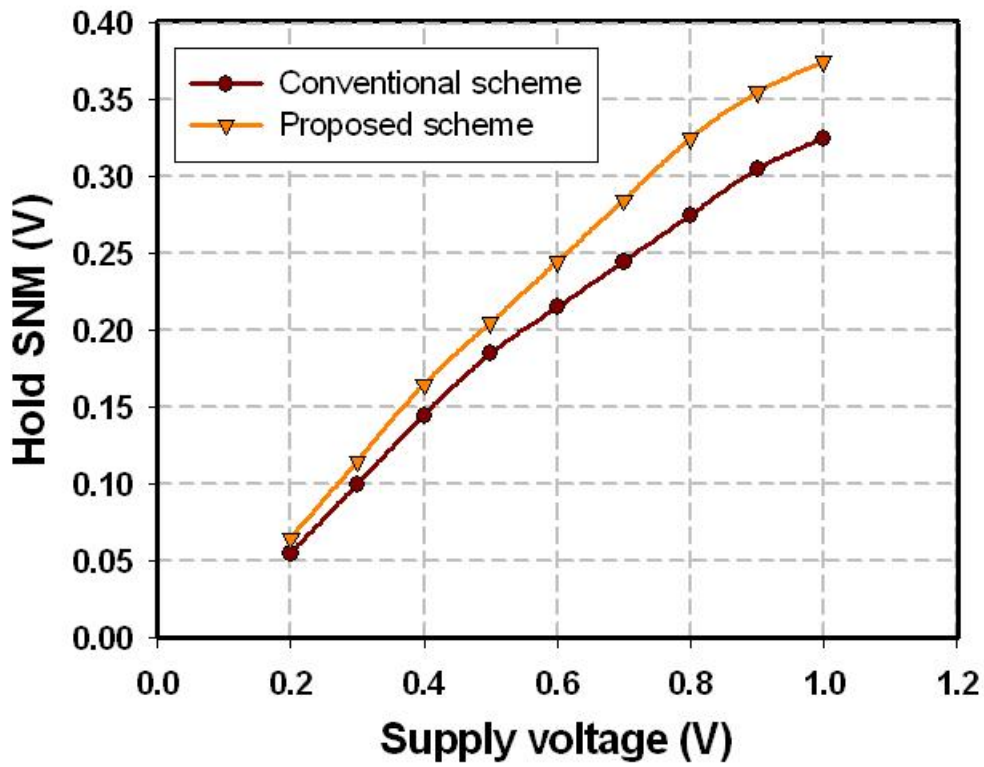


Figure 4.34: Hold SNM comparisons under different supply voltage.

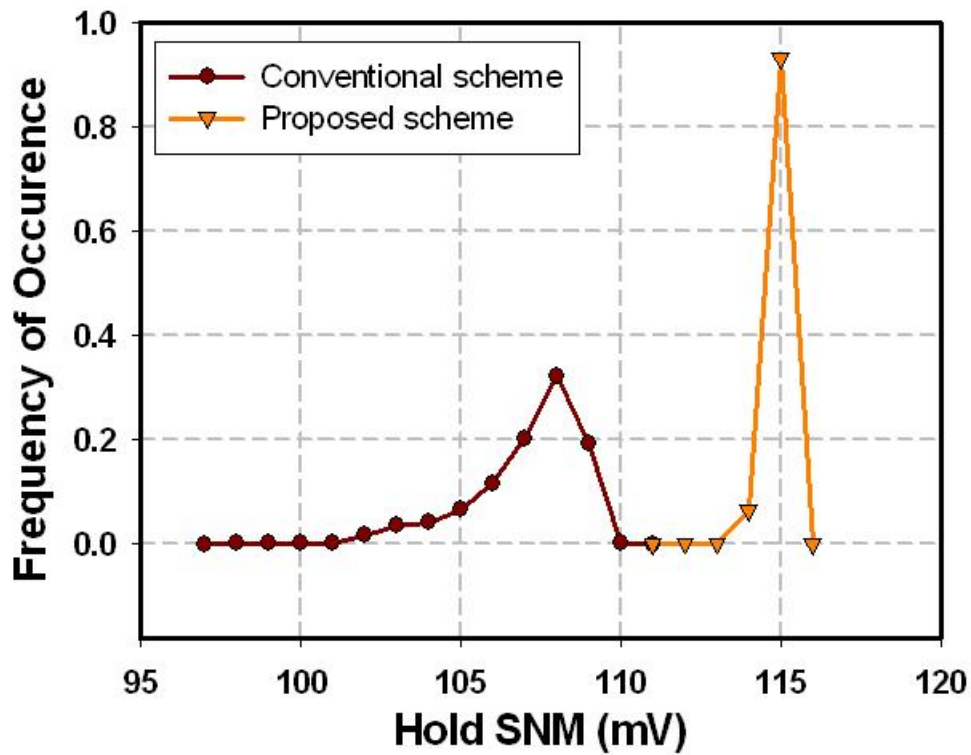


Figure 4.35: Distribution of hold SNM at 300mV.

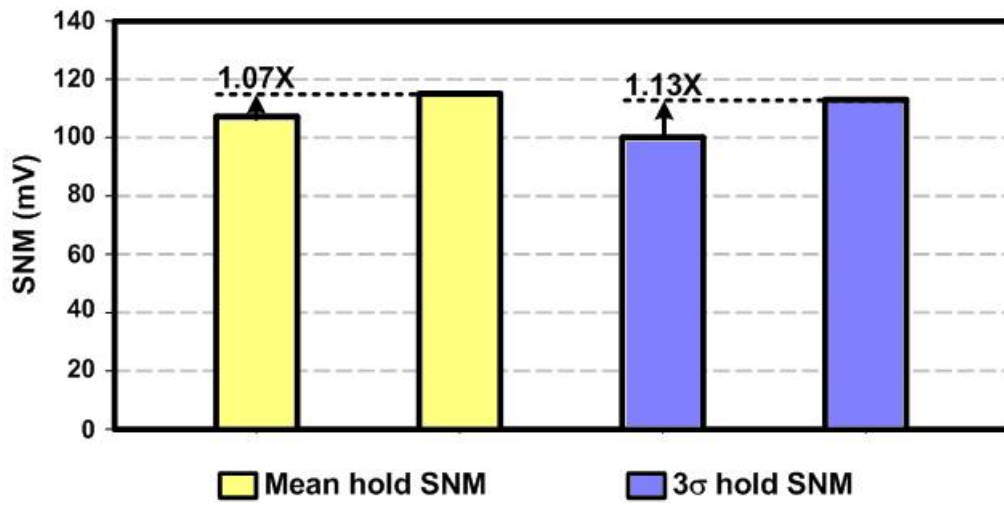


Figure 4.36: Hold SNM comparisons at 300mV.

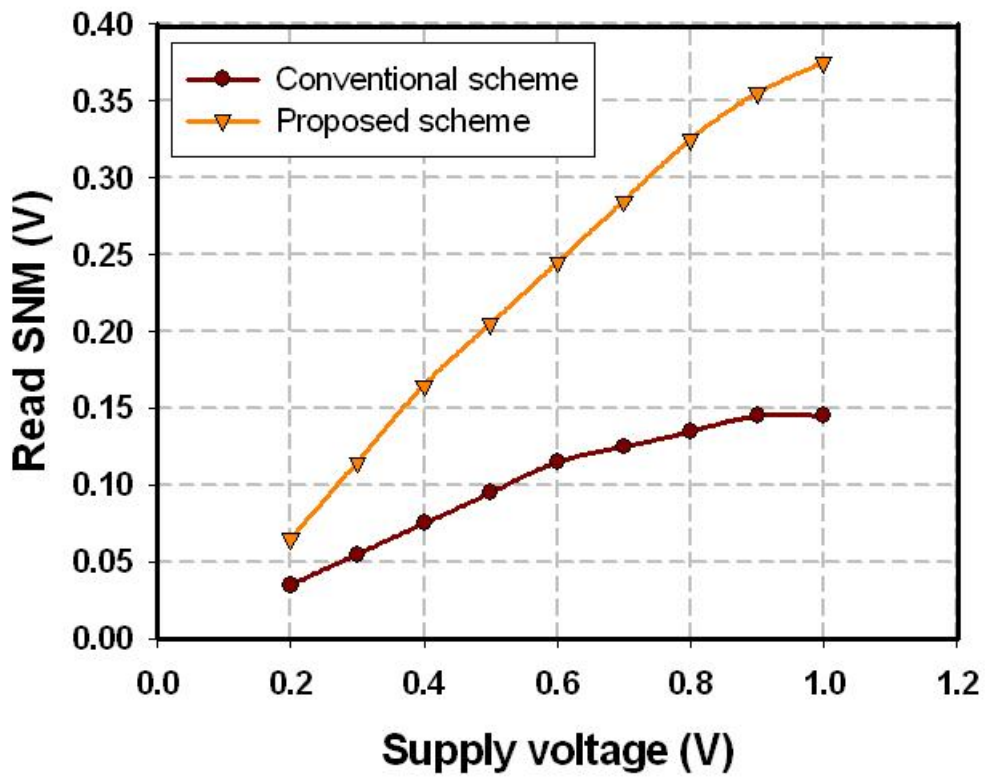


Figure 4.37: Read SNM comparisons under different supply voltage.

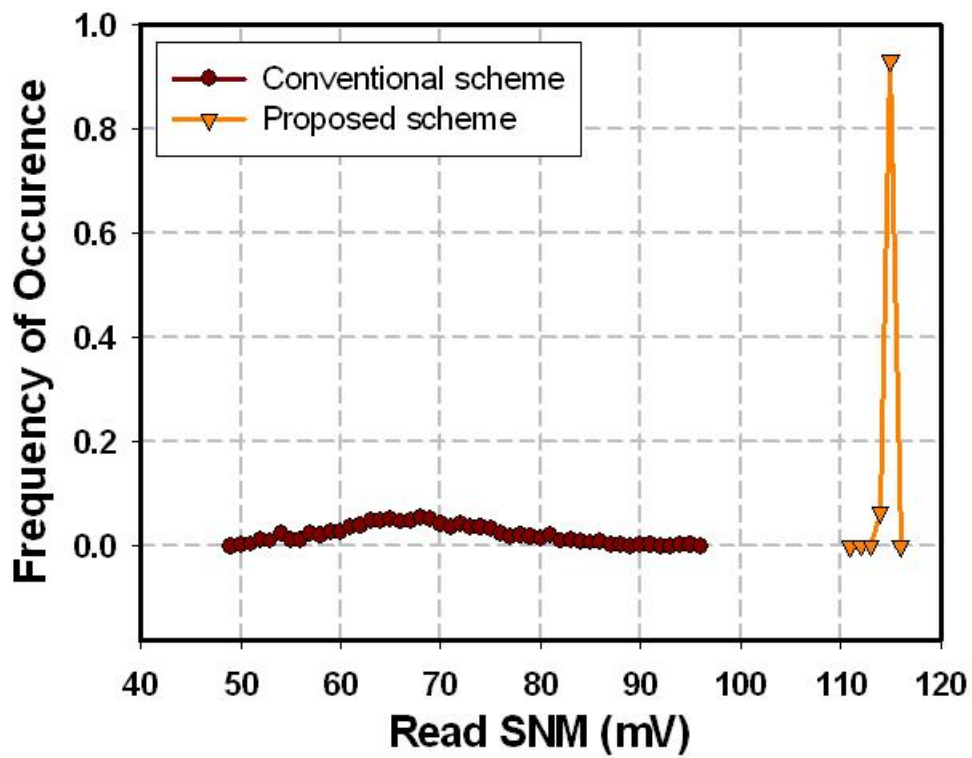


Figure 4.38: Distribution of read SNM at 300mV.

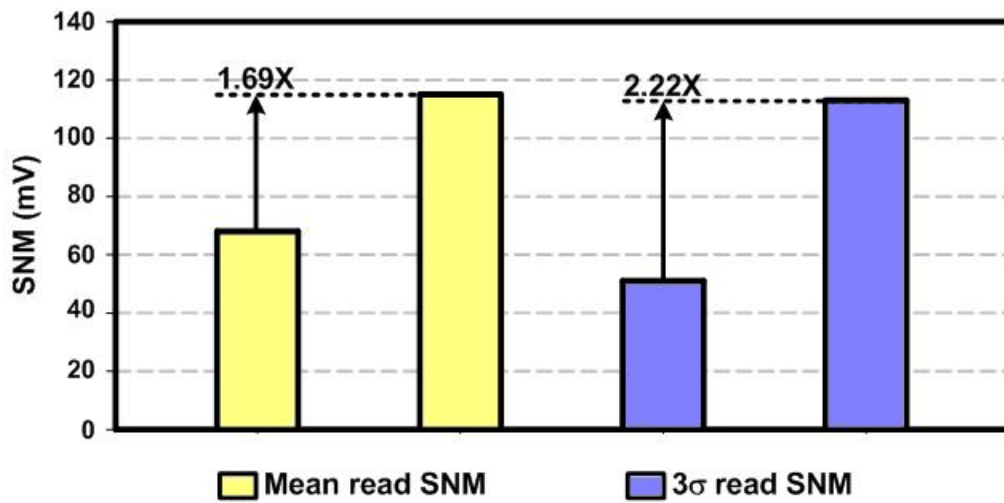


Figure 4.39: Read SNM comparisons at 300mV.

margin comparisons versus different supply voltage. Figure 4.41 shows the Monte Carlo simulations of write margin comparisons under 300mV supply voltage. Mean write margin and 3σ write margin are further shown in Figure 4.42. It is shown that the proposed scheme exhibits better write ability, with 1.25X mean write margin improvement and 1.13X 3σ write margin improvement, revealing the effectiveness of dual- V_T implementation.

