Design Optimization of pMOS only eDRAM Macro Cell at 28 nm Node

M.Tech. Thesis

By

ABU SAID PARVEJ ALAM

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 2025

Design Optimization of pMOS only eDRAM Macro Cell at 28 nm Node

A THESIS

Submitted in partial fulfillment of the
requirements for the award of the degree
of
Master of Technology

by

ABU SAID PARVEJ ALAM

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 2025

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **Design Optimization of pMOS only eDRAM Macro Cell at 28 nm Node** in the partial fulfillment of the requirements for the award of the degree of **MASTER OF TECHNOLOGY** and submitted in the **DEPARTMENT OF ELECTRICAL ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from July 2024 to May 2025. Thesis submission under the supervision of Prof. Abhinav Kranti, IIT Indore.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Alam 22-05-2025

Signature of the student with date
(ABU SAID PARVEJ ALAM)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

22/05/2025

Signature of the Supervisor of
M.Tech. thesis (with date)
(Prof. Abhinav Kranti)

ABU SAID PARVEJ ALAM has successfully given his M.Tech. Oral Examination held on **May 5, 2025.**

Signature of Supervisor of M.Tech. thesis
Date: 22/05/2025

Saptarshi Ghosh

Convener, DPGC
Date: 22-05-2025

ACKNOWLEDGEMENTS

First and foremost, I would want to sincerely thank Prof. Abhinav Kranti, my project supervisor, for the guidance and encouragement during the M.Tech. thesis research work. His motivation and mentorship played a crucial role in the successful completion of my M.Tech. project.

I am deeply grateful to IIT Indore for providing the necessary infrastructure and research environment to carry out this work. I also thank the Department of Electrical Engineering, IIT Indore. I am grateful to the Ministry of Education, Government of India, and IIT Indore for financial support in the form of stipend during my M.Tech. program.

I would like to thank Dr. Manish Gupta, BITS Pilani, K.K. Birla Goa Campus, for allowing access to the TCAD simulation tools. I am especially thankful to my senior, Mr. Rohit Kumar Nirala, for his continuous support, guidance, and assistance in the laboratory. I also appreciate the help and support provided by Mr. Aman Chandrakar during this work.

I express my heartfelt gratitude to my parents, Mr. Oliul Islam Mandal and Mrs. Firoja Bibi, for their unconditional love and support. I also want to thank my fellow classmates for their support, friendship, and the priceless memories we made together while we were in Indore.

Abu Said Parvej Alam

Dedicated to my family

Abstract

Design Optimization of pMOS only eDRAM Macro Cell at 28 nm Node

Embedded DRAM (eDRAM) architectures have been driven by growing need for energy-efficient and high-density memory. The 2-Transistor 0 Capacitor (2T0C) Gain Cell (GC) appears as compact logic-compatible substitute for traditional 1T1C DRAMs. However, a main drawback of the 2T0C GC is its low Data Retention Time (DRT), which is primarily degraded due to capacitive coupling and leakage currents at advanced technology nodes.

This thesis explores an effective approach to enhance DRT by using Double Gate (DG) topology with 2T pMOS-only GC structure at the 28 nm technology node. DG transistors offer better electrostatic control over the channel, boost driving current and reduced leakage current, and tunable threshold voltages through independent gate operation, all of which are crucial for GC. Important contribution of this thesis includes a two-step DRT enhancement approach. First, minimizing capacitive coupling by independent gate bias operation to allow a larger voltage difference (ΔV) between to logic levels. Second, suppressing leakage by write bit line (WBL) bias tuning to significantly lower the data degradation rate. The thesis work also analyses the effects of high temperature operation and supply voltage downscaling on DRT.

Table of Contents

Contents

List of Figures

List of Tables

NOMENCLATURE

Notation	*Description*	*Unit*
L	*Gate length*	*nm*
W	*Width*	*nm*
T_{OX}	*Oxide thickness*	*nm*
T_{Si}	*Silicon thickness*	*nm*
V_G	*Gate voltage*	*V*
V_{BG}	*Back gate voltage*	*V*
V_D	*Drain voltage*	*V*
V_S	*Source voltage*	*V*
V_{DD}	*Supply voltage*	*V*
I_{ON}	*On current*	*mA*
I_{OFF}	*Off current*	*nA*
I_{SD}	*Source to drain current*	*A*
V_{SD}	*Source to drain voltage*	*V*
V_{SG}	*Source to gate voltage*	*V*
V_{GS}	*Gate to source voltage*	*V*
V_{DS}	*Drain to source voltage*	*V*
V_{Th}	*Threshold voltage*	*V*
V_{WWL}	*Write word line voltage*	*V*
V_{WWL1}	*Write word line-1 voltage*	*V*
V_{WWL2}	*Write word line-2 voltage*	*V*
V_{WBL}	*Write bit line voltage*	*V*
V_{SN}	*Storage node voltage*	*V*
V_{RBL}	*Read bit line voltage*	*V*
V_{RWL}	*Read word line voltage*	*V*
V_{RWL2}	*Read word line-2 voltage*	*V*
μ_p	*Hole mobility*	$cm^2/V.s$
ΔV	*Voltage difference between two logic levels*	*V*

ΔV_{SN}	*Change in storage node voltage*	*V*
ΔV_{WWL1}	*Change in write word line-1 voltage*	*V*
ΔV_{WWL2}	*Change in write word line-2 voltage*	*V*
RAT	*Read access time*	*ns*
P	*Power*	*W*

ACRONYMS

1T1C	*One transistor and one capacitor*
2T	*Two transistor*
2T0C	*Two transistor and zero capacitor*
3T	*Three transistor*
3T0C	*Three transistor and zero capacitor*
4T	*Four transistor*
4T0C	*Four transistor and zero capacitor*
5T	*Five transistor*
6T	*Six transistor*
8T	*Eight transistor*
10T	*Ten transistor*
AI	*Artificial Intelligence*
CC	*Capacitive Coupling*
CIM	*Compute-in-memory*
CMOS	*Complementary Metal Oxide Semiconductor*
CPU	*Central Processing Unit*
DG-MOSFET	*Double Gate Metal Oxide Semiconductor Field Effect Transistor*
DG-pMOS	*Double gate p-type MOS*
DRAM	*Dynamic Random-Access Memory*
DRT	*Data Retention Time*
eDRAM	*Embedded DRAM*
FDSOI	*Fully Depleted SOI*
GC	*Gain Cell*
GC-eDRAM	*Gain Cell Embedded DRAM*
GPU	*Graphics Processing Unit*
IoT	*Internet-of-Things*
ML	*Machine Learning*

MOS	*Metal Oxide Semiconductor*
MOSCAP	*MOS Capacitor*
MOSFET	*Metal Oxide Semiconductor Field Effect Transistor*
MR	*Read Transistor*
MW	*Write Transistor*
nMOS	*n-type MOS*
pMOS	*p-type MOS*
PC	*Personal Computer*
PCB	*Printed Circuit Board*
PUF	*Physical Unclonable Function*
RBL	*Read Bit Line*
RWL	*Read Word Line*
Si	*Silicon*
SN	*Storage Node*
SOI	*Silicon on Insulator*
SoC	*System-on-Chip*
SRAM	*Static Random-Access Memory*
WBL	*Write Bit Line*
WWL	*Write Word Line*
WWL_1	*Write Word Line-1*
WWL_2	*Write Word Line-2*

Chapter 1

Introduction

Semiconductor memory is regarded as the central component of any digital system in the world. Every electronic device, from personal computers to smartphones, from artificial intelligence processors to tiny gadgets or Internet-of-Things (IoT) sensors, requires some kind of memory to store data. The features of memory impact the overall system performance, like speed, area, power consumption and even cost [1]. Semiconductor memories are designed for fast access by Central Processing Units (CPUs) to hold binary data and instructions so that processors can access them quickly. Specifically, embedded memories become an essential component of designs for System-on-Chip (SoC) [2]. As technology progresses, the need for denser, faster, and power-efficient memory devices has grown exponentially. These demands have driven innovations in Static Random-Access Memory (SRAM), Dynamic Random-Access Memory (DRAM) and Gain Cell embedded DRAM (GC-eDRAM) memory architectures [3].

1.1 Overview of Memory

Memory can be broadly classified into two primary groups: non-volatile and volatile memory [4]. Whereas non-volatile memories store information regardless of power, volatile memories destroy their retained data once the supply is removed. SRAM and DRAM are the two main forms of volatile memories [5].

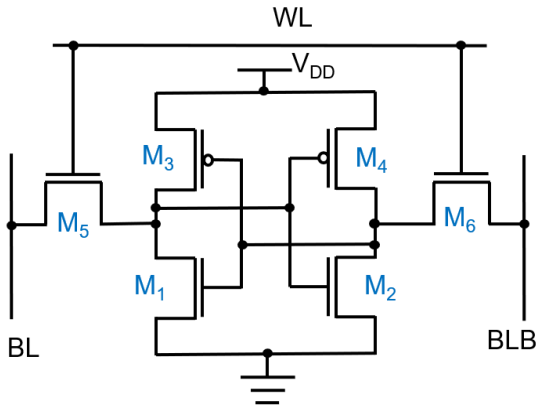


Fig. 1.1 Schematic diagram of a 6T SRAM bitcell [6].

1.1.1 Static Random-Access Memory (SRAM)

Over time, several memory technologies have been developed, each with unique benefits and drawbacks. One of the most used forms is SRAM, where data is stored in a bistable flip-flop circuit [6]. SRAM typically stores a single bit of information using six transistors (6T). A standard SRAM cell architecture consists of two access transistors and two cross-coupled inverters (Fig. 1.1). This configuration ensures that the kept data remains intact as long as power is accessible, without needing any periodic refresh operations. The benefits of SRAM include very quick read and write operations, great stability and reliability, and low dynamic power consumption during access. These qualities make SRAM a popular choice for small buffers, high-speed registers, and CPU cache memories. SRAM is still remained essential for applications requiring very quick access and minimal latency, like the L1 and L2 cache memories in modern microprocessors [6]. However, SRAM uses more area. Hence, memory density is reduced. Furthermore, SRAM is less suitable for very large memories (main memories) since leakage currents cause it to consume considerable static power, particularly in advanced technology nodes such as 28 nm or below [7]. Another challenge of 6T SRAM is the difficulty in scaling cells at lower supply voltages (V_{DD}). The noise margins get reduced, which makes SRAM cells more vulnerable to failure due to V_{DD} fluctuations [4]. In order to guarantee dependable subthreshold (sub-V_{Th})

operation, alternative SRAM cells with 8, 10, or even up to 14 transistors are necessary [8].

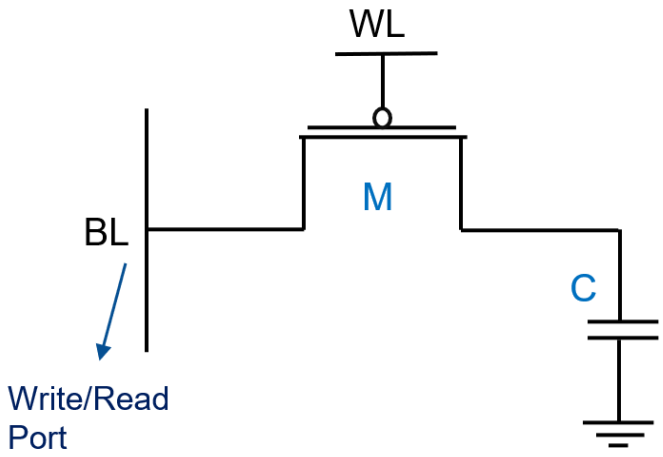


Fig. 1.2 Schematic diagram of a 1T1C DRAM bitcell [9].

1.1.2 Dynamic Random-Access Memory (DRAM)

Another common type is DRAM, which keeps each piece of information using one transistor and one capacitor (1T1C) [9]. Fig. 1.2 shows a schematic diagram of a 1T1C DRAM bitcell. The fundamental idea behind DRAM is very simple: charge either present or absent on the capacitor marks a binary '1' or '0', respectively. The biggest benefit of DRAM is its high density. Only one transistor and one capacitor allow DRAM for the formation of vast memory arrays. Because of this, DRAM is the recommended option for main memory (RAM) in servers, graphics cards, and Personal Computers (PCs) [5]. However, DRAM suffers from a leakage problem. Even if the device is kept powered on, the charge that is stored on the capacitor eventually tends to leak out. Because of this, DRAM cells need to be refreshed on a regular basis in order to recover lost charge and preserve data integrity. As a result, periodic refreshing increases energy consumption and complicates the memory controller [10].

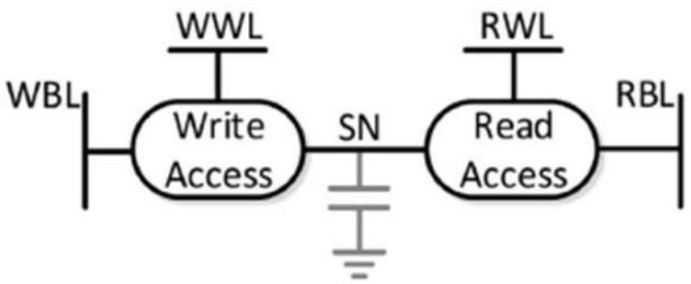


Fig. 1.3 Basic configuration of GC-eDRAM macrocell [11].

1.1.3 Gain Cell Embedded DRAM (GC-eDRAM)

As computers became more powerful and needed to handle larger amounts of data quickly, embedded DRAM (eDRAM) [5] was introduced to combine the advantages of DRAM (high density) with better integration in logic processes. Larger memory embedded straight on the logic process chip, such as CPUs or Graphics Processing Units (GPUs), always benefits rather than depending on outside main memory. The main reasons for this are: (1) higher system-level integration densities are facilitated by embedded memory, and (2) going off-chip connectivity via capacitive lines and input/output (I/O) terminals on printed circuit boards (PCBs) differs from on-chip connections and demands more power consumption [12]. Two forms of eDRAM are examined: (1) traditional 1T1C eDRAM, where a single access transistor and a unique, high-density, 3D capacitor (trench or stack capacitor) builds a macrocell [9], and (2) GC-eDRAM, whose macrocell is built from 2 to 5 MOSFETs [13], [14]. Fig. 1.3 shows the basic configuration of the GC-eDRAM bitcell. Traditional 1T1C eDRAM processes are not completely compatible with logic complementary metal-oxide-semiconductor (CMOS) technologies as it requires a special fabrication process to create a dense trench or stacked capacitors on chip [9], for which manufacturing becomes complex and costly. An alternative, GC-eDRAMs are thus completely logic compatible and readily implemented into SoCs with no additional price. GC-eDRAMs promise a good balance between storage density and speed, which makes it a low-

power alternative to SRAM [15], [16], [17], [18], [19], [20]. The following table summarises the key differences:

Table 1.1 Comparison of SRAM, DRAM and GC-eDRAM [1].

Feature	SRAM	DRAM	GC-eDRAM
Cell Size	Large	Small	Small
Refresh Needed	No	Yes	Yes
Process Complexity	Low	High	Low
Read Destructive	No	Yes	No
Leakage Power	High	Low	Low
Access Speed	Very Fast	Medium	Fast

1.2 Evaluation of Gain Cells (GCs)

The continuous scaling of CMOS technology has enabled a significant improvement in the performance of digital systems. But this aggressive scaling has also introduced some critical challenges, especially for embedded memories, which occupy at least half of SoCs' overall size and power budget [21]. As technology scaled down to nanometer sizes, traditional memory architectures like SRAM and DRAM have faced significant limitations. In this regard, eDRAM macrocell, also known as gain cell-based memories (GC memories), have developed, which provide a balanced solution between speed, storage density, and energy efficiency.

To store data as an electric charge, the traditional eDRAM bitcell needed a physically built storage capacitor [12]. It additionally utilizes the storage capacitor for write and read operations using a single MOSFET.

Unfortunately, making high-density trench or stacked capacitors for this kind of conventional eDRAM needs extra steps in the processing line. As feature sizes reduced, building sufficiently large and reliable capacitors became extremely difficult. That is why it is incompatible with common CMOS technologies [9]. On the other hand, traditional SRAM cells have lower noise margins and higher leakage currents in highly scaled nodes (65 nm, 45 nm, and 28 nm). As technology scales further, the transistor mismatch and reduced V_{DD} make SRAM less reliable unless an extra area and complexity are sacrificed, such as using 8T or 10T SRAM designs [8].

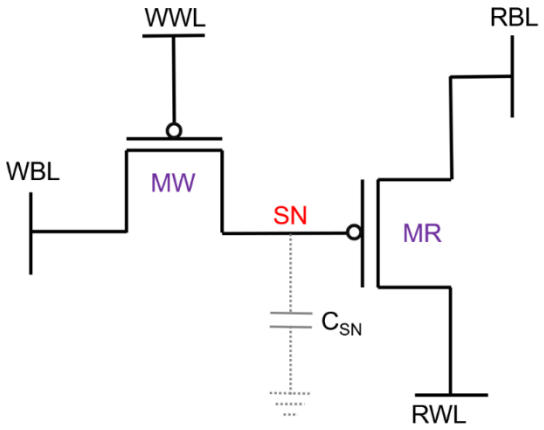


Fig. 1.4 Schematic diagram of a 2T0C DRAM macrocell [15].

The gain cell (GC) concept [22] was introduced to overcome the limitations of traditional SRAM and DRAM memories. It offers a middle-ground solution by providing logic compatibility, small area size, and reduced leakage without the need for any complicated process steps. A GC is a memory structure that stores data dynamically on the parasitic capacitance of a transistor's node (called the storage node, SN) rather than an explicit capacitor. A typical GC uses two transistors, a write transistor (MW) to write the data into the SN and a read transistor (MR) to read the stored data (Fig. 1.4). To improve stability or retention performance, some topologies use one or two additional transistors. GC-eDRAM is completely compatible with popular digital CMOS technologies, as the device is

fabricated solely from accessible metal and vias and MOSFETs. MOSFETs are utilized as Metal Oxide Semiconductor Capacitor (MOSCAP) and as access transistors (write and read transistors). One can improve the storage node capacitance by use of metal layers and vias [1]. Since GC-eDRAM offers many of the benefits of both SRAM (like standard digital CMOS logic compatibility) and 1T1C eDRAM (like better storage density) and prevents many of the problems of both SRAM (like the large bitcell) and 1T1C eDRAM (like destructive reading, restoring, and additional manufacturing expenses for specific processes) [12], it is an interesting substitute. That is why GC-eDRAM macrocells are easy to integrate into every digital system without extra production expenses. Dynamic data retention characteristics are the primary issue of GC-eDRAM over SRAM; hence, it demands periodic refresh operations.

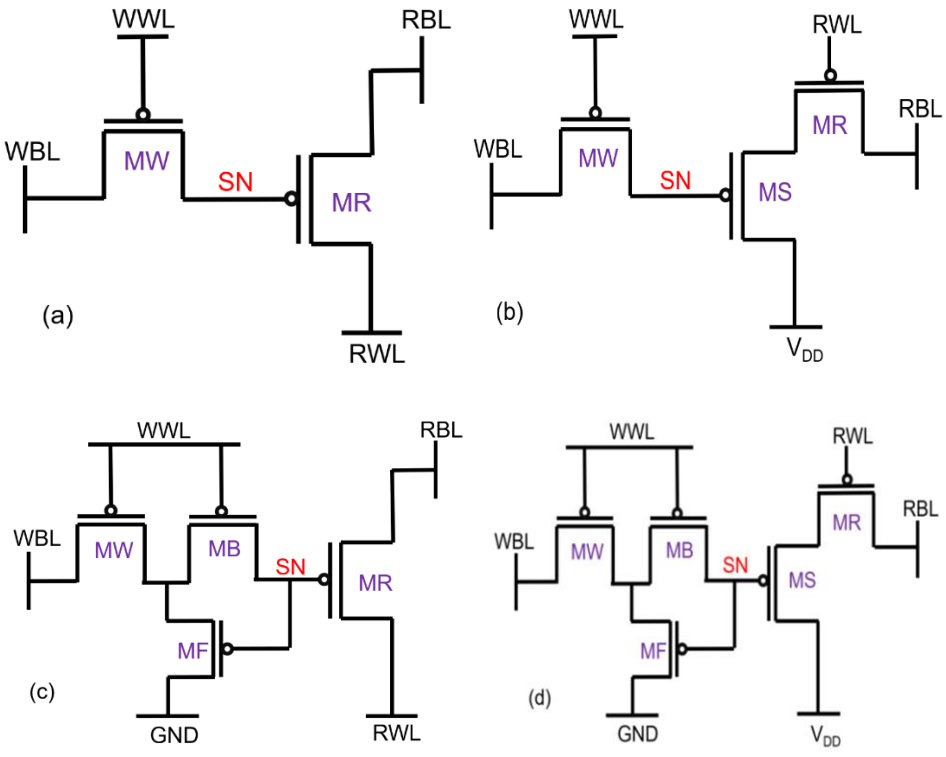


Fig. 1.5 Schematic representation of (a) 2T Gain Cell [16], (b) 3T Gain Cell [18], (c) 4T Gain Cell [23], and (d) 5T Gain Cell [14].

Over the past ten years, a wide range of GC topologies with 2-5 transistors have been proposed [14], [16], [18], [23]. A basic example of these GCs is displayed in Fig. 1.5. Every one of these designs utilizes a MW to access and store charge on the SN. Additionally, every GC topology has an SN capacitor made up of MW junction capacitance and MOSCAP. In the two-transistor (2T) configuration, the MOSCAP storage transistor can also be used as MR. The comparatively bigger three-transistor (3T) [18] version demonstrates a more reliable read operation because it uses a distinct MR that separates the read bit line (RBL) from MR. For enhanced read robustness, certain four-transistor (4T) [23], [24] GCs use a fourth transistor as MOSCAP capacitively connects the RBL to the SN and boosts the SN capacitor. The word ‘gain’ refers to the read transistor's transconductance gain, which indicates little changes at storage node voltage (gate voltage of MR) can result in a change in read path current (drain current of MR) [25].

1.3 Different Types of GC Topology

Several gain cell topologies have been proposed in the research community, each offering different trade-offs between area, speed, retention time, and complexity. While the fundamental principle of storing data dynamically using parasitic capacitance remains the same, the number of transistors in all of these circuits is lower than in standard SRAM circuits. This section will explore the popular GC topologies, such as 2T, 3T, 4T, and 5T, with their structure, fundamental principles, benefits, drawbacks, and trade-offs.

1.3.1 2T Gain Cell

The two-transistor (2T) gain cell [16], [17], [26], [27], [28] is the simplest and most area-efficient design. In 2T designs, the simple topology usually has a MW for writing and a MR that stores information and also reads it. The parasitic capacitance corresponding to SN is typically formed between the drain region and the substrate or adjacent interconnects.

As soon as the write word line (WWL) is asserted, MW links the write bit line (WBL) to SN and writes the available data at WBL to SN [16]. When the read word line (RWL) is asserted, MR drives a current through the RBL, and that is how we read the data from the SN without destroying the stored charge. During hold mode, both transistors remain off, and the charge (data) stored on SN is retained for a certain duration depending on leakage currents. However, data retention time (DRT) is comparatively limited because the storage node is weakly held [29].

One of the major advantages of 2T gain cells is area efficiency. With only two transistors, the cell footprint is very small compared to SRAM or DRAM, enabling higher memory density. Another benefit is non-destructive read, as the read transistor allows current through the readout path without lowering the stored charge [30].

The 2T topology suffers from lower retention time. Capacitive coupling (CC) effects between SN and the control lines (WWL, RWL, and RBL) arise from the 2T configuration. CC and leakage paths through MW and MR cause rapid data loss, which compromises the data integrity and reduces efficiency [31]. Thus, a third transistor is usually added, mostly to prevent RBL leakage and to avoid annoying capacitive couples between SN and RWL. In modern designs, implementation of several techniques, like using high-threshold (high-V_{Th}) write transistor, body biasing, and optimized WBL during hold, has been applied to extend DRT, which makes it suitable for embedded memory applications [32].

1.3.2 3T Gain Cell

To solve problems observed in 2T, the three-transistor (3T) gain cell [33], [34], [35], [36], [37], [38], [39] is utilized. This design reduced threshold voltage drop during write [40] and improved isolation between the write and read paths. Some designs also minimize RBL leakage and

prevent disrupting CC from RWL to SN. This helps maintain data for longer times without considerably expanding the area and makes it more robust than the 2T model. To improve retention and lower refresh requirements, certain 3T designs incorporate asymmetric transistor sizing or body biasing approaches [41]. Techniques like preferential boosting, where the read bit line or internal node is dynamically boosted by tying the MR drain to RWL instead of ground. Hence reducing some of the positive SN voltage in the pMOS MW design, further improve access times and prevent read-disturb [18]. The 3T GC improves noise margins and minimizes disturbance during read operations by having a dedicated read buffer. But it comes with a larger area and a little more control complexity. To maintain high-speed performance, surprisingly big cache memory designs choose the 2T topology at the expense of extra peripheral circuits [16], [26], [30].

1.3.3 4T Gain Cell

Turning now to the four-transistor (4T) gain cell, the design usually incorporates extra transistors for data integrity, boosting, or decoupling. These are frequently utilized in applications that require more stability and prefer refresh-free operation [23], [24]. Common designs use two transistors for write or access and two for read or isolation. The additional transistors contribute in the more separation of the write and read paths for non-destructive reads [20]. Another 4T design with three transistors and a ‘gated diode’ (MOS transistor serving as a storage device and amplifier for the cell voltage) enables techniques like preferential boosting for improved signal levels [42]. Even though 4T topologies have improved retention and better control over noise and leakage, they occupy more area and need complex circuit design because of more word lines or control signals.

1.3.4 5T Gain Cell

The five-transistor (5T) gain cell is a more modern invention designed for specialized applications like parallel sensing, ternary logic, or ultra-low power operation [43]. This design typically has three transistors for write/read and two for storage or level tuning. The extra transistor enables the storage node to be biased at an intermediate voltage level, so storing either '0,' '1,' or an intermediary 'X' logic [14]. This configuration allows more bits to be stored per cell. Other 5T designs, such as those in ultra-low power FDSOI, are perfect for IoT or biomedical edge applications because they achieve picowatt-level standby power [44]. Still, the area cost increases dramatically, and the design complexity increases as well.

In conclusion, the selection of the gain cell topology is an investigation of design space that balances performance, area, power, and retention. The 2T cells offer the highest density but the weakest retention, 3T maintains a balance with better control, 4T offers reliable read/write operation and long retention, and 5T creates opportunities for more sophisticated features like multi-bit storage and parallel sensing.

Table 1.2 Comparison of different types of GC topology [13].