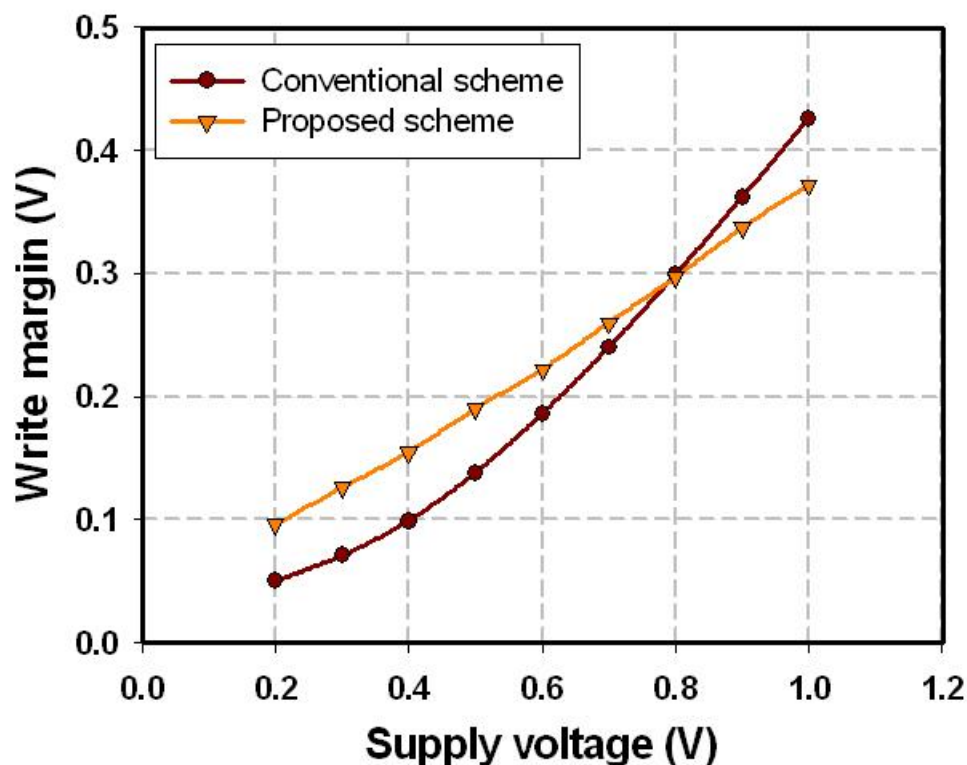


Figure 4.40: Write margin comparisons under different supply voltage.

Leakage Power

Due to high- V_T transistors, the proposed scheme consumes less leakage power, as shown in Figure 4.43. Note that this reduced bitline structure is able to achieve more power reduction in an SRAM array, making it suitable for ultra-low power design.

4.6 Summary

As SRAM continues to dominate the total area and power in modern SoC, subthreshold SRAM provides an effective strategy for total power saving. Stability issue, being the key concern in subthreshold designs, must be considered seriously. In this chapter, standard stability metrics and existing subthreshold SRAM designs are introduced. Moreover, a robust, fully differential, single-port subthreshold SRAM cell with auto-compensation is proposed. With the auto-compensation mechanism, the proposed cell is able to achieve better hold stability. Meanwhile, the cell structure eliminates conventional read SNM

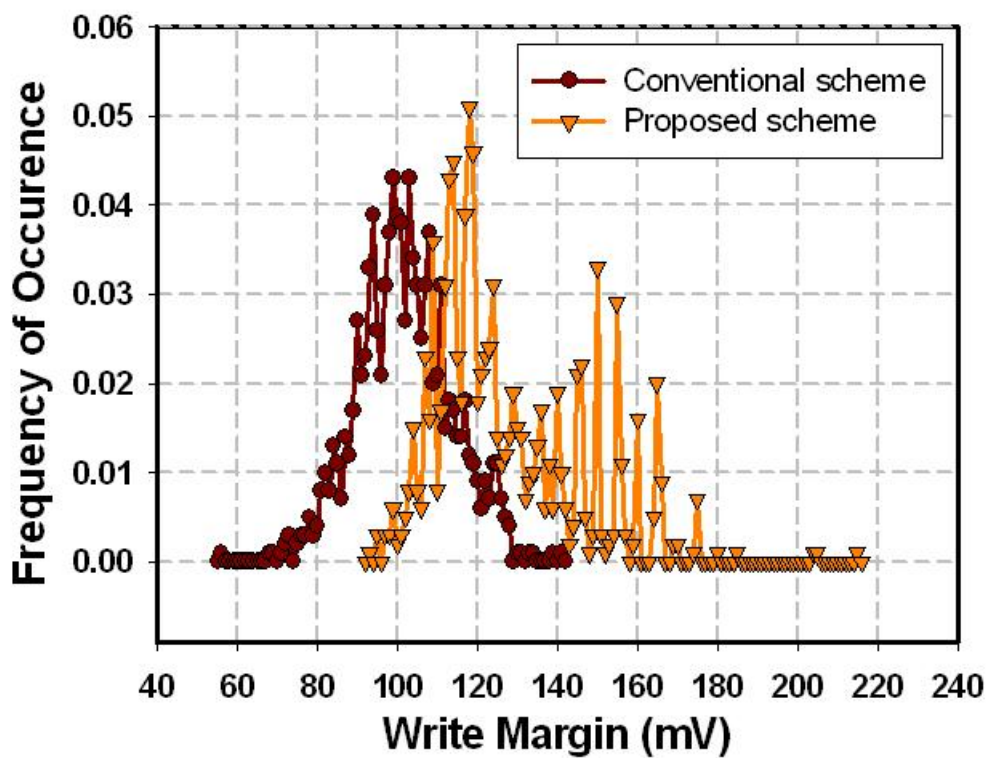


Figure 4.41: Distribution of write margin at 300mV.

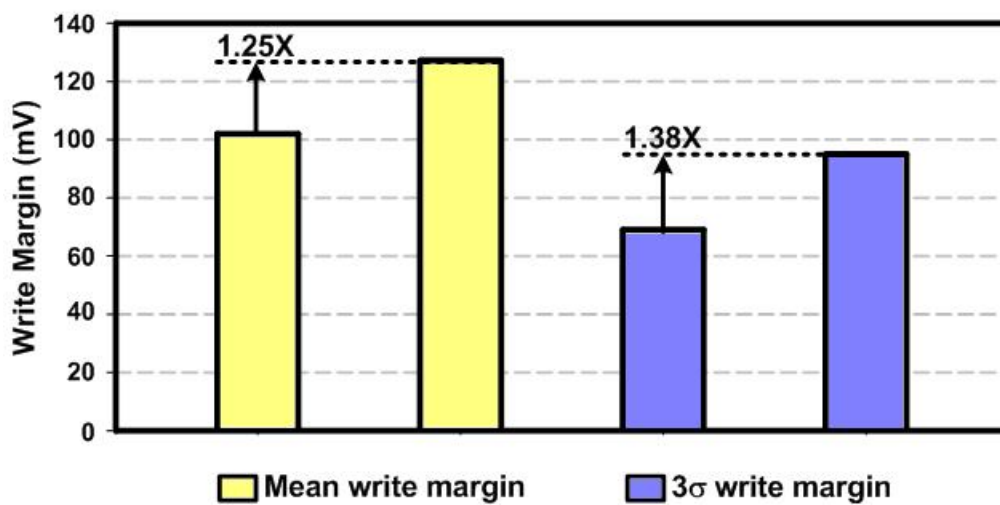


Figure 4.42: Write margin comparisons at 200mV.

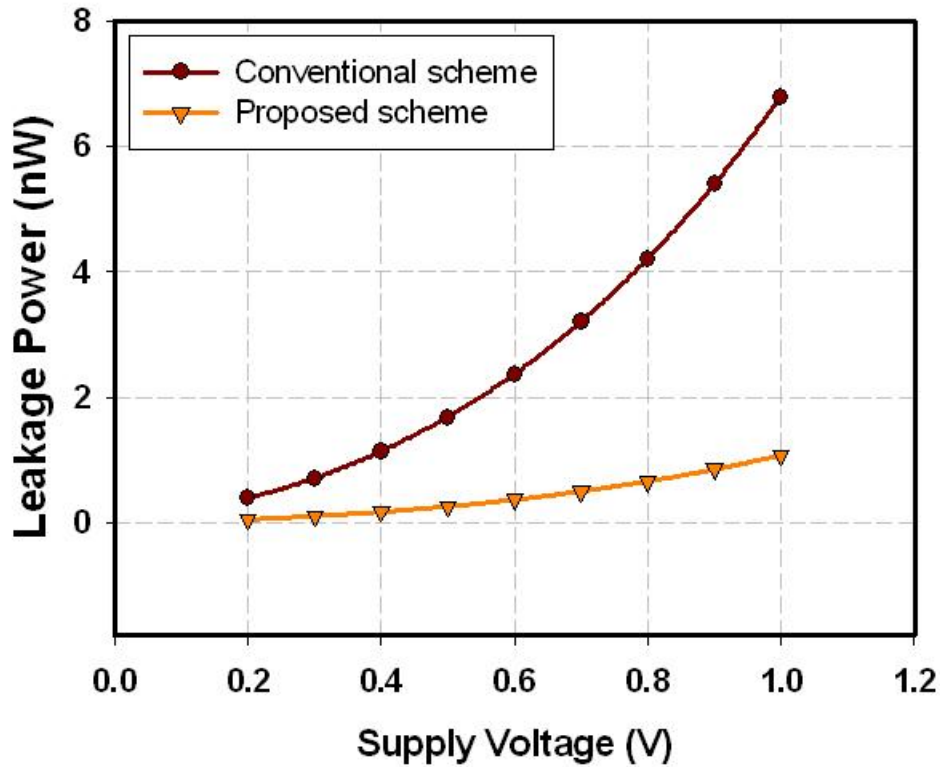


Figure 4.43: Cell leakage power comparison.

limitation by isolating storage nodes during read operation. Write ability is also improved by applying write assist technique. Compared to existing fully differential SRAM cell structures, simulation results show that at 200mV supply voltage, the proposed scheme achieves better hold stability, better read stability, and better write ability under process variation. On the other hand, a dual-port dual- V_T 7T subthreshold SRAM cell is proposed to enable simultaneous read/write operation. The proposed scheme exhibits better hold, read stability and write ability due to dual threshold transistor structure. Further, the dual- V_T 7T cell, along with the reduced bitline scheme, has significant power reduction compared to conventional design. As a result, the proposed schemes promise stable operation in subthreshold regime, enabling reliable and ultra-low power operation.

Chapter 5

A Robust Ultra-Low Power Asynchronous FIFO Memory

5.1 Introduction

First-in first-out (FIFO) memory is a key component of many SoC applications, which is commonly used for data buffering and flow control. An example is the emerging wireless body area network (WBAN), a breakthrough personal healthcare technology for body condition monitoring and diagnosis. Due to limited energy source and long-term stability requirement, robust ultra-low power designs are indispensable for a WBAN system [46]. As shown in Figure 5.1 [47], a major component of the system wireless sensor node (WSN) is a FIFO memory, which dominates the total die area and power. Therefore, reducing power consumption of the FIFO memory is an urgent design consideration for optimal WBAN.

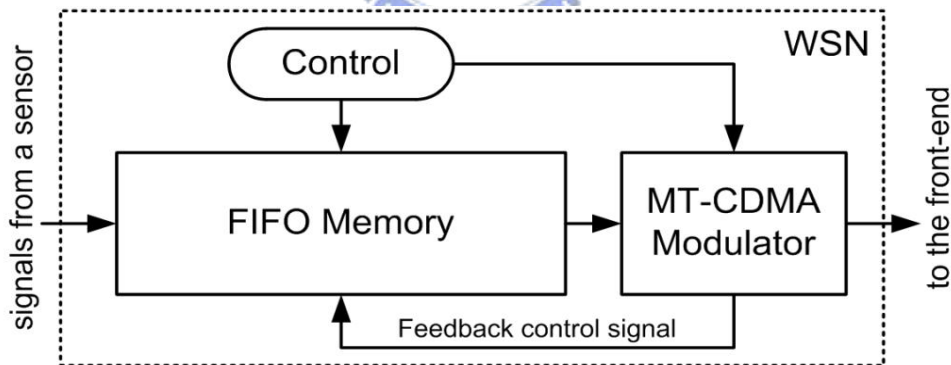


Figure 5.1: Block diagram of the wireless body network (WBAN) system wireless sensor node (WSN).

Due to loose timing constraint of the WSN, ultra-low supply voltage is suggested to be an effective method to gain ultra-low power operation. However, in nanometer CMOS technologies, where leakage power contributes a great portion of the total power consumption, the effectiveness of cutting off leakage power by supply voltage scaling is limited. Therefore, further leakage power reduction techniques must be applied to minimize power consumption. Moreover, as supply voltage scales down, CMOS circuit becomes sensitive to noise. Stability issue is especially important for storage elements that operate under ultra-low voltage. Therefore, when designing FIFO memory under

ultra-low supply voltage, stability improvement techniques must be applied to ensure functionality. In this chapter, three techniques are proposed to gain a robust ultra-low power FIFO memory, including the self-adaptive power control and the complementary power gating for FIFO memory array leakage power minimization, and the dual- V_T 7T SRAM cell (Section 4.5.3) for bitline overhead reduction and data stability improvement.