Feature	**2T GC**	**3T GC**	**4T GC**	**5T GC**
Area Efficiency	Highest	High	Medium	Low
DRT	Low	Medium	High	Very High
Read Stability	Moderate	High	High	Very High
Complexity	Very Low	Low	Medium	High

1.4 Features and Applications of Gain Cells

The GC-eDRAM shows particular characteristics and enables design choices that are uncommon to regular DRAMs and SRAM. The features of GC are described as follows —

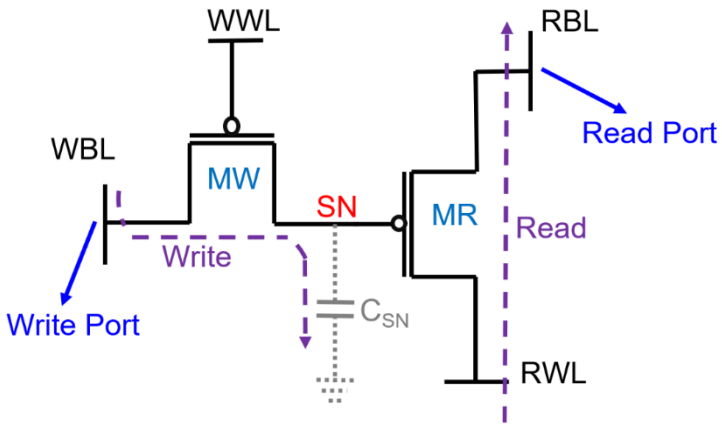


Fig. 1.6 Schematic representation of 2T GC showing write and read port [28].

An important aspect of gain cells is that they have two-port compatibility [28], which is their most notable characteristic. The GC-eDRAM is a dual port memory (Fig. 1.6) since a GC has both a write and a read port. As a result, it usually has separate address decoders for the write and read addresses.

With dual-port compatibility, another most advantageous characteristics of GC-eDRAM architecture is that it supports pipelined refreshing, which makes sure to simultaneously read and rewrite data for different cells within a single clock cycle [45]. This feature drastically lowers dynamic and leakage power compared to traditional DRAM architectures. With dual-port capability, the information can be refreshed by reading the content of the row at generic address n, while writing the previously read content back to the row at generic address n−1 within the same cycle. This capability is absent from traditional single-port DRAMs, which need two cycles to complete the same refresh operation [45].

As GC has two-port functionality for which reading becomes non-destructive and does not require any restore operation, unlike in 1T1C DRAM, where a single port is used for both write and read operations, and after every read operation, the data must be restored back to the capacitor [28]. Since the read path is isolated by another read transistor, the stored data can be sensed without changing the node voltage. By avoiding the need to restore the data, GC-eDRAM decreased power consumption per access cycle and reduced external circuit complexity and protected data during reading [46].

The small area footprint of gain cells is one of their most remarkable characteristics. Conventional 6T SRAM cells consume more space by using access transistors and cross-coupled inverters; gain cells are built by using only two transistors. Higher memory density is especially important in cache memory arrays, Artificial Intelligence (AI) and Machine Learning (ML) accelerator designs where vast amounts of on-chip memory are needed [26]. Apart from area, gain cells show logical process compatibility as discussed earlier.

In addition to features, GC has a wide and expanding range of applications, from low-power IoT sensors to high-speed logic cores. Because of its smaller footprint, GC has recently gained a lot of popularity as an on-chip cache memory, particularly as last-level cache in high-end processors [26]. GC is also used as eDRAM in portable electronic gadgets [47] and IoT deployments for sensor networks [48] because of low power consumption. GCs are utilized in Compute-in-memory (CIM) macros in AI accelerators and edge inference engines in order to reduce data movement overhead, which is one of the main sources of power consumption in neural accelerators [49]. Where power budgets are very limited in biomedical and IoT applications, subthreshold GC designs have shown great success; operating below 0.5V these designs can retain data using minimum power, like pico-watts per bit range [44]. Physical Unclonable Functions (PUFs),

where the inherent variations in leakage and retention time across gain cells are utilized as a form of silicon fingerprinting, is another new application [32].

1.5 Motivation for 2T pMOS-only GC

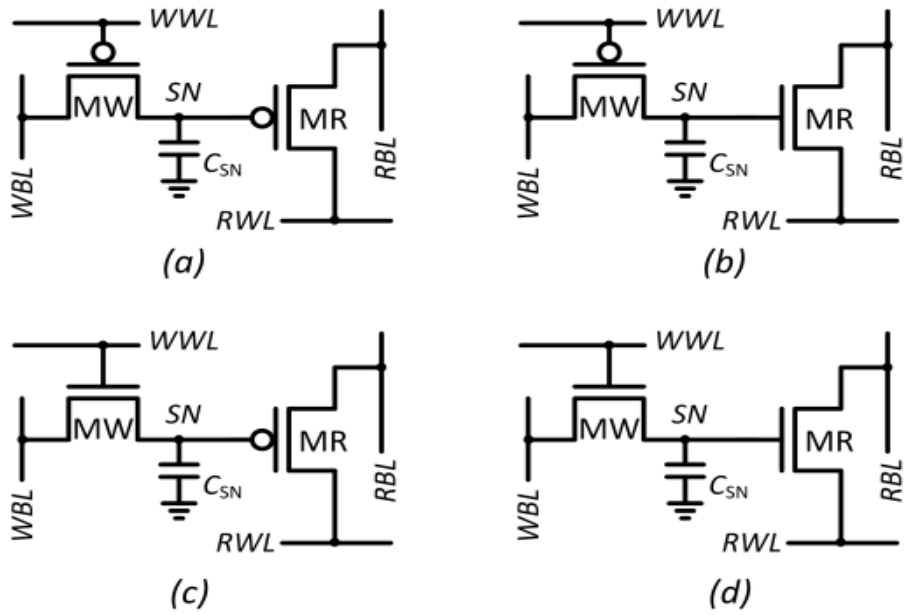


Fig. 1.7 Overview of 2T gain cell implementation options: (a) All pMOS Cell, (b) Mixed pMOS-nMOS Cell, (c) Mixed nMOS-pMOS Cell, and (d) All nMOS Cell [28].

In memory design, selecting the device architecture is vital, particularly when the objective is high data retention, low leakage power, and compatibility with deeply scaled technologies such as 28nm. Depending on the type of devices used (pMOS or nMOS) for these transistors, four 2T configurations are possible, which are shown in Fig. 1.7. The 2T pMOS-only gain cell is the effective option among different gain cell topologies for embedded dynamic memories intended for better retention and high-performance low-power applications [17], [28].

The most convincing reasons for choosing a 2T structure are its simplicity and compactness. It achieves an extremely compact area with just

two transistors, MW and MR, which allows more bits stored in a given silicon area. For applications like on-chip caches and portable electronics, where memory space frequently limits overall chip size, the 2T gain cell is an especially interesting choice [47].

Along with simplicity, choosing to use pMOS-only devices offers several benefits, especially at lower nodes such as 28nm and even below. When compared to nMOS devices, pMOS transistors inherently have lower subthreshold leakage currents. This is because pMOS devices have lower leakage when they are turned off due to their lower carrier mobility (μ) and higher V_{Th}. The main cause for SN charge loss is subthreshold leakage via MW. Thus, selecting the all-pMOS arrangement helps to increase the capacity of the memory to retain information without frequent refresh [16].

1.6 Conclusion

In conclusion, the expansion of memory technologies has created demand for low-power, high-density structures with simple logic circuit integration. GC-eDRAM offers a promising solution by combining the advantages of SRAM and DRAM while dealing with their main drawbacks. Two major benefits of using a 2T pMOS-only gain cell in cutting-edge technologies like 28 nm are simplicity and leakage reduction. While many GC topologies have been investigated, the 2T design stands out for its compactness and retention efficiency. This chapter discussed the background and motivation behind GC memories, compared several topologies, and highlighted their features and uses.

1.7 Organization of The Thesis

This thesis aims to explore and optimize a 2T pMOS-only gain cell based on 28nm double gate pMOSFET, with a focus on enhancing DRT and ensuring reliable operation at low supply voltages. There are five chapters in this thesis. The description of which is outlined below.

The study starts with chapter 1, which discusses the evaluation of GC memories, various topologies such as 2T, 3T, 4T, and 5T, and their key features and applications.

The detailed working of the 2T0C gain cell and the importance of DRT is presented in chapter 2.

Chapter 3 presents an analysis of 2T0C GC using 28 nm double gate pMOSFET. This chapter highlights the importance of back bias to reduce capacitive coupling and subthreshold leakage to get enhanced DRT.

The effects of temperature and supply voltage downscaling on GC performance are reported in Chapter 4.

The findings of this work are summarized in chapter 5.

Chapter 2

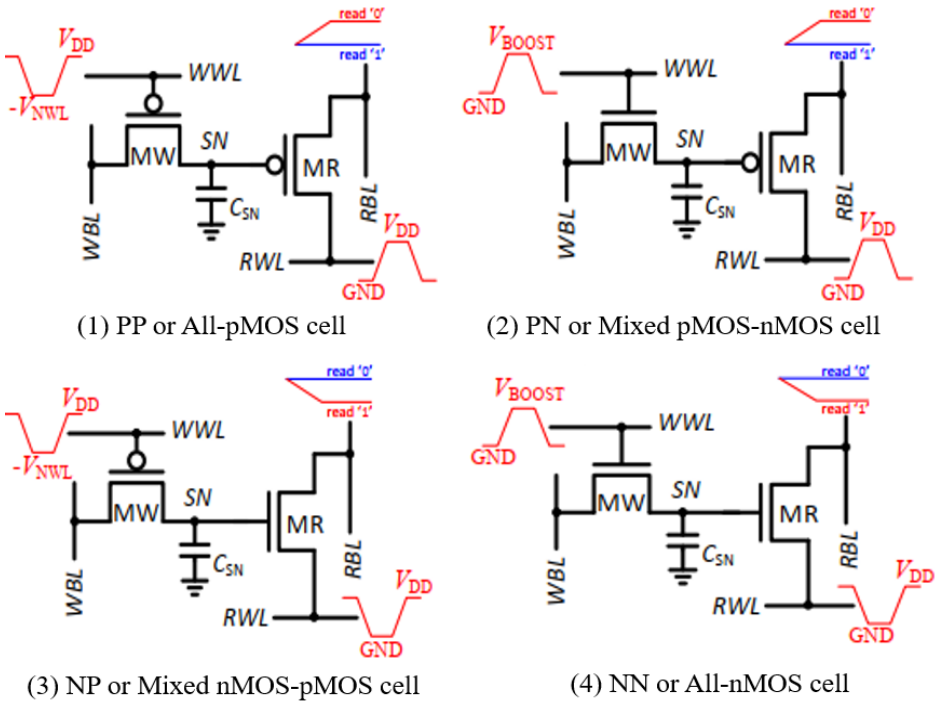
Operation of 2T0C DRAM

First chapter discussed GC as a promising solution for on-chip cache memory, especially as last-level cache in high-end processors [30]. This chapter will cover a detailed explanation of the operation of memory cell, explaining the fundamental write, hold, and read operations, as well as data storage using a parasitic node capacitor instead of a physical capacitor. Since its working depends on stored voltage levels at a floating node (SN), it is very important to understand charge leakage and device characteristics. This charge degradation at SN governs the retention time. The concept of data retention time (DRT), and the parameters influencing it, will also be discussed.

2.1 Working of Gain Cell

The fundamental idea behind any GC is to store digital information dynamically using charge on an internal node. This is achieved using two transistors, one for write access and another for read access. These two devices together form a memory element that stores data dynamically on SN. The working of GC happens in three operations: (1) writing data, (2) holding data, and (3) reading the stored data. Depending on the type of devices used (pMOS or nMOS) for these transistors, four 2T configurations are possible: (1) PP or All-pMOS cell where both MW and MR are pMOS transistors, (2) PN or Mixed pMOS-nMOS cell where MW is a pMOS-type transistor and MR is an nMOS-type transistor, (3) NP or Mixed nMOS-pMOS cell where MW is an nMOS transistor and MR is a pMOS transistor, and (4) NN or All-nMOS cell where both MW and MR are nMOS-type transistors (Fig. 2.1). Even though four configurations are different, all of these topologies follow a similar operating principle [29]: during write, MW

is turned ON, and it charges or discharges SN depending on the data available on WBL for writing; in the hold phase, both transistors are OFF, and the charge is expected to stay at SN until the next refresh; during the read phase, MR is turned ON, and the voltage at SN determines whether current will flow through MR to the read bit line (RBL), thus reading the stored value.



(1) PP or All-pMOS cell (2) PN or Mixed pMOS-nMOS cell

(3) NP or Mixed nMOS-pMOS cell (4) NN or All-nMOS cell

Fig. 2.1 Two transistor gain cell implementation options including the schematic waveforms [29].

2.1.1 Write Operation

In a 2T gain cell, writing a bit involves turning ON MW and transferring available WBL voltage (data) to the SN [45]. The WWL, which is the gate terminal of MW, is activated (set high for nMOS MW and set low for pMOS MW), which turns ON the MW transistor (Fig. 2.1). The MW now connects the WBL to SN and transfers either logic '1' (high) or logic '0' (low) depending on the value of WBL. MW either pulls the storage node SN toward V_{DD} (logic 1) or lets it fall closer to the ground (logic 0).

For a pMOS-only GC topology, a negative control bias is fed to the WWL node so that the gate to source potential (V_{GS}) becomes more negative than the threshold voltage of the MW, and the MW turns ON. However, due to the threshold voltage drop of the pMOS transistor, pMOS MW passes a weak ‘0’ and a strong ‘1’. This weak logic '0' is a vulnerability, as it will rise further during the hold phase due to leakage. This incomplete voltage swing affects retention time, which will be discussed in detail later.

2.1.2 Hold Operation

After the data has been written, WWL is de-asserted (turned low for nMOS and high for pMOS MW), which turns OFF the MW transistor, and the SN holds the stored voltage. Since there is no physical capacitor, the only element holding the charge is the parasitic capacitance at SN, which is made up of MW junction capacitance, MOSCAP, and any metal interconnects. Unfortunately, over time, leakage currents from the source/drain junctions through MW slowly destroy the charge. This process causes the voltage at SN to drift away from its original value - a phenomenon known as data degradation, and this degradation defines DRT. Factors like temperature and supply voltage downscaling directly impact this degradation, and that will be discussed later. The aim of this work is to minimize data degradation, i.e., maximizing DRT.

For a pMOS-only GC topology, the control bias is changed to positive so that the V_{GS} becomes less negative than the threshold voltage of the MW, and the MW turns OFF. This configuration is favorable because pMOS transistors leak less than nMOS, thanks to high V_{Th} and lower hole mobility (μ_p). Since pMOS passes weak '0', that is the reason for logic ‘0’ degrading much faster than logic ‘1’ in pMOS-only topology.

2.1.3 Read Operation

The read transistor allows data to be sensed by the sense amplifier (SA) from RBL, depending on the charge present on SN. To read out the data, we applied a drain to source (V_{DS}) potential across MR so that MR turns ON and current can flow. Keep in mind that during write and hold, we kept V_{DS} equal to zero so that no current would flow through the readout path even if MR was activated by the SN (gate of MR). Before reading, to keep V_{DS} of MR zero, both RWL and RBL are either pre-charged (for nMOS MR) or pre-discharged (for pMOS MR). As the nMOS passes a strong '0' and the pMOS passes a strong '1'. Before the read operation, RBL is pre-charged for nMOS (pre-discharged RBL for pMOS) to enable changes in RBL to be sensed during readout.

In the read phase, RWL is pulled low (for nMOS MR) or high (for pMOS MR), and V_{DS} becomes nonzero. For a pMOS-only GC topology, if SN holds a logic ‘0’, then the gate to source potential (V_{GS}) of MR becomes more negative than the threshold voltage, which turns ON MR, and current will flow from RWL to RBL [45]. This charges RBL, and thus read data as logic '0'. If SN holds a logic ‘1’, then the V_{GS} of MR is less negative than the threshold voltage, and MR remains OFF. No current will flow from RWL to RBL, RBL voltage will remain at zero, and thus we fetch data as logic '1'. The schematic waveforms describing GC working for different topologies are shown in Fig. 2.1.

2.2 Data Retention Time (DRT)

DRT is an important characteristic parameter of volatile memory because after this time period, memory needs refreshing. In GC, data is stored as electrical charge on a SN, which degrades over time. DRT is the time up to which the sense amplifier can read the data correctly [40].

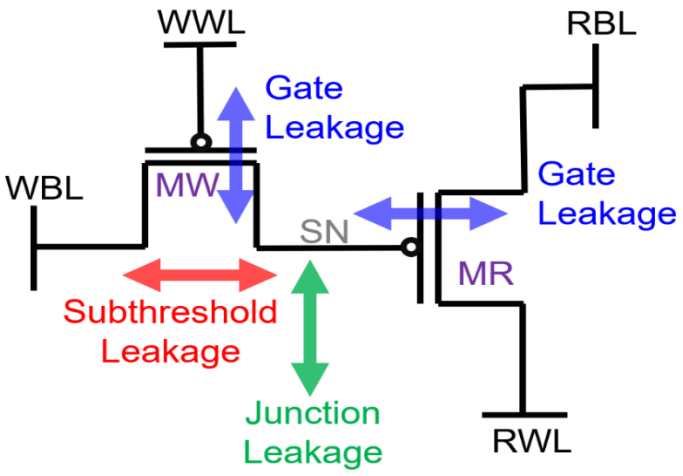


Fig. 2.2 Three types of leakages that can destroy stored data in a conventional 2T gain cell [17].

The primary reason for stored data (V_{SN}) degradation is leakage through MW and MR. Fig. 2.2 shows the structure of a conventional 2T eDRAM cell with its three types of leakages [17]:

(1) Subthreshold leakage: The dominant leakage mechanism that destroys the stored data the most. Even when the MW is OFF, some current penetrates through its channel. This mechanism weakens both logic '0' and logic '1'.

(2) Gate leakage: In advanced technology, thin gate oxides allow tunneling current to pass through the gate terminal. This leakage mechanism weakens one logic and strengthens another. To turn OFF the pMOS MW, we need a positive voltage at the gate terminal, which charges both logic '0' (weakens) and logic '1' (strengthens).

(3) Junction leakage: Reverse-biased diodes between drain/source and substrate leak small currents, especially at high temperatures, which also weakens one logic and strengthens another [29]. These leakage phenomena pull the voltage on SN away from its stored state, and if not refreshed in time, the stored value flips.

2.2.1 Importance of DRT

The importance of DRT is that power consumption is inversely proportional to DRT [50]. One might think that if DRT is low, then why not just refresh more frequently? The problem is refreshing consumes power, and in GC the total power consumption is dominated by the retention power due to periodic refreshing. A system/array with many cells, the required energy for refreshing every few microseconds will be substantial. Moreover, in low-power systems, where the system spends long idle periods in sleep or near-zero power modes, a short DRT would lead to unnecessary wakeups just to refresh memory, and the refresh energy may exceed the energy used to compute. Therefore, it is always better to improve DRT rather than refresh more often. Equation 2.1 shows that memory will consume less power if DRT is high [50]. Hence, DRT is a critical design parameter that must be understood and optimized.

$$P_{retention} = P_{leakage} + P_{refresh} = V_{DD} I_{leak} + \frac{E_{refresh}}{t_{refresh}} \qquad (2.1)$$

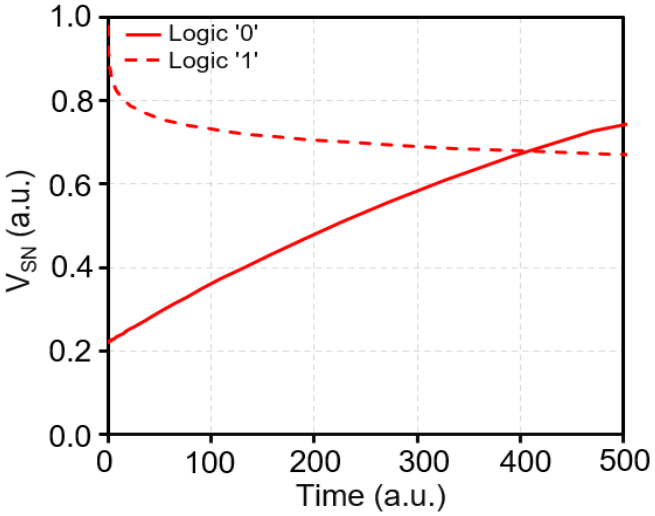


Fig. 2.3 Storage node voltage (V_{SN}) degradation of 2T GC after write operation.

2.2.2 Method to estimate DRT

There are two popular methods [10], [27] to estimate the data retention time of a gain cell. One is the threshold voltage-based method, and the other is the 200 mV differential method. Fig. 2.3 shows the degradation of the stored information (V_{SN}) over time. As lower voltage is considered logic '0' and higher voltage is taken as logic '1', so after a write operation, the storage node voltage for logic '0' increases and for logic '1' decreases over time.

2.2.2.1 Method 1 — Threshold Voltage of Read Transistor (MR)

This method [27] is based on determining the rate of decay of the storage node before a read error occurs. In this method, DRT is defined as the time up to which the voltage on the SN remains below (for logic ‘0’) or above (for logic ‘1’) the threshold voltage (V_{Th}) of the read transistor MR.

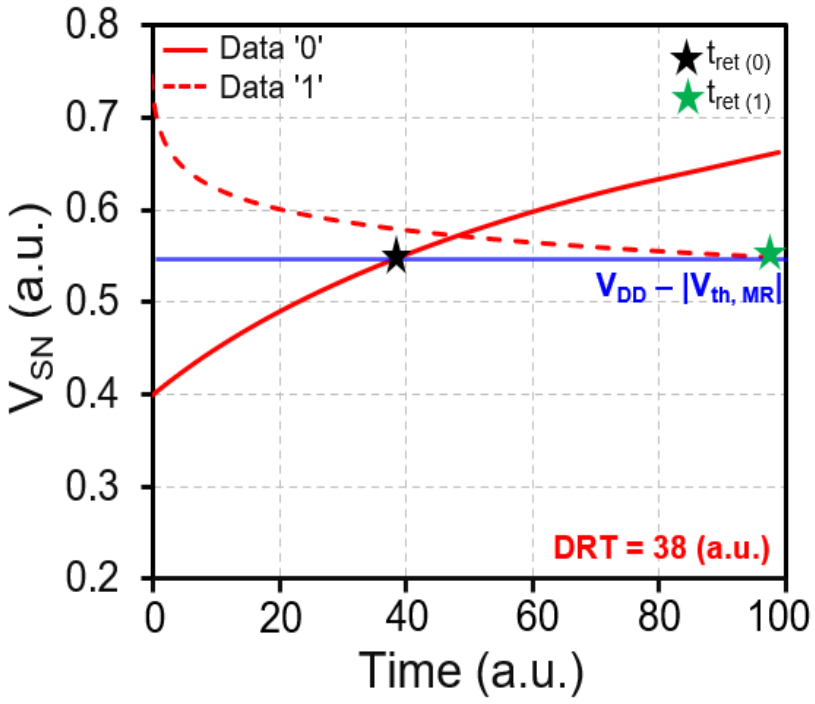


Fig. 2.4 Storage node voltage (V_{SN}) degradation of an arbitrary pMOS-only GC after write operation and DRT estimation using 1st method [27].

During read, the current flowing through MR depends on the V_{GS}) and V_{Th}. For one stored logic (depending on the type of MR), at the gate of MR will activate the read transistor ON, which results in a current flow through MR, and for other logic, MR will remain OFF and no current will follow. In this way, SA can distinguish between logic '0' and logic '1', and reading occurs. As an example, a pMOS read transistor will follow current as long as stored node voltage ($V_{GS,\ MR}$) remains below (logic '0') its threshold value. As leakage changes SN voltage, so logic ‘0’ drifts upward, and once it crosses the critical point (V_{DD} - $|V_{Th,\ MR}|$), the read transistor can’t distinguish logic ‘0’ from a logic ‘1’ anymore, and data will be lost. This method estimated DRT for both logic '0' and logic '1' separately, and the overall DRT is taken as the minimum of that. Fig. 2.4 shows the V_{SN} degradation after the write operation and presents the estimated DRT value according to this method. This approach is accurate and straightforward for simulation because it reflects actual read failure conditions, but there is a limitation to this method. If, just after writing, V_{SN} for both logic lies either below or above the critical value (V_{DD} - $|V_{Th,\ MR}|$), then MR cannot distinguish logic '0' and logic '1' levels. That is why alternate method for estimates DRT is required. In this method, DRT is the time when the voltage difference between two logic levels reaches a value of 200 mV [10].

2.2.2.2 Method 2 — 200 mV Differential Method

In this simpler approach, DRT is evaluated by tracking how long the voltage difference between logic ‘1’ and ‘0’ stays above 200 mV [10]. It is assumed that data is still readable as long as there is a voltage difference of at least 200 mV between logic ‘0’ and logic ‘1’ on SN. This strategy is based on the assumption that modern sense amplifiers can reliably detect a potential difference of 200 mV.

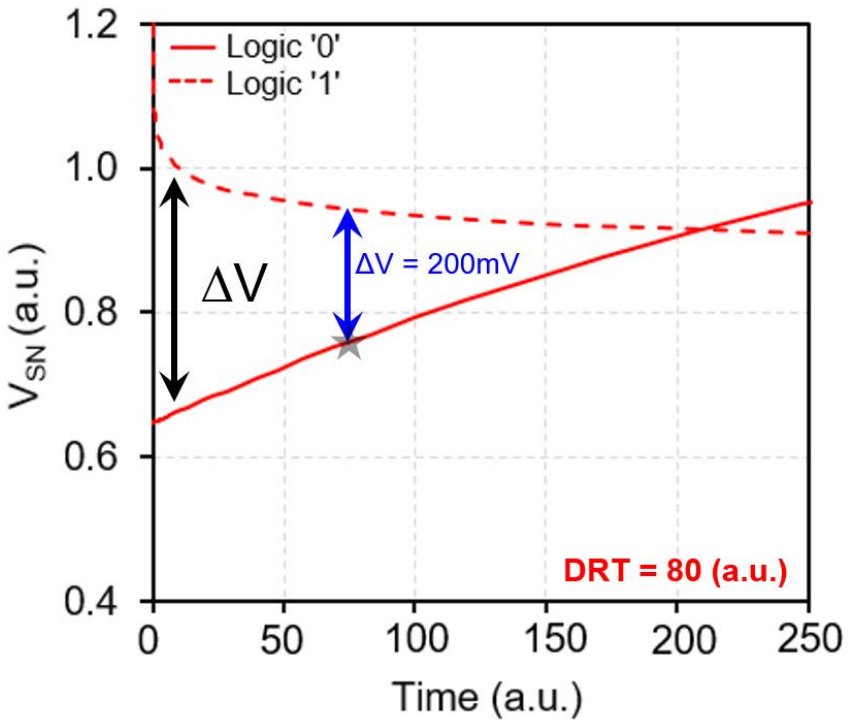


Fig. 2.5 Storage node voltage (V_{SN}) degradation of an arbitrary pMOS-only GC after write operation and DRT estimation using 2nd method [10].