This chapter is organized as follows. A robust ultra-low power FIFO memory is proposed in Section 5.2. WBAN, the target application of the proposed FIFO, is described in Section 5.3. Design implementation and simulation results are shown in Section 5.4 and Section 5.5. Finally, Section 5.6 sums up this chapter.

5.2 Proposed Ultra-Low Power FIFO Memory

A FIFO primarily consists of a set of read and write pointers, storage and control logic. Storage may be SRAM, flip-flops, latches or any other suitable form of storage. A dual-port SRAM is usually used where one port is used for writing and the other is used for reading [48]. A synchronous FIFO is a FIFO where the same clock is used for both reading and writing. An asynchronous FIFO uses different clocks for reading and writing.

The overall proposed FIFO architecture is shown in Figure 5.2. It is composed of SRAM cell array, read circuit and read control, write circuit and write control, logic pointers (read/write pointers), and adaptive power control. The equivalent FIFO symbol of size 256-word by 16-bit is shown in Figure 5.3. The signal description is shown in Table 5.1, and the command truth table is further shown in Table 5.2.

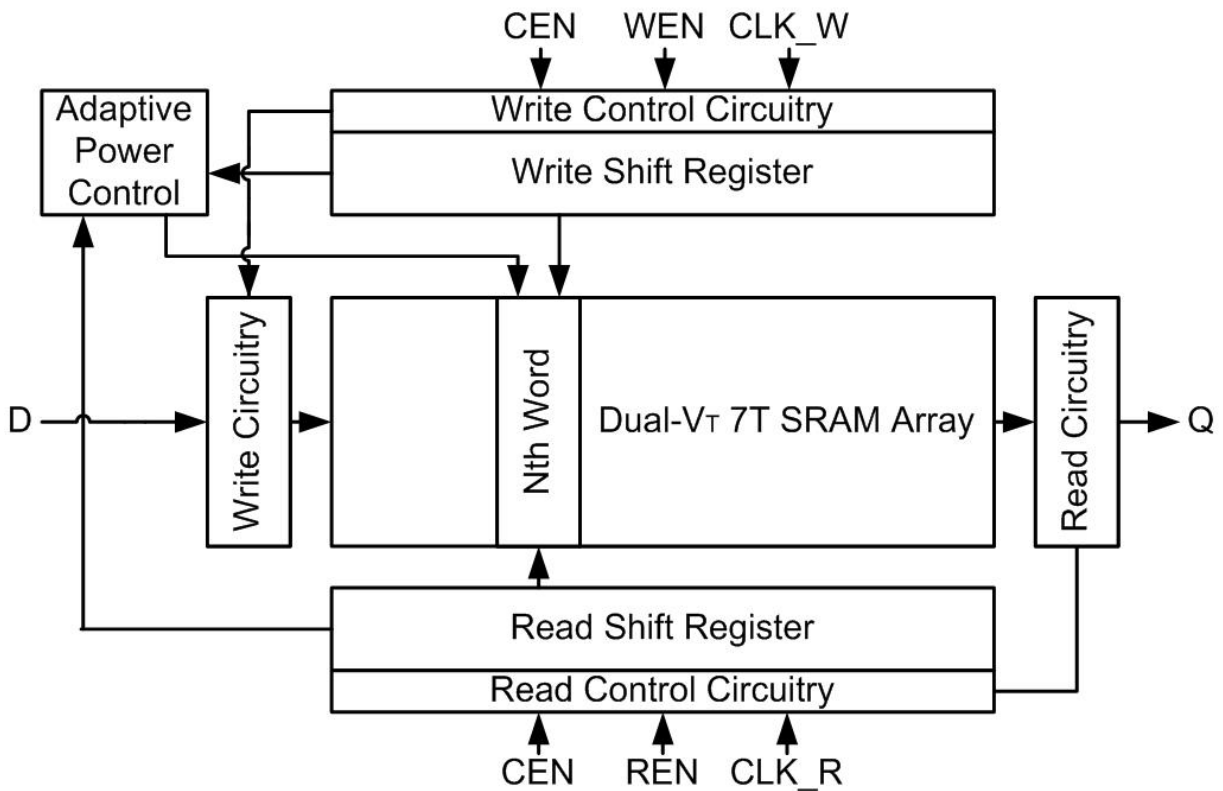


Figure 5.2: Block diagram of the proposed FIFO memory.

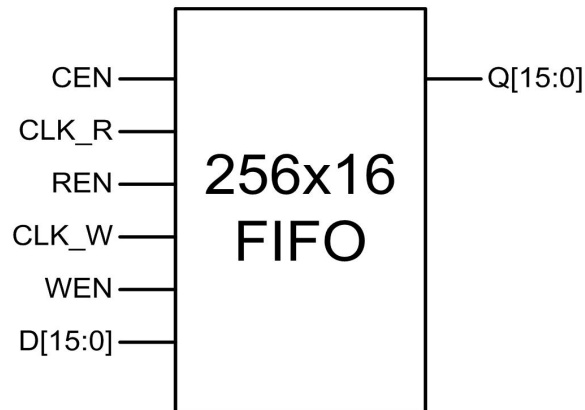


Figure 5.3: 256x16 FIFO symbol.

Table 5.1: Signal descriptions.

Input	CEN	Chip enable
	CLK_R	Read clock
	REN	Read enable
	CLK_W	Write clock
	WEN	Write enable
	D[15:0]	16-bit input data
Output	Q[15:0]	16-bit output data

Table 5.2: Command truth table.

CLK_R	CLK_W	CEN	REN	WEN	Operation
X	X	High	X	X	Chip disabled
X	X	Low	High	High	Hold
Positive edge	X	Low	Low	High	Read
X	Positive edge	Low	High	Low	Write
Positive edge	Positive edge	Low	Low	Low	Simultaneous read/write

5.2.1 Logic Pointer

Logic pointers (read pointer and write pointer) are used as the address pointers of the FIFO memory. An effective way to construct a logic pointer is the utilization of shift registers. The use of shift registers eliminates the use of counters and decoders, which reduces power consumption and increase operation speed significantly. Figure 5.4 shows a common logic pointer representation. When the FIFO is activated, the address will form a shift-register-like function to select the proper wordline to read or write data. Only one register will output data-1 to indicate the selected wordline, while the rest of the registers will output data-0, indicating the corresponding wordlines are not selected.

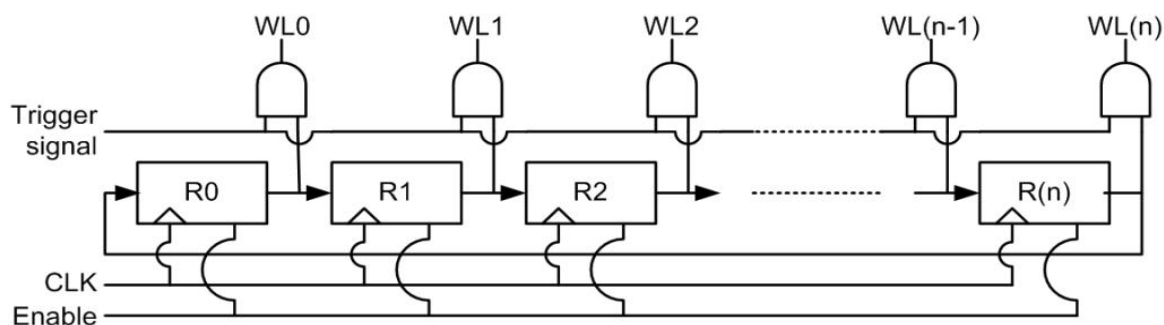


Figure 5.4: Logic pointer composed by shift registers.

Ultra-Low Voltage Flip-Flops

Logic pointers account for a relatively large portion of the total power consumption. The major part of the logic pointer is flip-flop, therefore, ultra-low power flip-flops is an important design consideration. Figure 5.5–Figure 5.8 shows four widely used flip-flop designs: PowerPC master-slaver latch [49], modified C2MOS (mC2MOS) master-slaver latch [50], hybrid-latch flip flop (HLFF) [51] and sense-amplifier-based flip-flop (SAFF) [52]. PowerPC and mC2MOS are composed of two identical cascaded latches, which are active at different phases of clock signal. HLFF belongs to the class of pulse triggered flip-flops. The input data is latched during a short pulse at the rising edge of the clock. SAFF with two coupled NAND gates as the output latch is a true single clock phase operation.

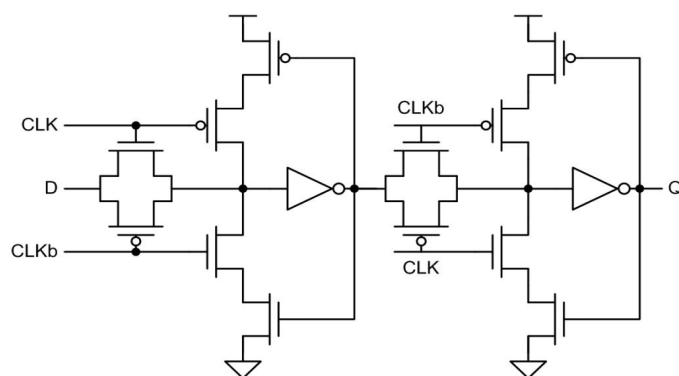


Figure 5.5: PowerPC master-slaver latch (PowerPC).

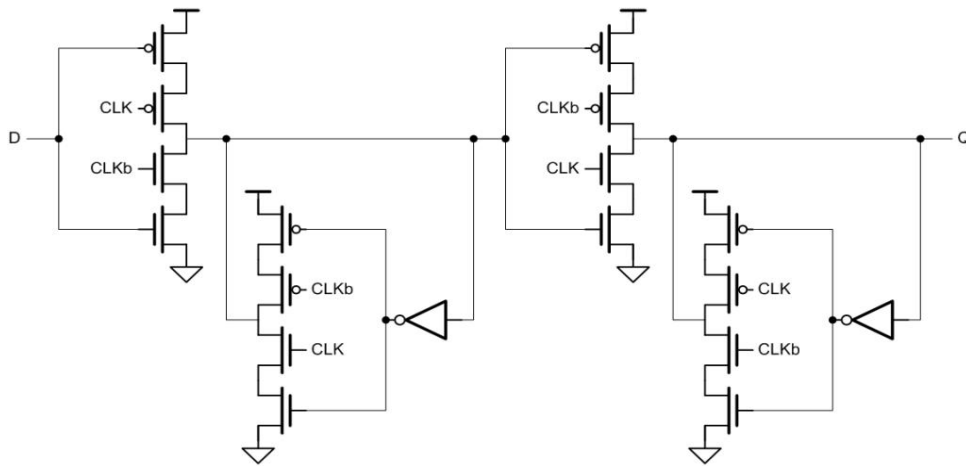


Figure 5.6: Modified C2MOS master-slaver latch (mC2MOS).

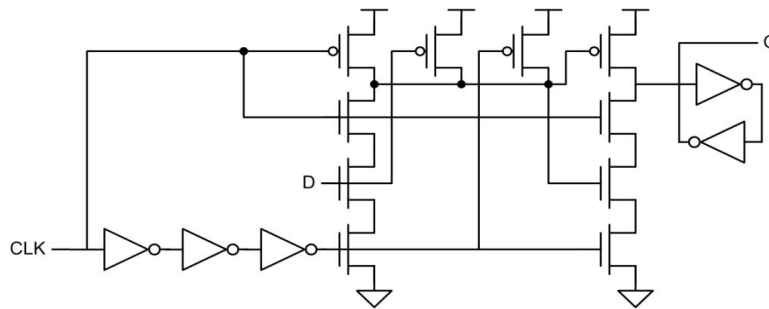


Figure 5.7: Hybrid-latch flip flop (HLFF).

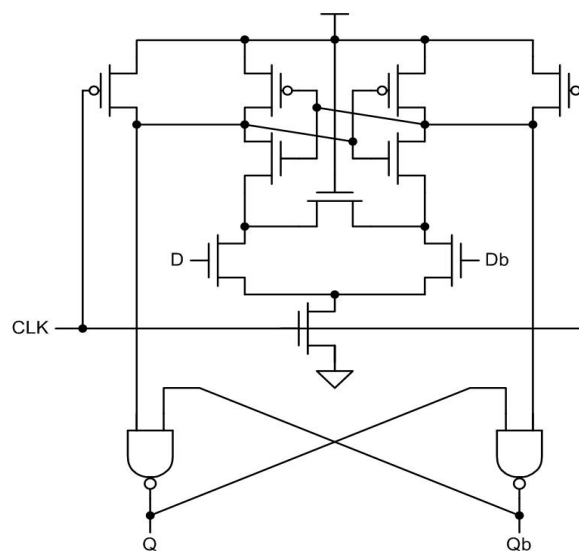


Figure 5.8: Sense-amplifier-based flip-flop (SAFF).