Fig. 2.5 illustrates the degradation of the stored information over time and presents the estimated DRT value according to this method. At the onset of the hold operation, we got a voltage difference (ΔV) between logic '1' and logic '0'. As lower voltage is considered logic '0' and higher voltage is taken as logic '1', hence after a write operation, the storage node voltage for logic '0' increases and for logic '1' decreases over time, and ΔV decreases over time. DRT estimated as the moment when the voltage difference between two logic levels reaches a value of 200 mV. In an ideal scenario this difference should be 0 V to distinguish between logic '0' and logic '1', but in practice this 200 mV gives a sufficient margin to the sense amplifier to distinguish between '0' and '1' even if there is any fluctuation. This method is more conservative and gives safer design margins, which is why it's often used in several reports [20], [44], [46]. In this work, this method is used to estimate DRT.

2.2.3 Improving DRT

Improving DRT is better approach to minimize frequent refresh. In a dynamic memory cell like the 2T gain cell, the primary cause of data degradation is leakage current, particularly the subthreshold leakage through MW. This current continuously drains the stored voltage on the floating storage node during the hold state. Therefore, the most direct and effective strategy for enhancing DRT is to suppress leakage as much as possible. This includes using transistors with higher threshold voltages [29] (such as pMOS) and applying reverse body biasing [51] to control V_{Th} such that MW enters into the hard-off state during hold.

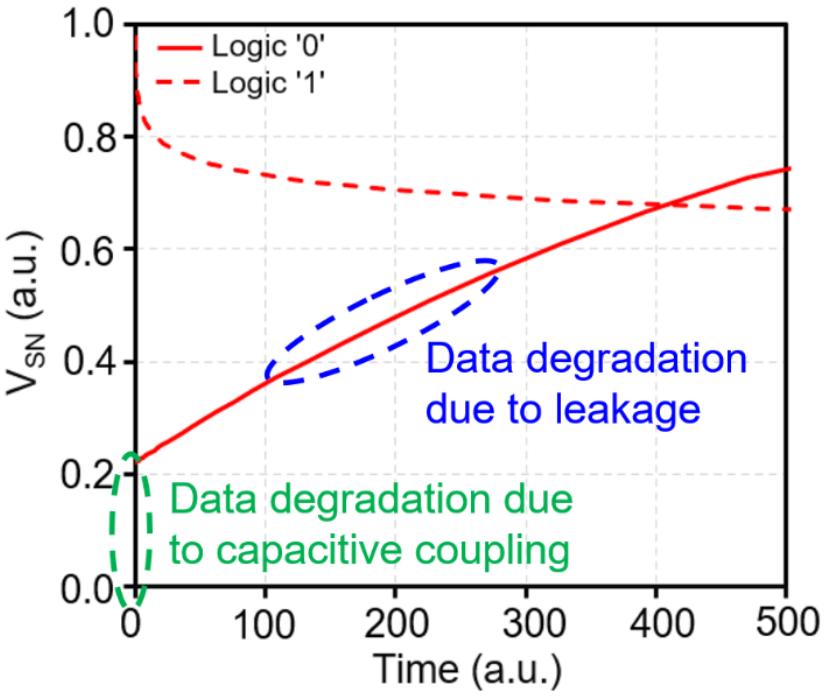


Fig. 2.6 Impact of leakage and capacitive coupling on storage node voltage (V_{SN}) degradation.

Fig. 2.6 illustrates the degradation of the stored information over time due to leakage. Along with this continuous degradation, there is a sudden degradation of data at the onset of the hold operation, and this degradation is because of capacitive coupling (CC).

Beyond leakage suppression, there are also architectural solutions that can enhance DRT. Gain cell topologies [14], [19], [39] such as 3T0C, 4T0C, and 5T0C include extra transistors to improve data stability. These designs improve write/read isolation and enhance retention capabilities. However, these improvements come at the cost of increased area and peripheral circuit complexity, which makes them less suitable for dense arrays or compact on-chip memory macros where space is vital [47]. Fig. 2.7 shows the reported DRT values for different architectures.

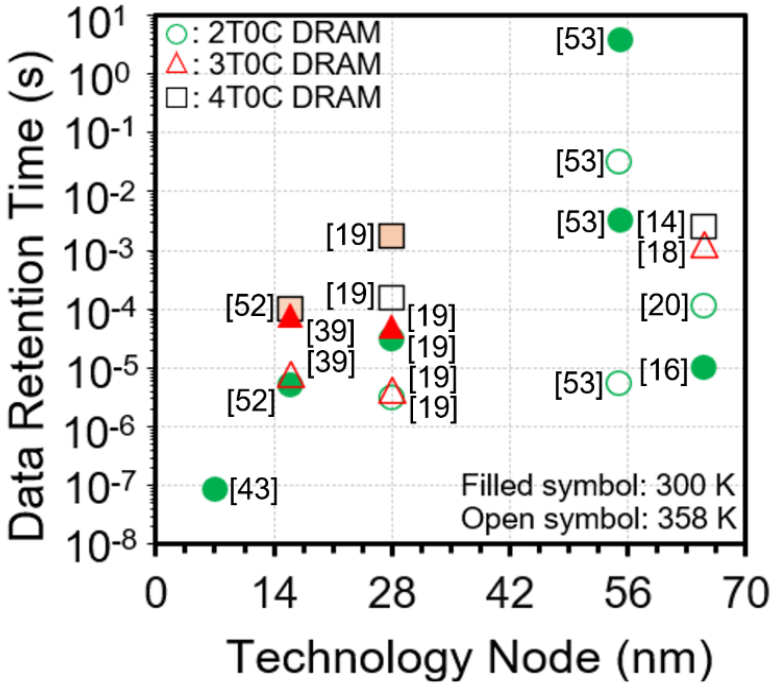


Fig. 2.7 Reported DRT values [14], [16], [18], [19], [20], [39], [43], [52], [53] of Si GC-eDRAM at different technology nodes.

To address this trade-off, the focus of this work remains on the 2T topology, which is the most compact among gain cell structures. Instead of adding extra transistors, this work explores the use of double gate (DG) pMOSFETs to improve DRT. DG devices offer better electrostatic control over the channel and significantly reduce subthreshold leakage and minimize CC effects at the storage node. These properties make them perfect for extending data retention without introducing any area overhead.

2.3 Conclusion

The functional behavior of the 2T gain cell memory has been thoroughly covered in this chapter, with a particular focus on the pMOS-only configuration. Study started with a general overview of how the gain cell operates, elaborating it into its three main states - write, hold, and read. This chapter explored how the stored charge is vulnerable to various leakage mechanisms. This chapter also had a focus on understanding data retention time, a critical metric for dynamic memory cells. Two commonly used methods for evaluating data retention time, based on the threshold voltage of the read transistor and the 200 mV voltage margin rule, were presented in detail. The understanding gained from this chapter prepares one for the next stage of this work.

Chapter 3

Double Gate based 2T pMOS-only GC

3.1 Double Gate (DG) p-type MOSFET

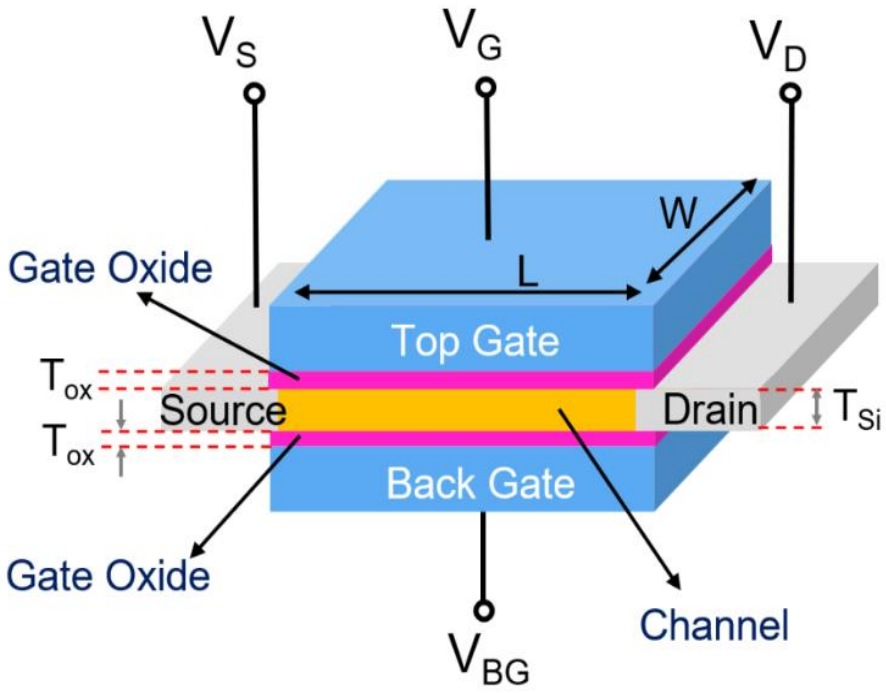


Fig. 3.1 Schematic diagram of a DG-MOSFET.

The increasing need for high-density and low-power digital systems is driving CMOS transistors to continue scaling into the nanoscale regime [54], [55]. However, as CMOS technology scales down beyond 28nm, controlling leakage currents and improving electrostatic integrity become more challenging due to the occurrence of short channel effects (SCEs) [56]. Traditional bulk MOSFETs have higher leakage and threshold voltage variability due to their poor channel control and increasing SCEs. To overcome this issue, a multi-gate transistor, specifically a DG-MOSFET

architecture is used in this work for the design of a 2T gain cell. Compared to traditional planar CMOS technology, the DG concept can significantly improve SCE mitigation at shorter channel lengths [57].

Fig. 3.1 shows the schematic of a double gate MOSFET, where the source and drain are connected at either end of the channel. Unlike single-gate MOSFETs (SG-MOSFETs), where the gate controls the channel from one side, DG MOSFETs employ two gates, a top gate (also known as the front gate) and a back gate (also known as the bottom gate), to control the channel simultaneously from both sides. V_G and V_{BG} are the control voltages at the top gate and at the back gate, respectively. V_S and V_D are the voltages at the source and the drain terminals.

Table 3.1 Dimensions of DG-pMOS used in this work.

W (nm)	**L (nm)**	**T_{OX} (nm)**	**T_{Si} (nm)**
1000	28	1.3	7.2

Table 3.1 shows the dimensions of DG-pMOS used in this work, where W is width, L is the gate length, T_{OX} is oxide thickness of SiO_2 layer, and T_{Si} is the silicon thickness of the transistor.

The utilization of double gates effectively enhances the coupling between the gate and channel, thereby facilitating the suppression of SCEs [57]. Furthermore, dual gate architecture creates a symmetric electrostatic field on the channel, which suppresses unwanted influence from the drain and source. As a result, DG-MOSFET exhibits reduced off-current, contributing to overall enhanced performance and efficiency in circuit design [58]. DG-MOSFETs offers higher ON current and reduced OFF current, enabling operation at lower supply voltage [59]. Operating at threshold voltages also reduces the power consumption. Thus, DG-

MOSFETs can improve DRT in gain cell applications by reducing SCEs and the off-state leakage current.

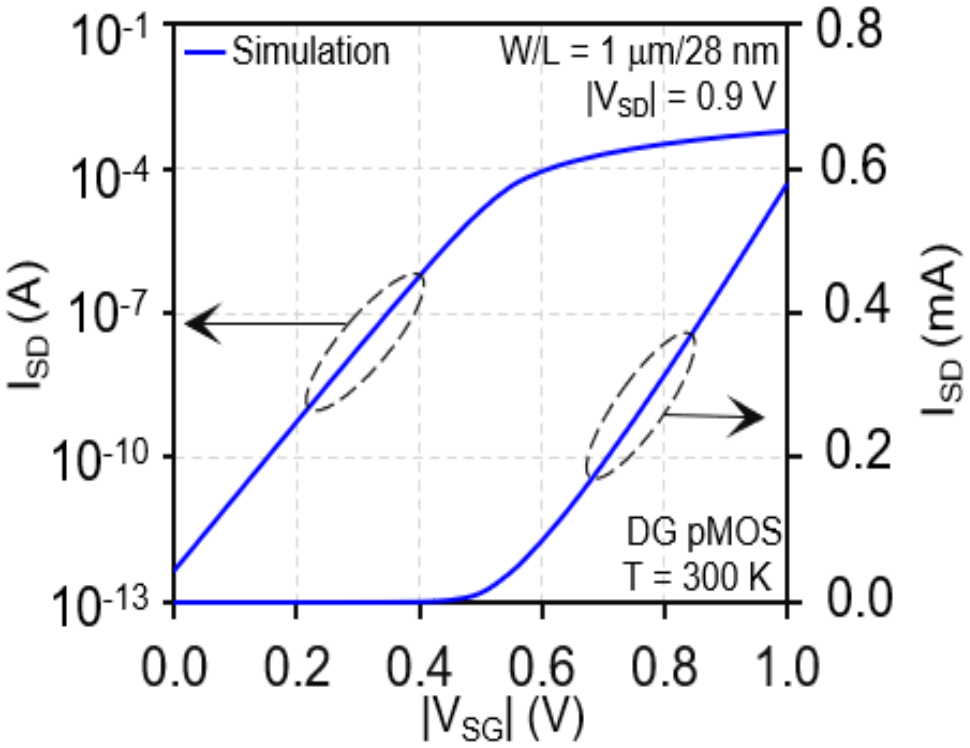


Fig. 3.2 Transfer characteristics of DG-pMOS.

As illustrated in Fig. 3.2, simulated transfer characteristics (I_{SD} vs V_{SG}) of DG-pMOS transistor. The simulation [60] included field, concentration and temperature dependent carrier mobility besides module for generation-recombination and quantum confinement effects. Temperature dependent carrier lifetime was also considered.

Two important parameters for transistor are on-current (I_{ON}) and off-current (I_{OFF}). I_{OFF} is a source to drain current (I_{SD}) when source to gate voltage (V_{SG}) at 0 V and source to drain voltage (V_{SD}) at V_{DD}. In this work, V_{DD} is equal to 1 V. I_{ON} is a source to drain current (I_{SD}) when source to gate voltage (V_{SG}) and source to drain voltage (V_{SD}) at V_{DD}. Mathematically, the same can be expressed as

$$I_{OFF} = I_{SD} \text{ @ } (V_{SG} = 0\ V\ and\ V_{SD} = V_{DD} \quad (3.1)$$

$$I_{ON} = I_{SD} \text{ @ } (V_{SG} = V_{SD} = V_{DD} \quad (3.2)$$

The most significant advantage shown in the above plot (Fig. 3.2) is that the DG-pMOS device achieves almost twice the I_{ON} compared to the SG FDSOI [61]. This means that the DG transistor will finish the transition more quickly during a write operation when a logic value needs to be stored on the floating storage node of the gain cell. This leads to a significant improvement in the write speed capability of the device. Furthermore, the I_{OFF} in DG-pMOS is decreased by nearly two orders of magnitude in comparison to SG FDSOI [61]. This reduction is extremely helpful during the hold phase, where subthreshold leakage has to be minimized to maintain the voltage level on the SN. The main aim of this work is improving data retention time, which increases with the duration of the SN able to preserve its value. The next sections will investigate the utilization of this DG-pMOS device in 2T pMOS-only gain cell design and way of enhancing DRT through effective biasing.

3.2 Double Gate 2T pMOS-only Gain Cell

3.2.1 Introduction to DG 2T pMOS-only GC

This section presents the proposed gain cell (GC) architecture which utilizes the advantages of a double gate (DG) 2T pMOS-only configuration. The cell is designed to reduce leakage, suppress capacitive coupling, and extend DRT without significantly increasing the area beyond the classical 2T footprint.

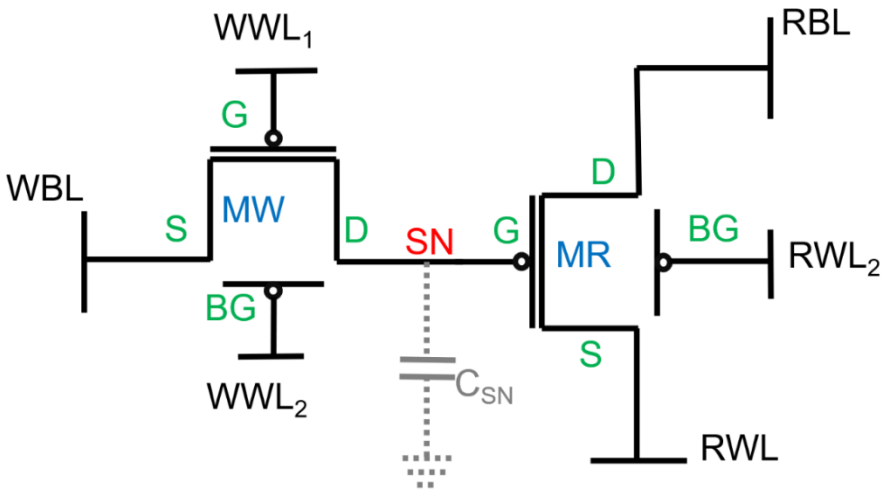


Fig. 3.3 Schematic representation of double gate 2T pMOS only gain cell.

Fig. 3.3 shows the suggested architectural schematic of DG 2T pMOS-only gain cell. The fundamental design is still the same as that of a traditional 2T gain cell. It consists of a MW and a MR, with the SN in between. Both transistors are implemented using double gate pMOS and connected such that the front and back gates of every device can be independently biased to optimize leakage control and capacitive behavior. The MW on the left is controlled by WWL_1 (front gate) and WWL_2 (back gate). The front gate of MR is connected with SN. During read, MR is controlled by RWL (source) and RWL_2 (back gate). A single SN connecting MW and MR, storing logic '1' or '0' depending on the last write operation.

The WBL provides data for writing, and from the read bit line RBL the data can be fetched. This cell is fully compatible with standard logic CMOS and offers scalable leakage control using the double gate structure. The use of pMOS-only transistors reduces gate and subthreshold leakage.

3.2.2 Working of DG 2T pMOS-only GC

3.2.2.1 Write Operation

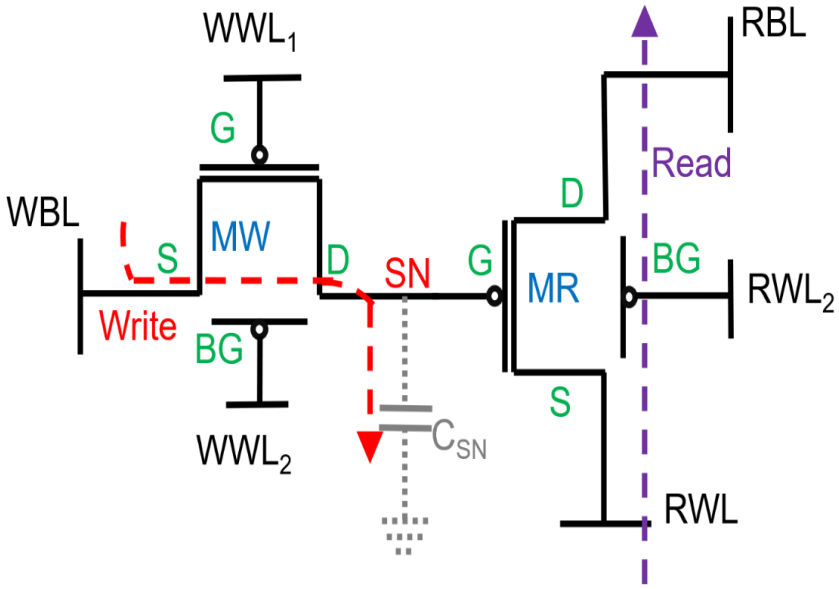


Fig. 3.4 Schematic representation of DG 2T pMOS-only GC showing write and read paths.

Fig. 3.4 illustrates schematic with write and read paths of DG 2T pMOS-only GC. In a 2T gain cell, writing a bit involves turning ON the MW and transferring available write bit line voltage (data) to the SN [45].

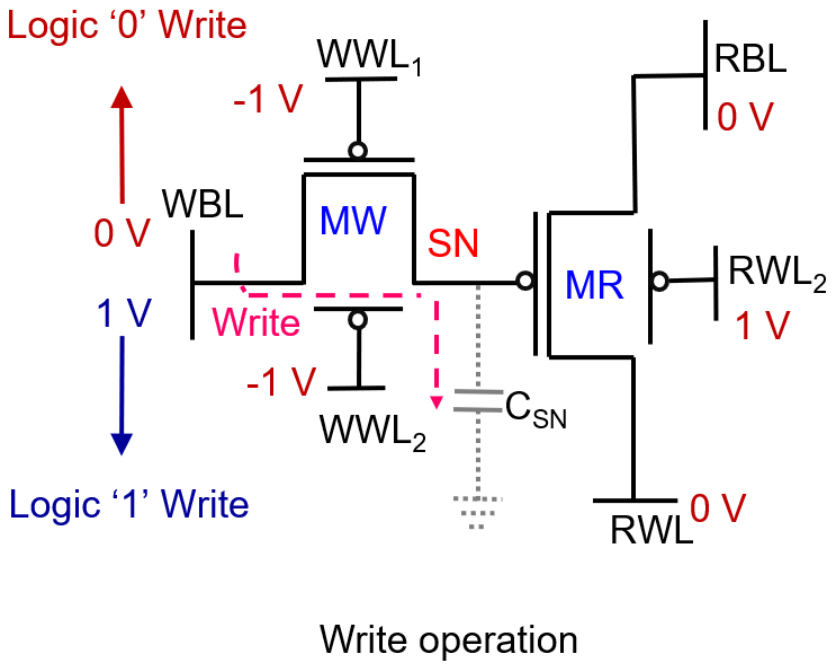


Fig. 3.5 Bias representation of DG 2T pMOS-only GC during write operation.

In DG 2T pMOS-only GC, the write operation is enabled by activating both WWL_1 and WWL_2 (pulling them LOW for pMOS) to turn ON the MW. To write logic '1', WBL is pulled to V_{DD} (= 1 V), and SN is pulled toward the V_{DD} through the conducting MW. Similarly, to write logic '0', WBL is grounded to 0 V, and SN is pulled down to a level near ground. As shown in Fig. 3.5, WWL_1 and WWL_2 are pulled to -1 V to ensure strong conduction through MW. This helps write the data very fast. Meanwhile, the MR remains OFF since RWL_2 is held high and also zero potential difference between RWL and RBL ($V_{SD,\ MR}$) makes sure no current will flow through MR.

3.2.2.2 Hold Operation

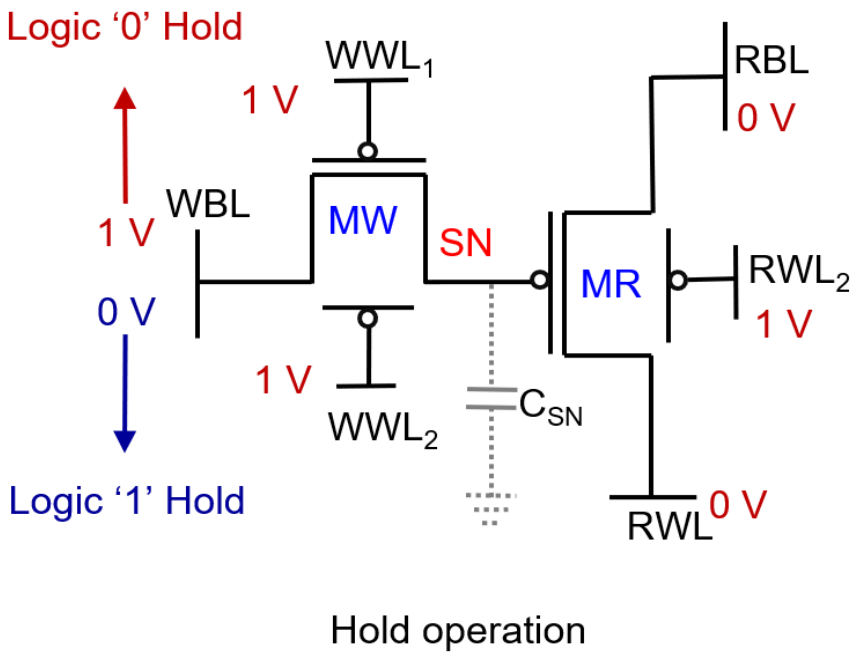


Fig. 3.6 Bias representation of DG 2T pMOS-only GC during hold operation.

Once the data is written to the SN, the write word lines WWL_1 and WWL_2 are deactivated (set to 1 V, which is high for pMOS and thus turns MW OFF). At the same time, RWL_2 is also high to turn OFF MR (Fig. 3.6). This isolates the SN completely, allowing it to retain its charge on the storage node capacitance C_{SN}.

This mode is highly sensitive to leakage and capacitive noise. Here the pMOS-only configuration becomes favorable because pMOS transistors leak less than nMOS due to their higher V_{Th} and lower μ_p. Furthermore, due to the double gate pMOS structure, an independent bias at WWL_2 can further increase the V_{Th} of the MW transistor and suppress leakage.

3.2.2.3 Read Operation

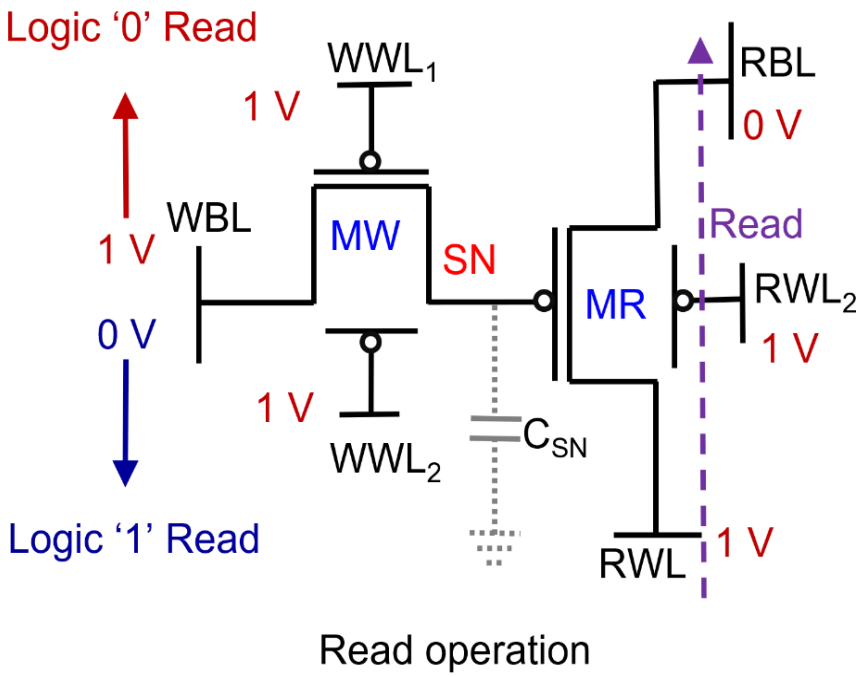


Fig. 3.7 Bias representation of DG 2T pMOS-only GC during read operation.