Figure 5.9 [53] compares the T_{setup} , $T_{C \rightarrow Q}$ and $T_{D \rightarrow Q}$ of various flip-flops as the supply voltage is scaled. It is shown that delay time increases as the supply voltage decreases, where delay time increases significantly when supply voltage is in the sub-threshold regime. It is also shown that HLFF achieves the least delay time. This fact is apparent when supply voltage is in the subthreshold regime. Figure 5.10 [53] shows

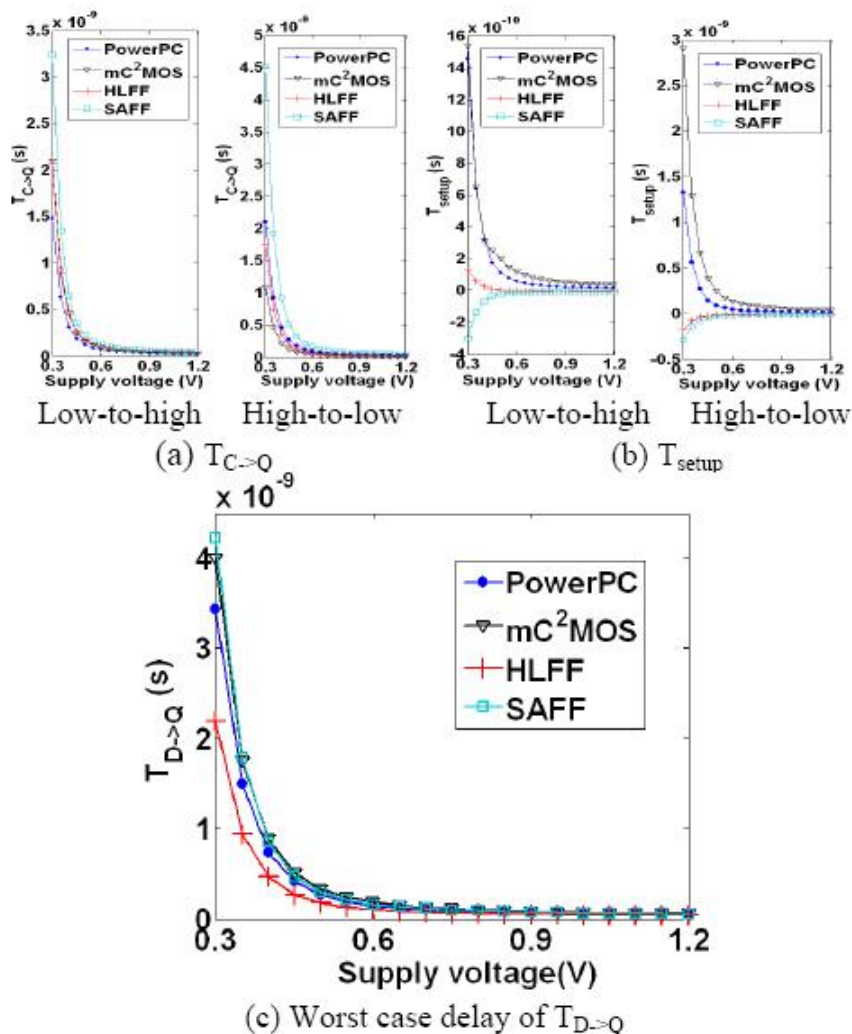


Figure 5.9: Timing parameters of the flip-flop as a function of the supply voltage.

the energy dissipation of selected flip-flop designs, where PowerPC consumes the least energy. Simulations of energy-delay-product (EDP), an examination vector of the balance between speed and energy consumption, is shown in Figure 5.11 [53]. The PowerPC flip-flop achieves the smallest EDP at high voltages for all switching activities due to its minimal energy consumption and relatively small delay. As supply voltage decreases, the most energy efficient flip-flop architecture depends on switching probabilities, where PowerPC achieves better EDP at low activities, and HLFF achieves better EDP at high activities. Therefore, from the analysis shown above, due to ultra-low voltage operation and low activity of the logic pointer, the PowerPC flip-flop is chosen to be the basic element.

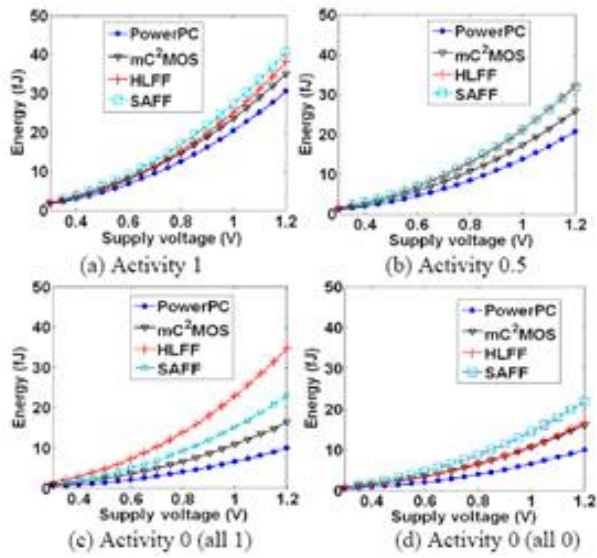


Figure 5.10: Energy dissipation as a function of the supply voltage for different switching activities.

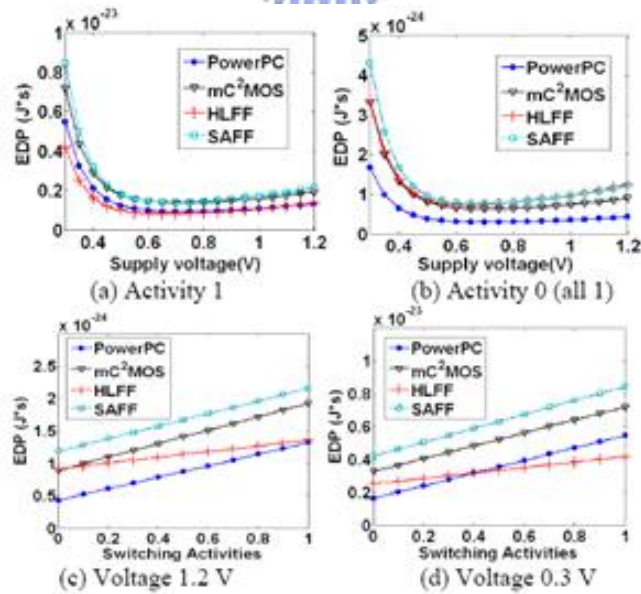


Figure 5.11: EDP as a function of supply voltage and switching activities.

Read Pointer

The read pointer is shown in Figure 5.12. Each block in the figure represents a register (flip-flop). READ and RD signal are generated by the read control circuit, which will be described afterwards. CLK_R represents the clock for read operation. When CLK_R is at the positive edge and READ=1, the shift register will shift its data. In the 256 bit shift register, only one register stores data-1, along with the AND gates for signal RD to control timing properly, the selected read wordline will be turned on at the desired time.

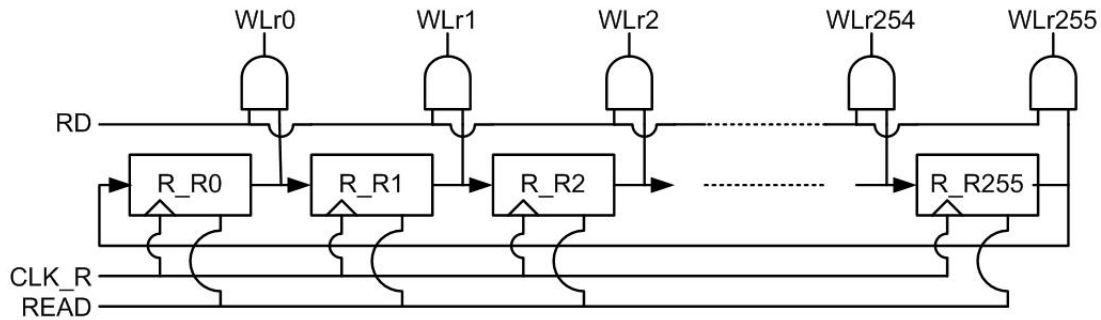


Figure 5.12: Read pointer.

Write Pointer

The write pointer is shown in Figure 5.13. It functions similar to the read pointer. WRITE and WR signal are generated by the write control circuit, which will be described afterwards. CLK_W represents the clock for write operation. When CLK_W is at the positive edge and WRITE=1, the shift register will shift its data. In the 256 bit shift register, only one register stores data-1, along with the AND gates for signal WR to control timing properly, the selected write wordline will be turned on at the desired time.

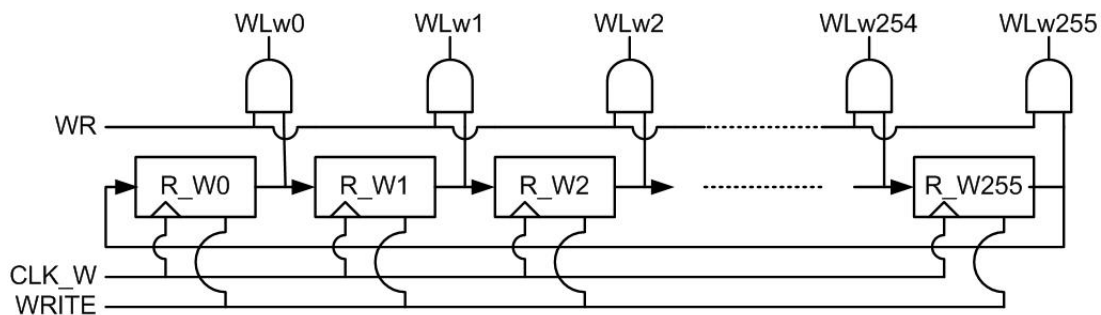


Figure 5.13: Write pointer.

5.2.2 Read Operation

Read Control Circuit

For read control signal generation, at the positive CLK_R edge, if CEN=0 and REN=0, DFF (D flip-flop) will generate a READ signal with pulse width equal to one CLK_R

cycle. RD and R2 signal are generated by logic combination of CLK_R and READ, which is a pulse with pulse width equal to half CLK_R cycle. RD is high at the first half of READ, and R2 is high at the second half of READ. READ, RD, R2 can properly control FIFO read behavior and the adaptive power control circuitry. The read control circuit is shown in Figure 5.14.

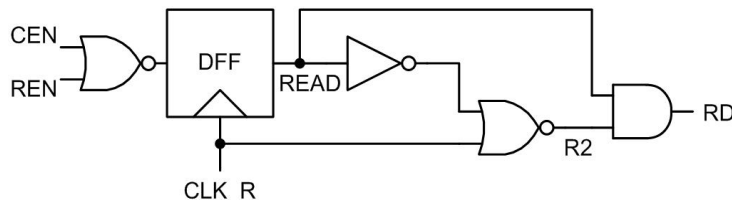


Figure 5.14: Read control circuit.

Precharge Circuit and Sense Amplifier

For the 7T cell (Section 4.5.3) to perform read, precharge circuit and sense amplifier are needed. Figure 5.15 shows a FIFO column. Read bitline is precharged to V_{DD} . During read access, RD turns on the selected read wordline, voltage swing on the read bitlines will be generated according to the word content. Sense amplifiers at the read output end will amplify the signals on the read bitlines, and hold the output data by a latch. The designed sense amplifier utilizes dual- V_T transistors to reduce power consumption and reduce read delay at the same time.

Note that the read failure probability issue discussed in Chapter 4 will be solved by applying the complementary power gating technique, which will be presented afterwards. Number of Cells per read bitline is also limited (256 cells per bitline) to control the read failure rate.

5.2.3 Write Operation

Write Control Circuit

For write control signal generation, at the positive CLK_W edge, if CEN=0 and WEN=0, DFF will generate a WRITE signal with pulse width equal to one CLK_W cycle. WR signal is generated by logic combination of CLK_W and WRITE, which is a pulse with pulse width equal to half CLK_W cycle, and is high at the second half of WRITE. WRITE and WR can properly control FIFO write behavior and the adaptive power control circuitry. The write control circuit is shown in Figure 5.16.

Write Driver

Figure 5.15 shows a FIFO column with single-ended write. The write bitline does not need to be precharged for successful write. However, for non-precharged write bitline, the driving ability of the write driver must be strong enough to drive both data-0 and data-1. Therefore, buffers with proper size are placed. During write access, WR turns on the selected write wordline, and the write driver inputs data to the selected word, enabling successful write.

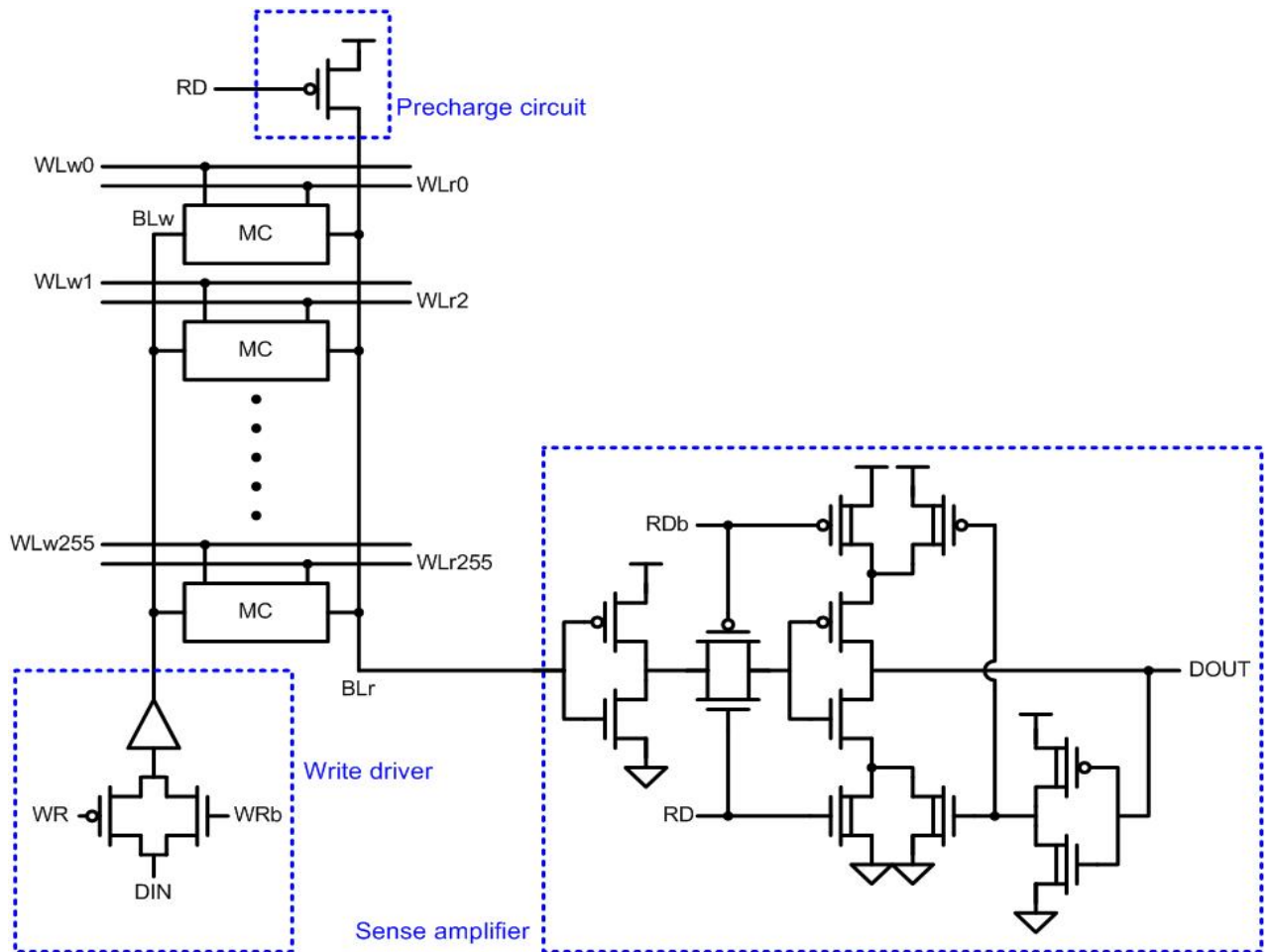


Figure 5.15: A FIFO memory Column.