As shown in Fig. 3.7, to perform a read, a drain to source potential is applied across MR in order for current flow, and at this time WWL_1 and WWL_2 remain high (MW is OFF). The read word line (RWL) pulled HIGH (1 V), keeping RBL at 0 V. If SN holds a logic '0' (0 V), then the V_{GS} of MR becomes more negative than the threshold voltage, which turns ON MR, and current (I_{DS}) will flow from RWL to RBL. This charges RBL, and thus read data as logic '0'. If SN holds a logic '1' (1 V), then the V_{GS} of MR remains less negative than the threshold voltage, and MR stays OFF. No current will flow from RWL to RBL, RBL voltage will remain at zero, and thus data read as logic '1'. This operation ensures no disturbance to SN during read because of two port compatibility.

3.2.2.4 Timing Diagram

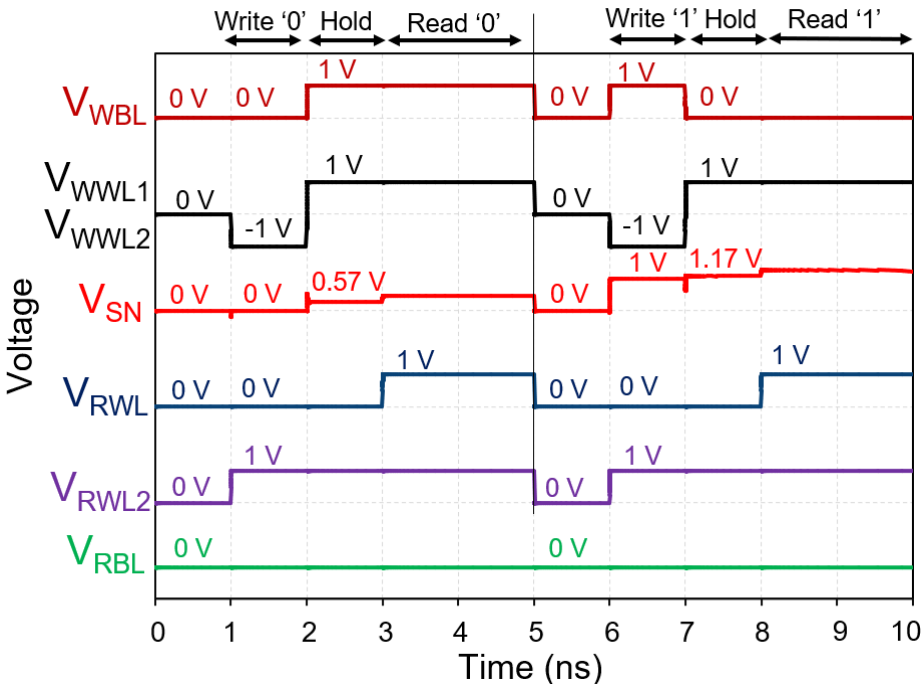


Fig. 3.8 Timing diagrams demonstrating DG 2T pMOS-only GC operations.

The timing diagram in Fig.3.8 displays voltages over time for WBL, WWL_1, WWL_2, SN, RWL, RWL_2, and RBL during a complete Write → Hold → Read sequences for both logic '0' and logic '1'. During the write operation, V_{WWL1} and V_{WWL2} = -1 V turn ON MW, and SN is driven to 0 V for logic '0' (V_{WBL} = 0 V) and to 1 V for logic '1' (V_{WBL} = 1 V). In the hold operation, V_{WWL1} and V_{WWL2} = 1 V turn OFF MW, and SN is expected to remain data. SN drifts the data over time because of leakage and capacitive coupling. One can notice that although 0 V is written for logic '0', but at the onset of hold, SN voltage instantly degraded to a higher value such as 0.57 V. This is because of capacitive coupling. The detailed analysis has been present in the next section. For read '0' operation, MR turns ON and RBL charges towards 1 V.

3.2.3 Capacitive Coupling (CC)

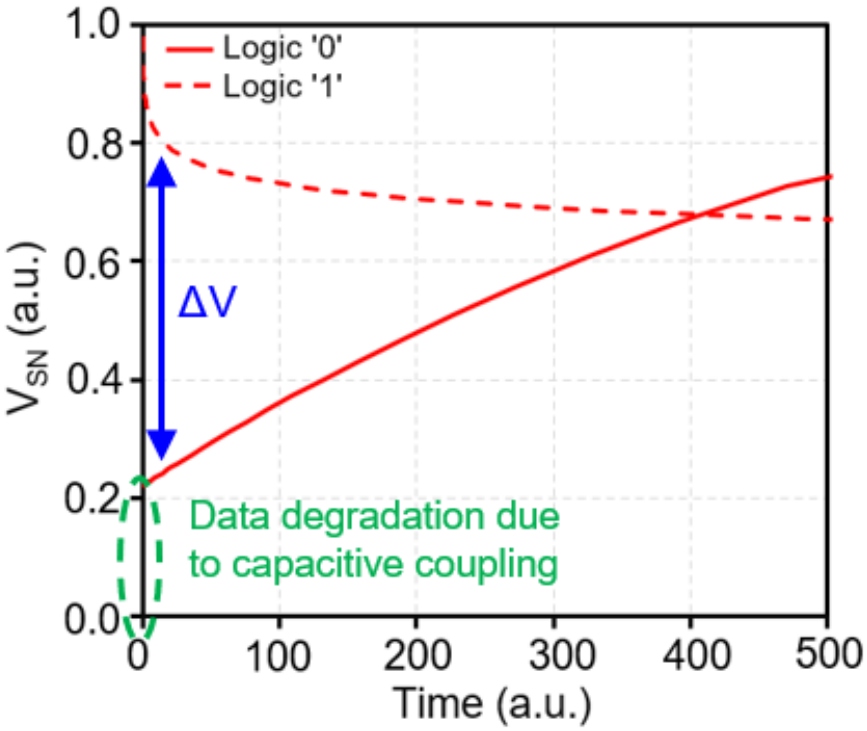


Fig. 3.9 Impact of capacitive coupling (CC) on storage node voltage (V_{SN}) at the beginning of hold state.

In nanoscale CMOS circuits, capacitive coupling is an important phenomenon that refers to a voltage change in one node affecting the voltage of an adjacent node through a parasitic capacitor [62]. The impact of CC is most prominent right after a write operation, at the beginning of the hold phase. The Fig. 3.9 shows logic '0' on SN has a sharp voltage level jump. This degradation is not gradual and cannot be assigned to leakage currents. Instead, it is the result of rapid gate voltage transitions, particularly when WWL or RWL switches states. This sudden jump compromises the voltage difference (ΔV) between logic levels and reduces DRT. Thus, suppressing CC becomes an essential strategy for enhancing retention.

Capacitive coupling refers to the unintentional transfer of electrical energy between two nodes without having any direct electrical connection between them [63]. This transfer occurs due to the presence of parasitic capacitance between the nodes, where a voltage transition at one node induces a voltage shift at another node. The coupled voltage can be approximated by ohm's law-based equation 3.3, which represents larger bias changes over shorter timescales ($\frac{dV}{dt}$) induce higher capacitive coupling.

$$V_{coupled} \cong (C\frac{dV}{dt})R \tag{3.3}$$

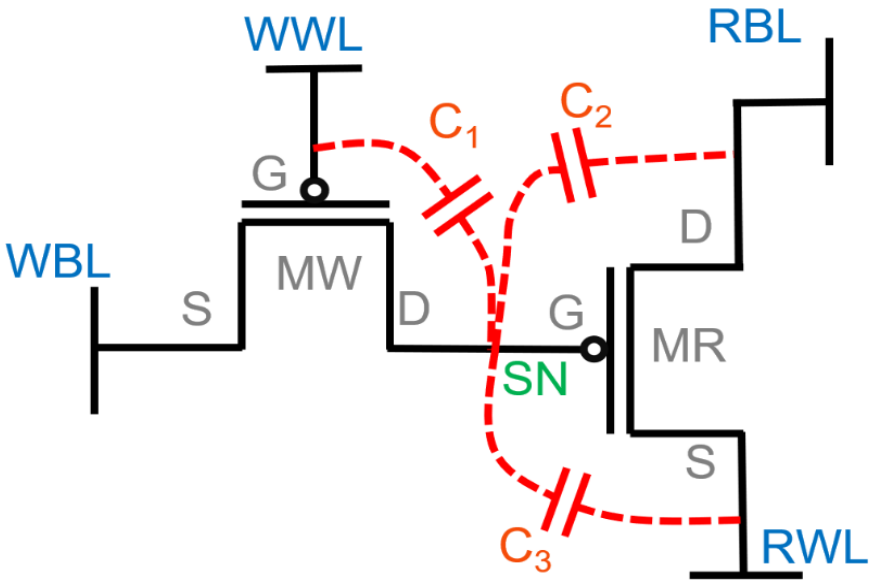


Fig. 3.10 Schematic of a 2T GC with the coupling capacitances between SN, WWL, RBL, and RWL [62].

MOSFET contains many parasitic capacitances, but here the schematic highlights 3 major parasitic capacitances that are directly connected to the storage node and other control nodes (Fig. 3.10). In gain cells, the SN is surrounded by several control and data lines, such as WWL, RWL, and WBL — all of which are capable of coupling the charge on the SN during transitions. The effect of CC is that if a node voltage changes rapidly, then it induces a change in the storage node voltage [63].

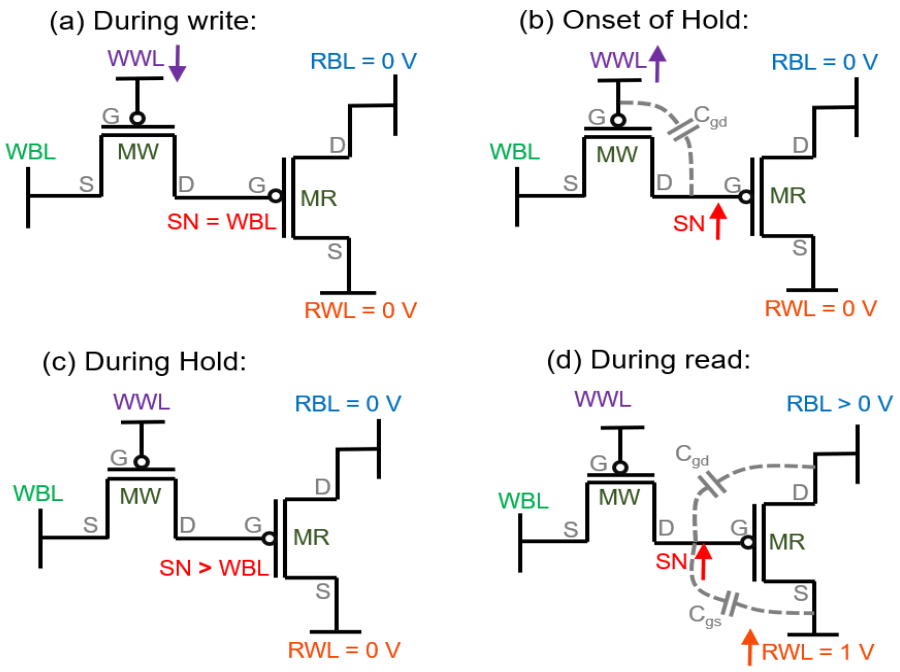


Fig. 3.11 Impact of capacitive coupling on storage node voltage during transitions of control node biases [63].

The Fig. 3.11 above illustrates the effect of CC on SN voltage throughout memory operation. During write, a lower WWL voltage activates the pMOS MW, and SN voltage is driven to V_{WBL}. At the onset of the hold operation, when we increased the WWL voltage to turn OFF MW, then some voltage is induced through gate to drain capacitance (C_{gd}) into the storage node. When the RWL bias is changed to read the data, it again coupled an unwanted shift on storage node voltage [63]. This phenomenon is captured by the following mathematical models [62], which conclude a larger change in control node biases (such as ΔV_{WWL} and ΔV_{RWL}) results in a larger change in storage node voltage (ΔV_{SN}).

$$\Delta V_{SN} \cong \frac{C1}{C1+C2+C3} \Delta V_{WWL} \tag{3.4}$$

$$\Delta V_{SN} \cong \frac{C3}{C1+C2+C3} \Delta V_{RWL} \tag{3.5}$$

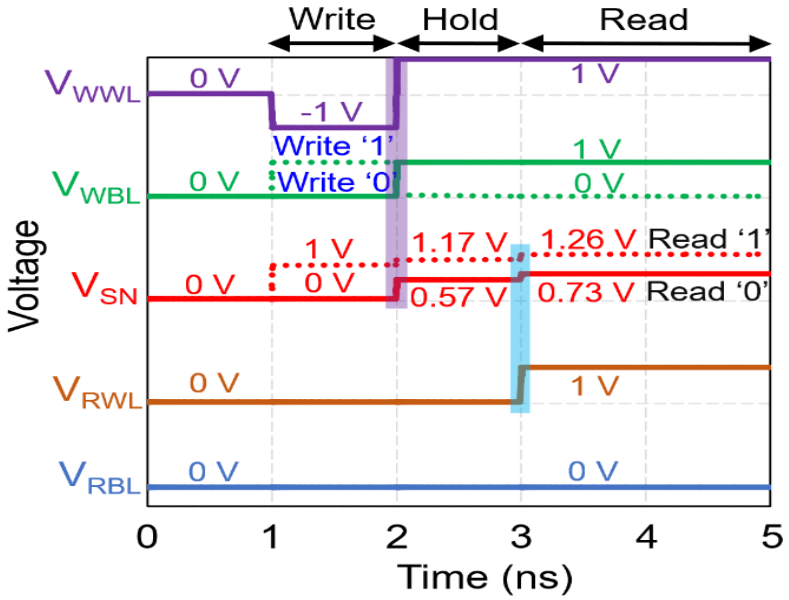


Fig. 3.12 Timing diagrams demonstrating capacitive coupling in 2T GC.