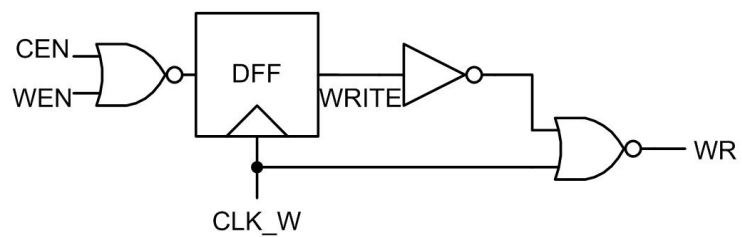


Figure 5.16: Write control circuit.

5.2.4 Self-Adaptive Power Control

The key idea of leakage power minimization is to reduce voltage swing on un-functioning hardware. To demonstrate this idea, Figure 5.17 is a FIFO operation example. Grey blocks represent words that contain data, while white blocks represent words that are empty. Empty words does not need data retention ability, thus, the cross voltage of the word can be reduced to zero for leakage power minimization. Moreover, due to first-in first-out data behavior, status for each word is completely predicable, i.e. a word changes state only when read or write occurs, where read/write pointer follows circular shifting characteristic, thus allowing self-adaptive power control with acceptable power overhead.

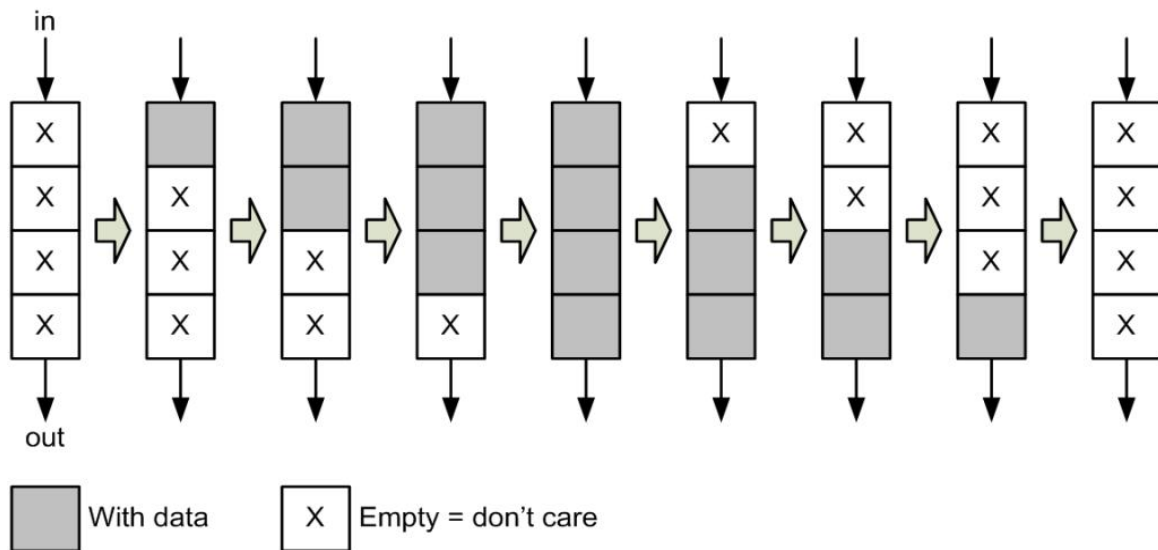


Figure 5.17: FIFO operation example.

The basic function of the adaptive control signal generation for each word is described as follows

```

if(CEN) CTRL_CELL=1; //chip disable
else if(write) CTRL_CELL=0; //write occurs
else if(read) CTRL_CELL=1; //read occurs
else CTRL_CELL=CTRL_CELL; //else

```

where CTRL_CELL=1 means that the cross voltage of the word is zero, representing a turned off state, and CTRL_CELL=0 means that the word has normal voltage supply for data retention.

The corresponding schematic of the adaptive power control circuit is shown in Figure 5.18. As shown, CTRL_CELL and CTRL_READ are generated by read pointer, WLr, WLw, CEN, and R2. The read pointer signal is the output signal of each flip-flop (Figure 5.12); WLr is the read wordline signal (Figure 5.13); WLw is the write wordline signal (Figure 5.14); CEN is the chip enable signal; R2 is generated by the read control circuit (Figure 5.14). CTRL_READ=1 means that the foot of the 7T cell's read buffer is high to minimize the cross voltage of the read buffer and reduce read bitline leakage, and CTRL_READ=0 means that read access is functioning.

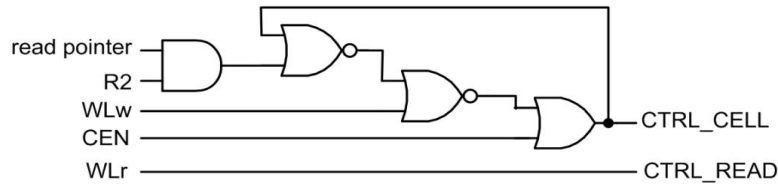


Figure 5.18: Adaptive power control circuit.

Figure 5.19 and Figure 5.20 is an example to illustrate the control scheme. "_0" represents signals related to the first memory word, and "_255" represents signals related to the 256th (the last) memory word. In this example, the write frequency is 200kHz, while the read frequency is 5MHz, a read clock rate much faster than the write clock rate. In Figure 5.19, *ctrl_0* (CTRL_CELL_0) is initially high due to the chip disable signal, and the corresponding word is therefore turned off. When write occurs, that is, *wl_w_0* (WLw0) is on, the *ctrl_0* (CTRL_CELL_0) is triggered and drops to low, and the corresponding word is therefore turned on for data retention. *ctrl_0* (CTRL_CELL_0) stays at low to hold the data until read occurs. After read operation is complete, the signal *r2_0* is pulsed to high to indicate the corresponding word contains don't care data, and *ctrl_0* (CTRL_CELL_0) is high again to turn off the word. Figure 5.20 shows the 256th word operation, which is similar to Figure 5.19 just described. All the words not shown also share the same operating characteristic.

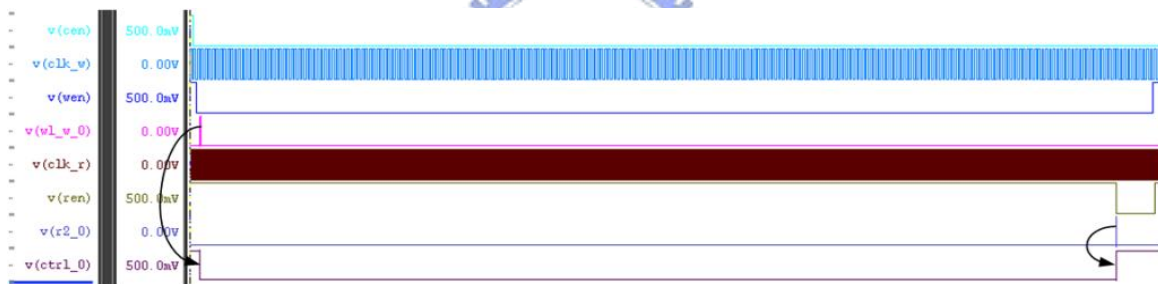


Figure 5.19: Waveform of the adaptive power control related signal (1st word).

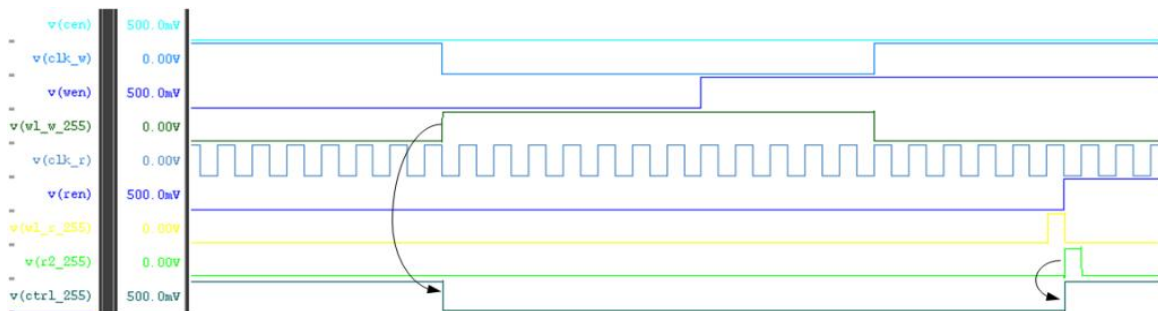


Figure 5.20: Waveform of the adaptive power control related signal (256th word).

5.2.5 Complementary Power Gating

To support the self-adaptive power control, complementary power gating is inserted into each FIFO memory word, as shown in Figure 5.21. If the word has data storage,

CTRL_CELL is off to turn on the PMOS transistor so that V_VDD has a direct path to VDD for data preservation. On the other hand, if the word is empty, CTRL_CELL is on to turn on the NMOS transistor, so that V_VDD drops to GND, which minimizes the voltage swing across SRAM cells. Complementary power gating is also applied on V_GND, which is the ground of the read buffer of each SRAM cell. When the word is not in read operation, CTRL_READ is off to turn on the PMOS transistor, so that V_GND is boosted to VDD, which minimizes voltage swing between read bitlines and V_GND. If the word is functioning read, CTRL_READ is on to turn on the NMOS transistor so that a read path can be created. Table 5.3 summarizes relations between each FIFO memory word status and its corresponding control signals. The generations of CTRL_CELL and CTRL_READ signals are described in the self-adaptive power control section. Complementary power gating utilizes high- V_T devices and ensures no floating node occurs, which promise robust and low leakage operation.

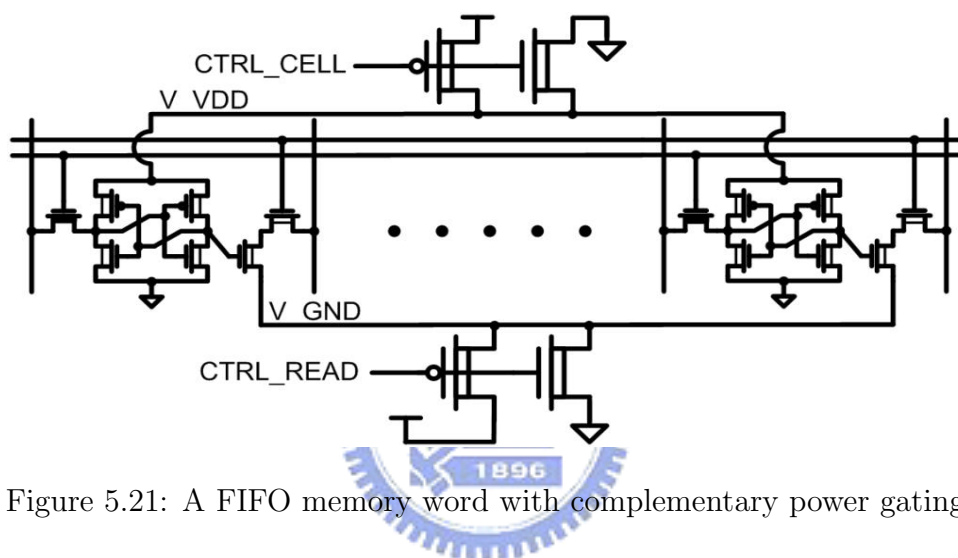


Figure 5.21: A FIFO memory word with complementary power gating.

Table 5.3: Summary of FIFO word states and corresponding control signals.

		CTRL_CELL/CTRL_READ	V_VDD/V_GND
Write	1st half cycle	High/Low	GND/VDD
	2nd half cycle	Low/Low	VDD/VDD
Read	1st half cycle	Low/High	VDD/GND
	2nd half cycle	High/Low	GND/VDD
With data storage		Low/Low	VDD/VDD
Without data storage		High/Low	GND/VDD

5.2.6 Storage Element

The basic storage element of the ultra-low power FIFO memory utilizes the dual- V_T 7T SRAM cell proposed in Chapter 4. As stated previously, the 7T SRAM cell has the following characteristics which are suitable for robust, ultra-low power FIFO operation.

- **Reduced bitline:** A major power consumption source of a memory array is caused by the voltage swing on bitlines. This is because bitlines usually contribute large loading capacitance [54]. Therefore, reducing bitline loading results in significant active power reduction.

- Dual- V_T structure: In the 7T cell, the dual- V_T structure improves hold stability and write ability, enabling more reliable operation under ultra-low supply voltage. Moreover, only one transistor is a low- V_T device, the other six transistors are high- V_T devices. High- V_T device results in smaller leakage current, thereby reducing leakage power.
- Read buffer: With the read buffer structure, the 7T cell isolates storage nodes from being directly interfered by the read bitline noise, allowing better read stability for more robust operation under ultra-low supply voltage.

5.3 Case Study: Ultra-Low Power Wireless Sensor Node for WBAN Application

Wireless body area network (WBAN) is one of the most suitable technology for building unobtrusive, scalable, and robust wearable health monitoring systems [55]. It consists of a set of mobile and compact intercommunicating sensors, either wearable or implanted into human body, which monitor vital body parameters and movements. These devices, communicating through wireless technologies, transmit data from the body to a home base station, from where the data can be forwarded to a hospital, clinic or elsewhere, real time. Figure 5.22 shows the generic concept of wireless body area network of intelligent sensors for patient monitoring [56].

To be unobtrusive, the sensors must be lightweight with small form factor. The size and weight of sensors is predominantly determined by the size and weight of batteries. Requirements for extended battery life directly oppose the requirement for small form factor and low weight. This implies that sensors have to be extremely power efficient, as frequent battery changes for multiple WBAN sensors would likely hamper users' acceptance and increase the cost. In addition, low power consumption is very important toward future generations of implantable sensors that would ideally be self-powered, using energy extracted from the environment. As a result, the ultra-low power wireless sensor node (WSN) is the most crucial design target to achieve.

Intelligent on-sensor signal processing has the potential to save power by transmitting the processed data rather than raw signals, and consequently to extend battery life. A careful trade-off between communication and computation is crucial for an optimal design. A dual mode transceiver supporting OFDMA and Multi-Tone CDMA modulation is proposed for different communication benefits. OFDMA is able to achieve high data rate while Multi-Tone CDMA has higher interference-immune ability. Thus, the modulation mode can be chosen as "Normal Mode" (Multi-Tone CDMA) or "Turbo Mode" (OFDMA). The block diagram of WSN baseband design is shown in Figure 5.23. The operation scenario composes of four modes. First, the "Down Link Mode", WSN receives commands from the central processing node for the upper layer protocol and the calibration information. The OFDMA DL receiver is busy in this phase. When the "Data Gathering Mode" is entered, the FIFO memory stores signals from a sensor until the amount of the data is nearly full. After that, the control unit asks the baseband modulator to do signal modulation either in "Normal Transmission Mode" or "Turbo Transmission Mode". After all the data are transmitted, the WSN returns to "Down Link Mode".

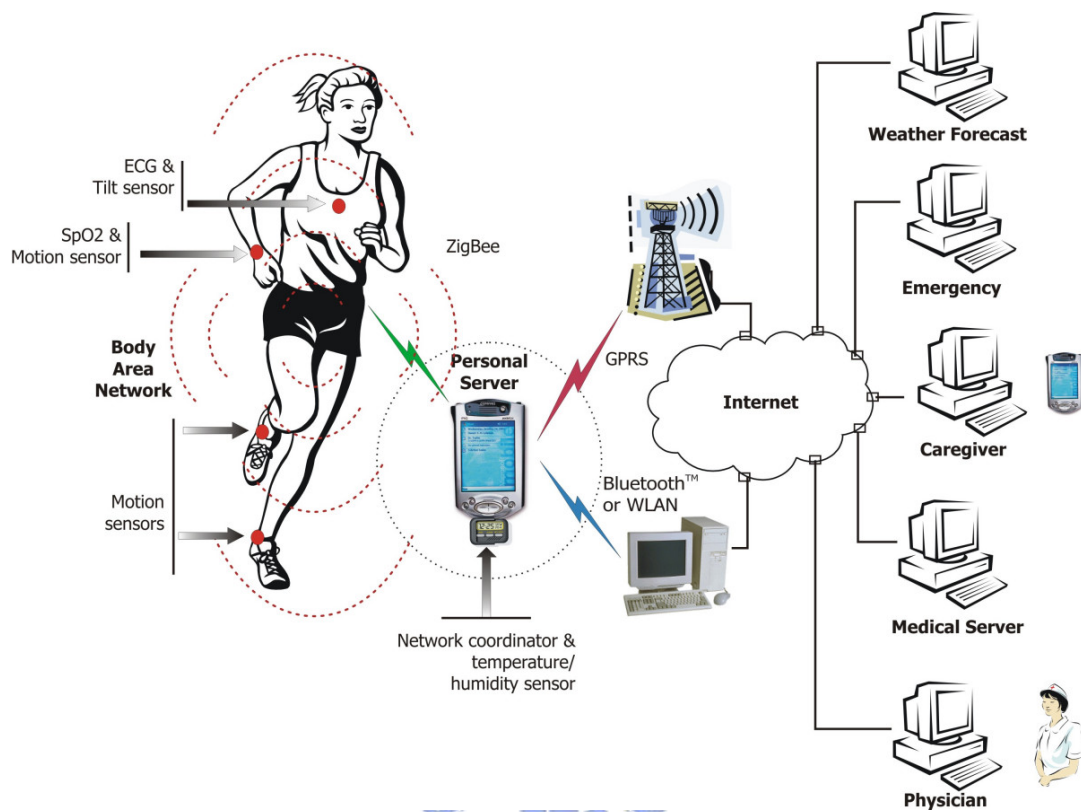


Figure 5.22: Wireless body area network of intelligent sensors for patient monitoring.

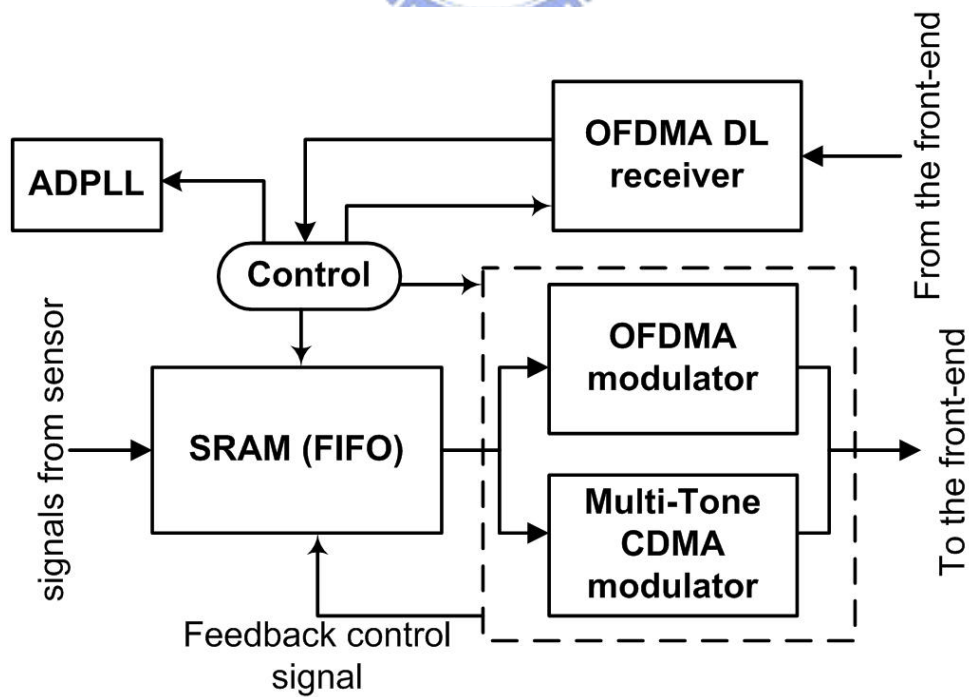


Figure 5.23: Block diagram of the proposed WSN.

5.4 Design Implementation

An optimize operation behavior is calculated for WSN power efficiency. The calculation result indicates a 512-word by 16-bit FIFO memory block with 200kHz write frequency and 5MHz read frequency best suits the system. Therefore, the proposed FIFO is designed to meet this speculation. Moreover, the proposed WSN provides three voltage levels: 1V, 0.8V, and 0.5V. Conventional SRAM block fails to operate below 0.7V [57] due to stability issue, while the proposed FIFO with stability improvement works well under ultra-low supply voltage. 0.5V is also sufficient for the proposed FIFO to meet the timing constraint; therefore 0.5V is set for the FIFO's supply voltage.

A 256-word by 16-bit robust ultra-low power asynchronous FIFO memory is implemented in UMC 90nm CMOS technology with 0.5V supply voltage. This sizing configuration is suitable for being a basic FIFO block. The target application is the new generation WSN for WBAN, where chip design is required to be highly integrated in a tiny area, and power consumption is limited to μW scale. The proposed design is fully functional within $\pm 10\%$ voltage variation, 0°C to 100°C temperature variation, and all process corners. The layout view of the FIFO block and the test chip are shown in Figure 5.24 and Figure 5.25, respectively. The design profile is summarized in Table 5.4.

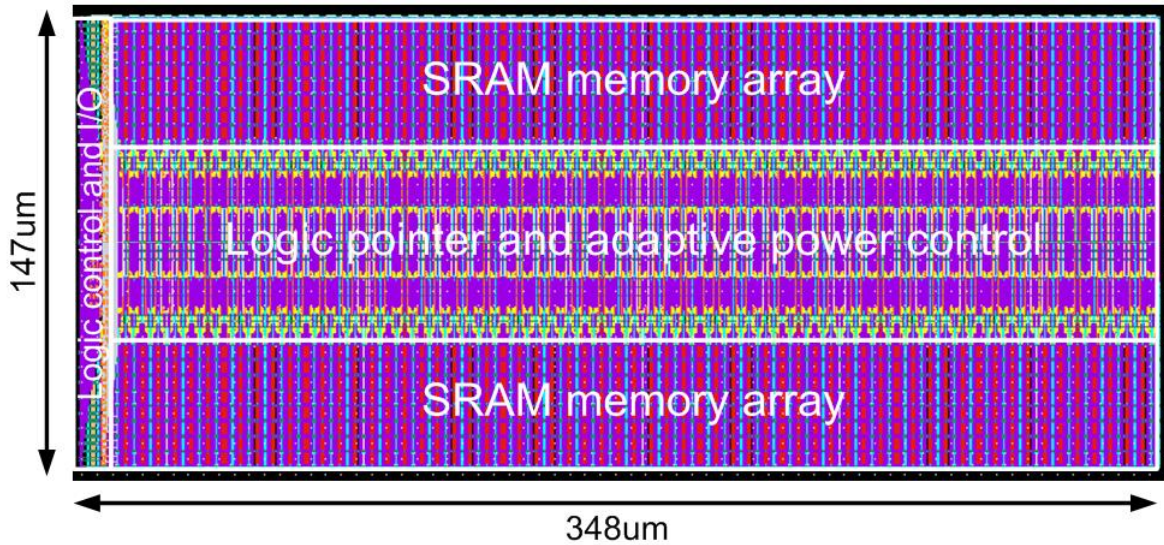


Figure 5.24: Layout view of the FIFO memory.

Table 5.4: Summary of the FIFO memory features.

Technology	UMC 90nm CMOS
FIFO memory Size	256-word by 16-bit (4kb)
Supply voltage	0.5V
Read frequency	5MHz
Write frequency	200kHz
Power consumption	2.21uW
Area	348um X 147um

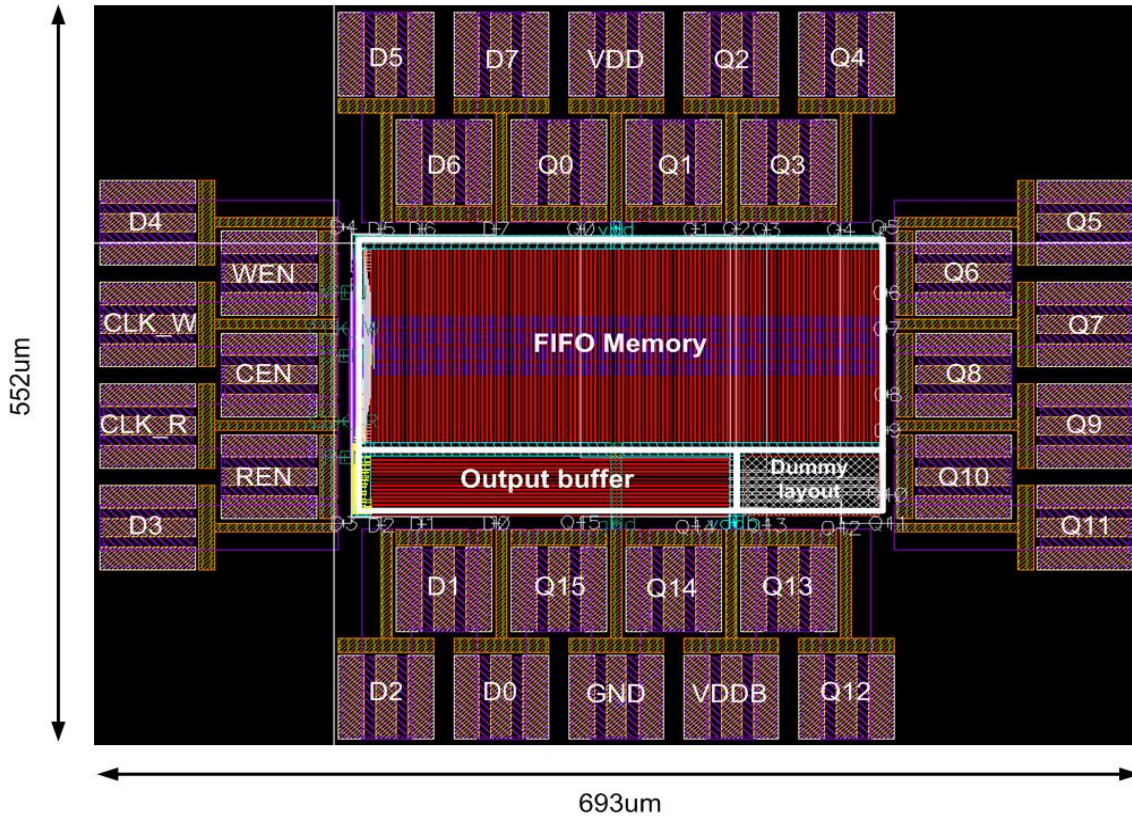


Figure 5.25: Layout view of the test chip.

5.5 Simulation Results

As shown in Table 5.4, memory size, supply voltage, read and write frequency are given by the system speculation. The output loading of the test chip is 0.5pF, which is the capacitance of the output pad. Therefore, output buffers are inserted to drive this large loading. Note that the simulated power consumption excludes the power of the output buffers.

Due to the system data behavior, power consumption of the FIFO block can be approximately described as

$$\begin{aligned}
 \text{Power} = & \quad 1\% \quad \text{standby power} \\
 & + \quad 93\% \quad \text{write power} \\
 & + \quad 5\% \quad \text{simultaneous read/write power} \\
 & + \quad 1\% \quad \text{read power}
 \end{aligned} \tag{5.1}$$

meaning that 98% of the total time the FIFO is collecting data in a slower rate (200kHz), and 6% of the total time the FIFO outputs data to the WSN signal processor in a faster rate (5MHz). This characteristic is revealed in Figure 5.26.

Process, voltage, and temperature variations are simulated. The proposed design is fully functional within +/-10% voltage variation, 0°C to 100°C temperature variation, and all process corners. Simulation results are summarized in Table 5.5, Table 5.6, and Table 5.7.

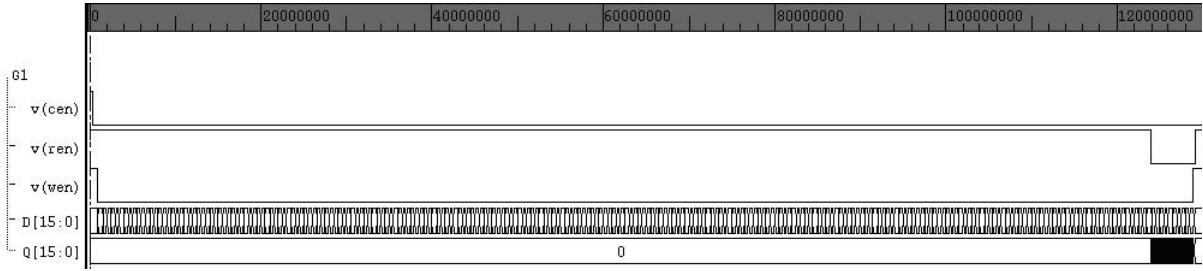


Figure 5.26: Waveform of a complete data collection and data output.

Table 5.5: Process corner simulation (@500mV; 25°C).

	Power (uW)
SS	1.46
SF	1.99
TT	2.21
FS	1.75
FF	4.54

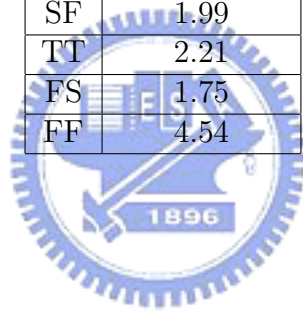


Table 5.6: Voltage variation simulation (@TT corner; 25°C).

	Power (uW)
450mV	1.46
500mV	2.21
550mV	7.42

Table 5.7: Temperature variation simulation (@TT corner; 500mV).

	Power (uW)
0°C	1.46
25°C	1.99
100°C	2.21

Power consumption is compared between various kinds of FIFO architecture, including the conventional register based FIFO, existing dual-port SRAM based FIFO memories shown in Figure 5.27 [43]–[45], and the proposed FIFO memory. All of them are implemented in UMC 90nm CMOS technology with 0.5V supply voltage, but it is worth notice that DP SRAM cell shown in Fig. 5.27 (a) has weak data preservation ability. Cells of existing designs use high- V_T alone for fair power comparison. If existing designs use regular- V_T instead, the power consumption improvement of the proposed scheme will be more significant.

Figure 5.28 shows the power consumption comparisons between conventional schemes mentioned above and the proposed scheme. For conventional register based FIFO, although long bitlines don't exist, a register consumes larger power than an SRAM cell, and the register array size should be two times of a dual-port SRAM array size in order to perform simultaneous read/write, therefore results in much larger power consumption. For conventional dual-port SRAM cell based FIFO memories, bitline overhead and the lack of self-adaptive power control, complementary power gating techniques, result in larger power consumption compared to the proposed scheme.

To sum up, the proposed design has 94%, 79%, 21%, 16%, 16% power reduction compared to conventional register based FIFO, DP SRAM based FIFO, 8T SRAM based FIFO, 10T_C SRAM based FIFO, and 10T_K SRAM based FIFO, respectively.

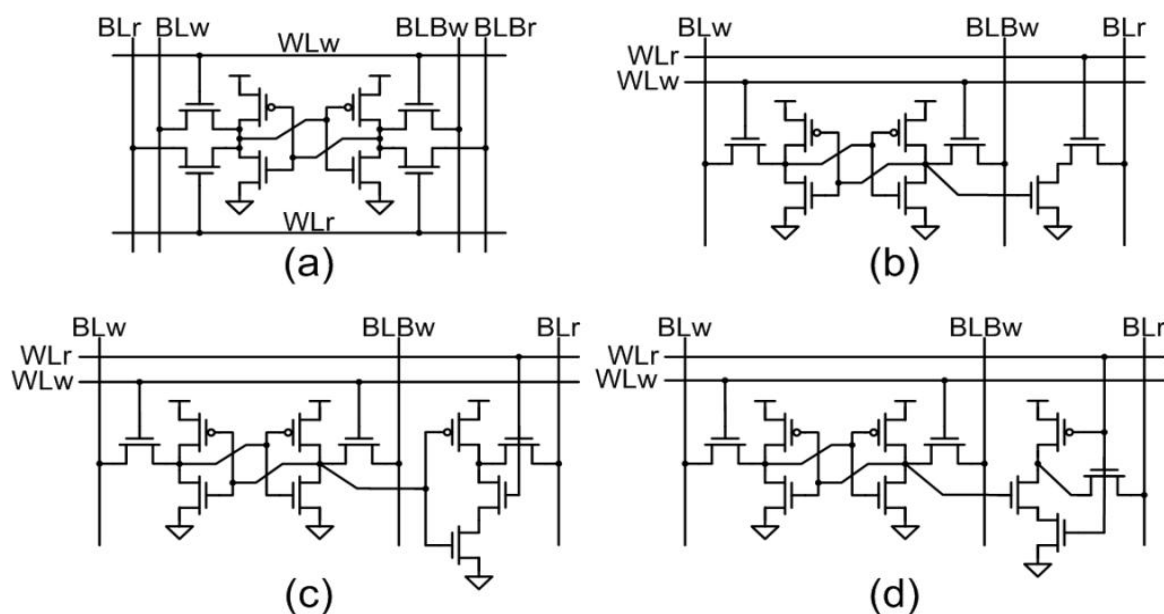


Figure 5.27: Dual-port SRAM cells. (a) DP SRAM cell. (b) 8T SRAM cell. (c) 10T_C SRAM cell. (d) 10T_K SRAM cell.

5.6 Summary

First-in first-out (FIFO) memories are widely used in SoC for data buffering and flow control. In this chapter, a robust ultra-low power asynchronous FIFO memory is proposed. With the self-adaptive power control and complementary power gating technique, leakage power in the FIFO memory array is minimized. Furthermore, to guarantee functionality under low voltage operation, a dual-port dual- V_T 7T SRAM cell with reduced

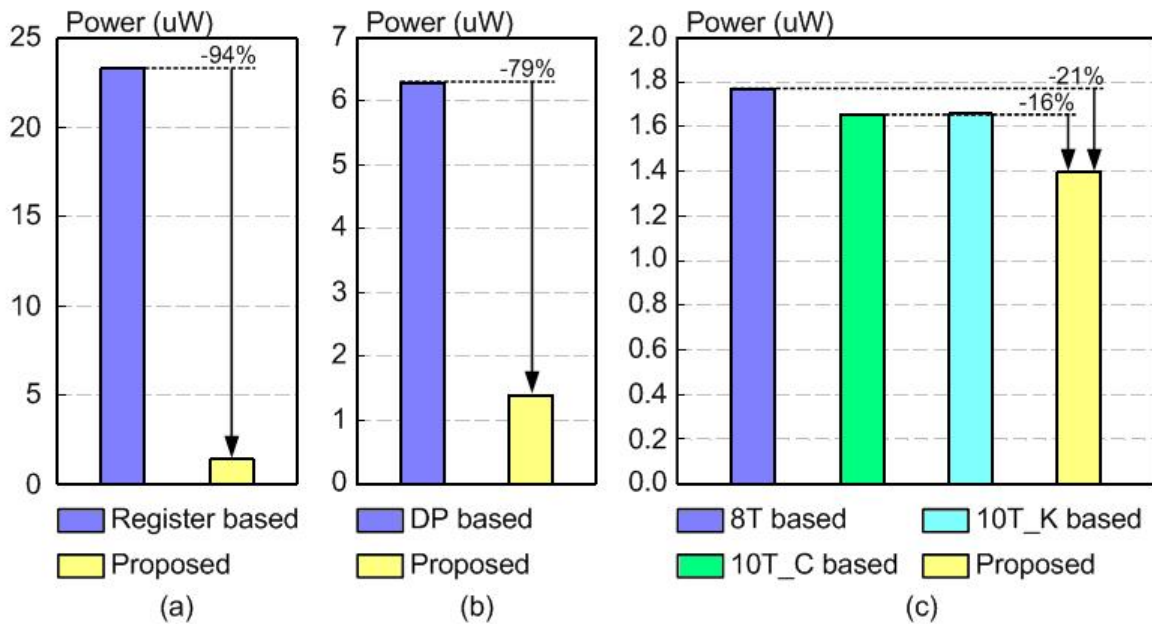


Figure 5.28: Power consumption comparisons between conventional schemes and the proposed scheme. (a) Register based FIFO. (b) DP SRAM based FIFO. (c) 8T SRAM based FIFO; 10T_C SRAM based FIFO; 10T_K SRAM based FIFO.

bitline overhead is utilized to gain stability improvement and power reduction.

The design is implemented in UMC 90nm CMOS technology, with $2.21\mu\text{W}$ power consumption at 5MHz reading frequency and 200kHz writing frequency. Simulation confirms that the proposed design functions successfully under PVT variations with ultra-low supply voltage. It is also shown that the proposed scheme has up to 94% power reduction over conventional designs. As a result, the proposed FIFO memory enables robust and ultra-low power operation for future SoC applications.

Chapter 6

Conclusions

As SRAM continues to dominate the total area and power in modern SoC, subthreshold SRAM provides an effective strategy for total power saving. Stability issue, being the key concern in subthreshold designs, must be considered seriously. In this thesis, a robust, fully differential subthreshold SRAM cell with auto-compensation is proposed. With the auto-compensation mechanism, the proposed cell is able to achieve better hold stability. Meanwhile, the cell structure eliminates conventional read SNM limitation by isolating storage nodes during read operation. Write ability is also improved by applying write assist technique. Compared to existing fully differential SRAM cell structures, simulation results show that at 200mV supply voltage, the proposed scheme achieves better hold stability, better read stability, and better write ability under process variation.

On the other hand, a robust ultra-low power asynchronous FIFO memory is proposed for energy constrained applications. With the self-adaptive power control and complementary power gating technique, leakage power in the FIFO memory array is minimized. Furthermore, to guarantee functionality under low voltage operation, a dual-port dual- V_T 7T SRAM cell with reduced bitline overhead is proposed to gain stability improvement. Simulation confirms that the proposed design functions successfully under PVT variations with ultra-low supply voltage. The proposed FIFO memory enables robust and ultra-low power operation for future SoC designs.

Bibliography

- [1] S. Hanson, B. Zhai, K. Bernstein, D. Blaauw, A. Brant, L. Chang, K. K. Das, W. Haensch, E. J. Nowak, and D. M. Sylvester, "Ultralow-Voltage, Minimum-Energy CMOS," *IBM Journal of Research and Development*, vol. 50, no. 4/5, pp. 469-490, July/September 2006.
- [2] W. Haensch, E. J. Nowak, R. H. Dennard, P. M. Solomon, A. Bryant, O. H. Dokumaci, A. Kumar, X. Wang, J. B. Johnson, and M. V. Fischetti, "Silicon CMOS devices beyond scaling," *IBM Journal of Research and Development*, vol. 50, no. 4/5, pp. 339-361, July/September, 2006.
- [3] Y. Nakagome, M. Horiguchi, T. Kawahara, and K. Itoh, "Review and Future Prospects of Low-Voltage RAM Circuits," *IBM Journal of Research and Development*, vol. 47, no. 5/6, pp. 525-552, September/November, 2003.
- [4] H. Qin, "Deep Sub-Micron SRAM Design for Ultra-Low Leakage Standby Operation," Ph.D. dissertation, University of California, Berkeley, 2007.
- [5] T. Norgall T, R. Schmidt, T. von der Grun, "Body Area Network, a Key Infrastructure Element for Patient-Centered Medical applications," *Biomed. Tech (Berl)*, 47: 365-368, 2002.
- [6] A. Wang, B. H. Calhoun, and A. P. Chandrakasan, *Sub-threshold Design for Ultra Low-Power Systems*. Springer US, 2006, ch. 1, pp. 1-6.
- [7] S. Roundy, P. Wright, and J. Rabaey, "A Study of Low Level Vibrations as a Power Source for Wireless Sensor Nodes," *Computer Communications*, vol. 26, no. 11, pp. 1131-1144, 2003.
- [8] S. Roundy, D. Steingart, L. Frechette, P. Wright, and J. Rabaey, *Power Sources for Wireless Sensor Networks*. Springer-Verlag Berlin Heidelberg, 2004.
- [9] R. Weinstien, "RFID: A Technical Overview and Its Application to the Enterprise," *IT Professional*, vol. 7, no. 3, pp. 27-33, May-June 2005.
- [10] C. C. Chang, D. Marculescu, "Design and Analysis of a VLIW DSP Core," *Proc. Emerging VLSI Technologies and Architectures*, March 2006.
- [11] M. Nakai, S. Akui, K. Seno, T. Meguro, T. Seki, T. Kondo, A. Hashiguchi, H. Kawahara, K. Kumano, and M. Shimura, "Dynamic Voltage and Frequency Management for a Low-Power Embedded Microprocessor," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 28-35, January 2005.

- [12] K. Romer and F. Mattern, "The Design Space of Wireless Sensor Networks," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 54-61, December 2004.
- [13] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," *IEEE Computer Networks*, vol. 38, no. 4, pp. 393-422, March 2002.
- [14] F. Fallah and M. Pedram, "Standby and Active Leakage Current Control and Minimization in CMOS VLSI Circuits," *IEICE Trans. Electron*, vol. E88-C, no. 4, pp. 509-519, April 2005.
- [15] K. Roy, S. Mukhopadhyay, and H. Mahomoodi-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 305-327, February 2003.
- [16] K. Roy and S. C. Prasad, *Low-Power CMOS VLSI Circuit Design*. New York: Wiley, 2000, ch. 5, pp. 214-222.
- [17] K. M. Cao, W. C. Lee, W. Liu, X. Jin, P. Su, S. K. Fung, J. X. An, B. Yu, C. Hu, "BSIM4 Gate Leakage Model Including Source-Drain Partition," in *IEDM Technical Digest*, December 2000, pp. 815-818, .
- [18] S. Mukhopadhyay, C. Neau, R. T. Cakici, A. Agarwal, C. H. Kim, and K. Roy, "Gate Leakage Reduction for Scaled Devices Using Transistor Stacking," *IEEE Trans. VLSI System*, vol. 11, no. 4, pp. 716-730, August 2003.
- [19] N. Yang, W. K. Henson, and J. Wortman, "A Comparative Study of Gate Direct Tunneling and Drain Leakage Currents in N-MOSFETS with Sub-2100-nm Gate Oxides," *IEEE Trans. Electron Devices*, vol. 47, pp. 1636-1644, August 2000.
- [20] K. Nii, Y. Tsukamoto, T. Yoshizawa, S. Imaoka, Y. Yamagami, T. Suzuki, A. Shibayama, H. Makino, and S. Iwade, "A 90-nm Low-Power 32-kB Embedded SRAM With Gate Leakage Suppression Circuit for Mobile Applications," *IEEE J. Solid-State Circuits*, vol. 39, no. 4, pp. 684-693, April 2004.
- [21] Semiconductor Industry Association, International Technology Roadmap for Semiconductors, 2003 ed., <http://public.itrs.net>.
- [22] H. J. M. Veendrick, "Short-Circuit Dissipation of Static CMOS Circuitry and Its Impact on the Design of Buffer Circuits," *IEEE J. Solid-State Circuits*, vol. sc-19, no. 4, pp. 468-473, August 1984.
- [23] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-Power CMOS Digital Design," *IEEE J. Solid-State Circuits*, vol. 27, no. 4, pp. 473-484, April 1992.
- [24] M. Horowitz, T. Indermaur, and R. Gonzalez, "Low-Power Digital Design," in *IEEE Symp. Low Power Electronics Technical Digest*, October 1994, pp. 8-11.
- [25] N. H. E. Weste and D. Harris, *CMOS VLSI Design*, 3rd ed., New York: Addison-Wesley, 2005, ch. 2, pp. 98-99.

- [26] J. Chen, L. T. Clark, and Y. Cao, "Ultra-low Voltage Circuit Design in the Presence of Variations," *IEEE Circuit and Devices Magazine*, vol. 21, no. 6, pp. 12-20, November/December 2005.
- [27] T. Kawahara, M. Horiguchi, Y. Kawajiri, G. Kitsukawa, T. Kure, and M. Aoki, "Sub-threshold Current Reduction for Decoded-Driver by Self-Reverse Biasing," *IEEE J. Solid-State Circuits*, vol. 28, no. 11, pp. 1136-1144, November 1993.
- [28] S. Mukhopadhyay, C. Neau, R. T. Cakici, A. Agarwal, C. H. Kim, and K. Roy, "Gate Leakage Reduction for Scaled Devices Using Transistor Stacking," *IEEE Trans. VLSI Systems*, vol. 11, no. 4, pp. 716-730, August 2003.
- [29] L. Wei, Z. Chen, M. Johnson, K. Roy, Y. Ye, and V. De, "Design and Optimization of Dual Threshold Voltage Circuits for Low Voltage Low Power Applications," *IEEE Trans. VLSI Systems*, vol. 7, no. 1, pp. 16-24, March 1999.
- [30] L. Wei, Z. Chen, K. Roy, Y. Ye, and V. De, "Mixed-V_{th} (MVT) CMOS Circuit Design Methodology for Low Power Applications", in *IEEE Proc. DAC*, 1999, pp. 430-435.
- [31] J. Y. Lin, L. R. Wang, C. L. Hu, and S. J. Jou, "Mixed-V_{th} (MVT) CMOS Circuit Design for Low Power Cell Libraries," in *IEEE Proc. SOCC*, 2007, pp. 181-184.
- [32] S. Shigematsu, S. Mutoh, Y. Matsuya, Y. Tanabe, and J. Yamada, "A 1-V High Speed MTCMOS Circuit Scheme for Power-Down Applications," *IEEE J. Solid-State Circuits*, vol. 32, No. 6, pp. 861-869, June 1997.
- [33] T. Sakurai, "Perspectives on Power-Aware Electronics," in *ISSCC Dig. Tech. Papers*, February 2003, pp. 26-29.
- [34] W. Hwang, (2008), "Embedded Memory Design", Lecture/Class, National Chiao Tung University.
- [35] E. Grossar, M. Stucchi, K. Maex, and W. Dehaene, "Read Stability and Write-Ability Analysis of SRAM Cells for Nanometer Technologies," *IEEE J. Solid-State Circuits*, vol. 41, no. 11, pp. 2577-2588, November 2006.
- [36] B. H. Calhoun and A. P. Chandrakasan, "Static Noise Margin Variation for Sub-threshold SRAM in 65-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, no. 7, pp.1673-1679, July 2006.
- [37] E. Seevinck, F. List, and J. Lohstroh, "Static Noise Margin Analysis of MOS SRAM Cells," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 5, pp. 748-754, October 1987.
- [38] A. Raychowdhury, S. Mukhopadhyay, and K. Roy, "A Feasibility Study of Sub-threshold SRAM Across Technology Generations," in *IEEE Proc. ICCD*, October 2005, pp. 417-412.
- [39] R. Heald and P. Wang, "Variability in Sub-100nm SRAM Designs," in *IEEE/ACM Proc. ICCAD*, November 2004, pp. 347-352.

- [40] N. Verma and A. P. Chandrakasan, "A 256kb 65nm 8T Subthreshold SRAM Employing Sense-Amplifier Redundancy," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 141-149, January 2008.
- [41] J. P. Kulkarni, K. Kim, and K. Roy, "A 160 mV Robust Schmitt Trigger Based Subthreshold SRAM," *IEEE J. Solid-State Circuits*, vol. 42, no. 10, pp. 2303-2313, October 2007.
- [42] I. J. Chang, J. J. Kim, S. P. Park, and K. Roy, "A32kb 10T Subthreshold SRAM Array with Bit-Interleaving and Differential Read Scheme in 90nm CMOS," in *ISSCC Dig. Tech. Papers*, February 2008, pp. 388-389.
- [43] L. Chang, D. M. Fried, J. Hergenrother, W. Sleight, R. H. Dennard, R. K. Montoye, L. Sekaric, S. J. McNab, A. W. Topol, C. D. Adams, K. W. Guarini, and W. Haensch, "Stable SRAM Cell Design for the 32nm Node and Beyond," in *Symposium on VLSI Dig. Tech. Papers*, June 2005, pp. 128-129.
- [44] B. H. Calhoun, and A. P. Chandrakasan, "A 256kb 65-nm Sub-threshold SRAM Design for Ultra-Low Voltage Operation," *IEEE J. Solid-State Circuits*, vol. 42, no. 3, pp.680-688, March 2007.
- [45] T. H. Kim, J. Liu, J. Keane, and C. H. Kim, "A High-Density Subthreshold SRAM with Data-Independent Bitline Leakage and Virtual Ground Replica Scheme," in *ISSCC Dig. Tech. Papers*, February 2007, pp. 330-331.
- [46] J. Y. Yu, W. C. Liao, and C. Y. Lee, "An MT-CDMA Based Wireless Body Area Network for Ubiquitous Healthcare Monitoring," in *IEEE BioCAS*, November 2006.
- [47] J. Y. Yu, C. C. Chung, W. C. Liao, and C. Y. Lee, "A sub-mW Multi-Tone CDMA Baseband Transceiver Chipset for Wireless Body Area Network Applications," in *ISSCC Dig. Tech. Papers*, February 2007, pp. 364-365.
- [48] N. Shibata, M. Watanabe, and Y. Tanabe, "A Current-Sensed High-Speed and Low-Power First-In First-Out Memory Using a Wordline/Bitline-Swapped Dual-Port SRAM Cell," *IEEE J. Solid-State Circuits*, vol. 37, no. 6, pp. 735-750, June 2002.
- [49] G. Gerosa, S. Gary, C. Dietz, P. Dac, K. Hoover, J. Alvarez, H. Sanchez, P. Ippolito, N. Tai, S. Litch, J. Eno, J. Golab, N. Vanderschaaf, J. Kahle, "A 2.2 W, 80 MHz Superscalar RISC Microprocessor," *IEEE J. Solid-State Circuits*, vol. 29, no. 12, pp. 1440-1454, December 1994.
- [50] V. Stojanovic and V. Oklobdzija, "Comparative Analysis of Master-Slave Latches and Flip-Flops for High-Performance and Low-Power Systems," *IEEE J. Solid-State Circuits*, vol. 34, no. 4, pp. 536-548, April 1999.
- [51] H. Partovi, R. Burd, U. Salim, F. Weber, L. DiGregorio, D. Draper, "Flow-through Latch and Edge-Triggered Flip-Flop," in *ISSCC Dig. Tech. Papers*, February 1996, pp. 138-139.
- [52] J. Montanaro, R. T. Witek, K. Anne, A. J. Black, E. M. Cooper, D. W. Dobberpuhl, P. M. Donahue, J. Eno, W. Hoepfner, D. Kruckemyer, T. H. Lee, P. C. M. Lin, L. Madden, D. Murray, M. H. Pearce, S. Santhanam, K. J. Snyder, R. Stehpany, S. C.

- Thierauf, "A 160-MHz, 32-b, 0.5-W CMOS RISC Microprocessor," *IEEE J. Solid-State Circuits*, vol. 31, no. 11, pp. 1703-1714, November 1996.
- [53] B. Fu and P. Ampadu, "Comparative Analysis of Ultra-Low Voltage Flip-Flops for Energy Efficiency," in *IEEE Proc. ISCAS*, May 2007, pp. 1173-1176.
- [54] Hao-I Yang, Ming-Hung Chang, Ssu-Yun Lai, Hsiang-Fei Wang, and Wei Hwang, "A Low-Power Low-Swing Single-Ended Multi-Port SRAM," in *IEEE VLSI-DAT*, April 2007.
- [55] C. A. Otto, E. Jovanov, and A. Milenkovic, "A WBAN-based System for Health Monitoring at Home," in *IEEE-EMBS Proc. International Summer School and Sym. Medical Devices and Biosensors*, September 2006, pp. 20-23.
- [56] E. Jovanov, A. Milenkovic, C. A. Otto, P. C. de Groen, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *J. NeuroEngineering and Rehabilitation*, 2:6, March 2005.
- [57] K. Takeda, H. Ikeda, Y. Hagihara, M. Nomura, and H. Kobatake, "Redefinition of Write Margin for Next-Generation SRAM and Write-Margin Monitoring Circuit," in *ISSCC Dig. Tech. Papers*, February 2005, pp. 478-479.