The impact of CC on SN is visualised in timing diagrams during memory operation [55]. Fig. 3.12 presents a timing chart showing the changes in SN voltage when transitions happen on various control nodes during write, hold, and read phases. During the write operation, 0 V for logic '0' and 1 V for logic '0' were successfully written, and after that at the beginning of hold operation the value increased due to CC. The violet highlighter represents CC due to WWL voltage changing, and the blue one represents CC due to change in RWL voltage. At the beginning of the hold, WWL is switched from -1 V to +1 V. This 2V change couples charge onto SN, causing it to jump from 0V to 0.57 V for logic '0' and from 1 V to 1.17 V for logic '1'. Again, due to a 1 V change in RWL at the beginning of the read operation, SN shifts to 0.73 V for logic '0' and to 1.26 V for logic '1'.

In summary, capacitive coupling is a critical challenge for maintaining robust DRT in 2T gain cells. It causes immediate SN disturbance at every transition point, particularly at the start of hold and activation of read. As nodes scale, parasitic capacitances become more prominent due to reduced dimensions, making CC effects more pronounced when technology is scaled.

3.3 Approach to Enhance DRT

Enhancing data retention time in a 2T gain cell design requires a specific understanding of how charge is lost from the storage node. As discussed in earlier chapters, DRT essentially depends on the leakage through MW, along with the parasitic capacitive coupling effects that disturb the stored voltage.

3.3.1 Motivation

The motivation behind this work arises from a fundamental question — Can DRT of a 2T gain cell be further improved, especially in advanced technology nodes like 28 nm? Although several device topologies [14], [39], including reverse body biasing [64] schemes and hybrid architectures [19], have tried to address DRT degradation, the maximum reported DRT at 28 nm remains limited to roughly 32 µs at 300 K and 3.2 µs at 358 K.

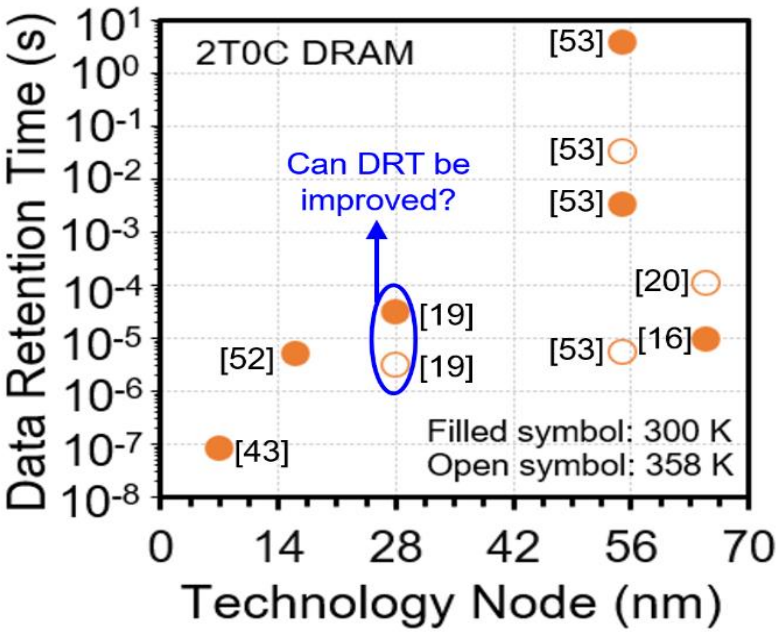


Fig. 3.13 Reported DRT values [16], [19], [20], [43], [52], [53] of 2T GC eDRAM at different technology nodes.

As shown in Fig. 3.13, which plots reported DRT values across technology nodes, DRT improves slightly with larger nodes due to reduced

leakage (as in 55 nm and 65 nm implementations). However, to achieve compact memory, it is important to optimise designs within more advanced technology, such as the 28 nm node. In this context, the objective of this thesis is to use a double gate (DG) pMOS-based 2T gain cell to achieve higher DRT. By using independent gate biasing, the DG structure can significantly reduce leakage current which can enhance DRT.

3.3.2 Two-Step Approach to Enhance DRT

Enhancing DRT requires minimizing the degradation of the voltage stored on the storage node. In 2T gain cell, this degradation is mainly caused by two effects:

(1) Capacitive coupling from word lines, especially during switching events.

(2) Subthreshold leakage current through the write transistor during the hold operation. The following section describes a two-step approach to address both capacitive coupling and leakage.

3.3.2.1 Step 1 — Maximize the voltage difference (ΔV)

The Fig. 3.14 (a) shows the typical degradation profile of SN voltage in a 2T gain cell without any optimisation. Here, both logic '1' and logic '0' begin at their respective post-write levels, but the voltage difference (ΔV) between the two logic levels is not the same as pre-write levels. The second chapter already covered how capacitive coupling and leakage cause this voltage difference to decrease over time, and DRT can be estimated as the time when this voltage window reaches a level equal to 200 mV (method 2). In order to increase the time required to reach a level of 200 mV, our first step is to maximize this ΔV level (Fig. 3.14 (b)). This can be achieved by reducing capacitive coupling. When CC is suppressed, the initial data degradation (ΔV_{SN}) is also reduced, which results in maximizing the ΔV.

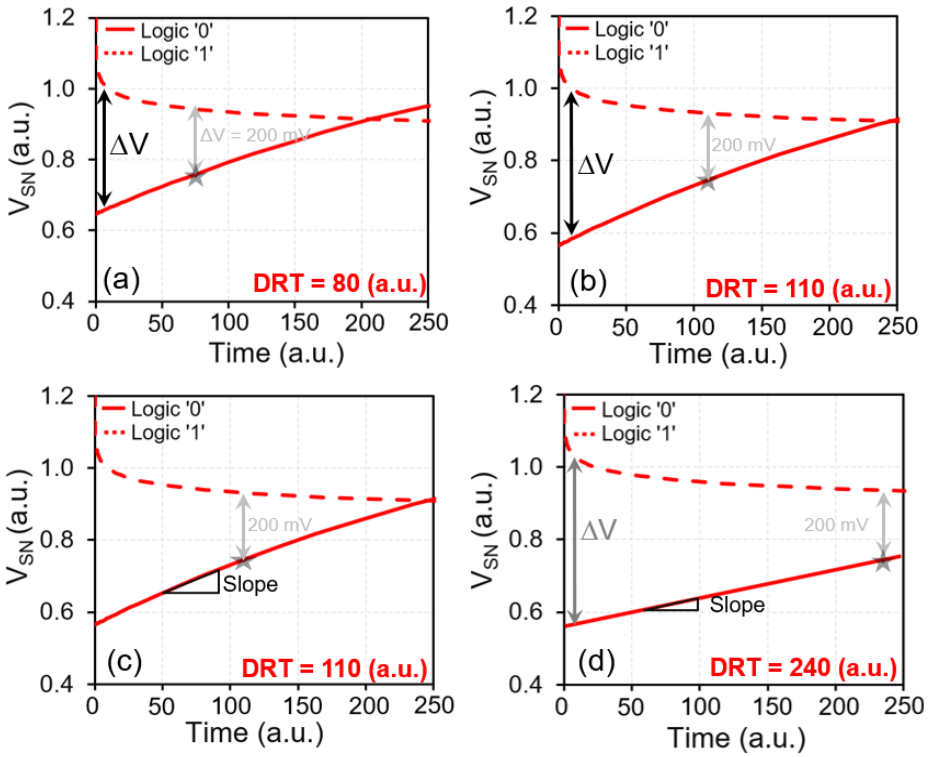


Fig. 3.14 Approach to enhance DRT — (a) a typical V_{SN} degradation with lower voltage difference (ΔV), (b) V_{SN} degradation with increased voltage difference (ΔV), (c) a pre-optimized V_{SN} degradation with higher slope, and (d) an optimized V_{SN} degradation plot having reduced degradation slope along with increased voltage difference (ΔV).

3.3.2.2 Step 2 — Lowering the data degradation slope (dV_{SN}/dt)

The second step is minimizing the data degradation slope along with maximizing the ΔV, so that it reduces the rate at which ΔV decreases. When the degradation rate of the ΔV is kept to a minimum, it will take more time to reach a level of 200 mV, which significantly enhances DRT. This step can be visualized in Fig. 3.14 (c) and Fig. 3.14 (d).

This two-step approach, first reducing capacitive coupling to maximize the voltage difference between two logic levels and second suppressing leakage to slow down voltage drift, forms the foundation of enhancing DRT in the proposed DG 2T pMOS Gain Cell.

3.4 Enhancing DRT – Bias optimization

The previous section introduced a two-step strategy for improving data retention time (DRT):

(1) Reducing capacitive coupling (CC).

(2) Suppressing leakage current.

This section will look at how adjusting the voltages applied to various nodes, particularly the write word lines (WWL_1, WWL_2) and write bit line (WBL), impacts our two-step strategy and thus influences overall DRT.

3.4.1 Optimizing WWL node bias

The first step is reducing CC, and that can be carried out by carefully optimizing write word line node bias. The write word lines WWL_1 and WWL_2 are connected to the front and back gates of the double gate pMOS MW. These nodes play a key role in enabling the write operation, but they also influence CC during switching. The WWL node needs to be biased in such a way that the following two conditions are satisfied:

(1) Allow full-swing voltage to be written on the SN with very fast access time. It enables reliable logic '0' and '1' voltage on the SN without having any threshold voltage drop.
(2) Minimize the CC by reducing abrupt voltage change at the WWL node during switching (write-to-hold transition). This maximizes the ΔV.

Moving forward with the first requirement, i.e., writing full-swing voltage to SN, the write transistor needs to be turned ON, and for that, applied WWL voltage should be greater than its V_{Th}. Back gate bias has an influence on the threshold voltage of MOSFET [65]. Fig. 3.15 shows threshold voltage as a function of back gate bias (V_{BG}). V_{Th} can be extracted by many methods, such as constant current, transconductance-to-current

(g_m/I_{DS}) ratio, and the linear extrapolation method [66]. In this work, the constant current method has been used to extract V_{Th}, as the gate voltage corresponds to a defined threshold current (I_{Th}). The mathematical equation to determine V_{Th} is as follows:

$$I_{Th} = \left(\frac{W}{L}\right) 100\ nA \tag{3.6}$$

$$V_{Th} = V_{GS} @ I_{SD} = I_{Th} \tag{3.7}$$

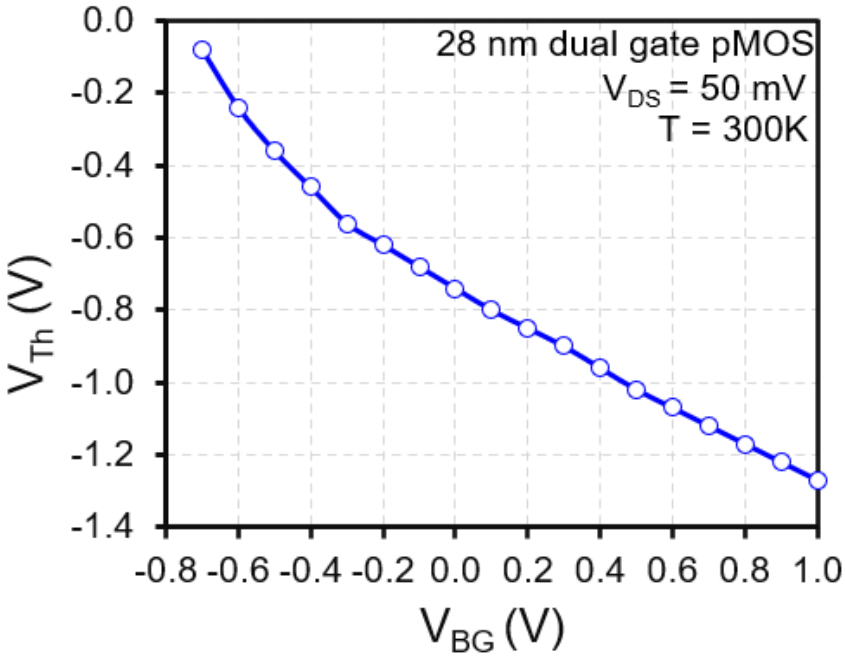


Fig. 3.15 Impact of back gate bias (V_{BG}) on threshold voltage (V_{Th}).

Fig. 3.15 illustrates that V_{Th} is nearly equal to -0.4 V when both gates (WWL_1 and WWL_2) are tied, and when the back gate (WWL_2) is grounded and only the front gate (WWL_1) is active, V_{Th} is equal to -0.7 V. Since MW is a p-type device, it will turn ON when gate voltage is more negative than V_{Th}. Hence, to perform a proper write operation, WWL_1 must be lower than -0.7 V (V_{WWL1} = -1 V while keeping V_{WWL2} = 0 V). This guarantees MW turns ON and allows a full-swing write operation. However, this introduces a new problem. During the hold state, WWL_1 must be brought back to a high voltage (typically to V_{DD} = 1 V) to turn MW OFF.

This transition from -1 V to +1 V creates a 2 V swing, which couples strongly into SN and minimizes ΔV. The result is shown in Fig. 3.16, where the SN experiences a sharp voltage change due to high CC. Here, the difference between logic '0' and logic '1' is ΔV = 0.48 V, and the estimated DRT is only 65 µs due to the strong CC disturbance.

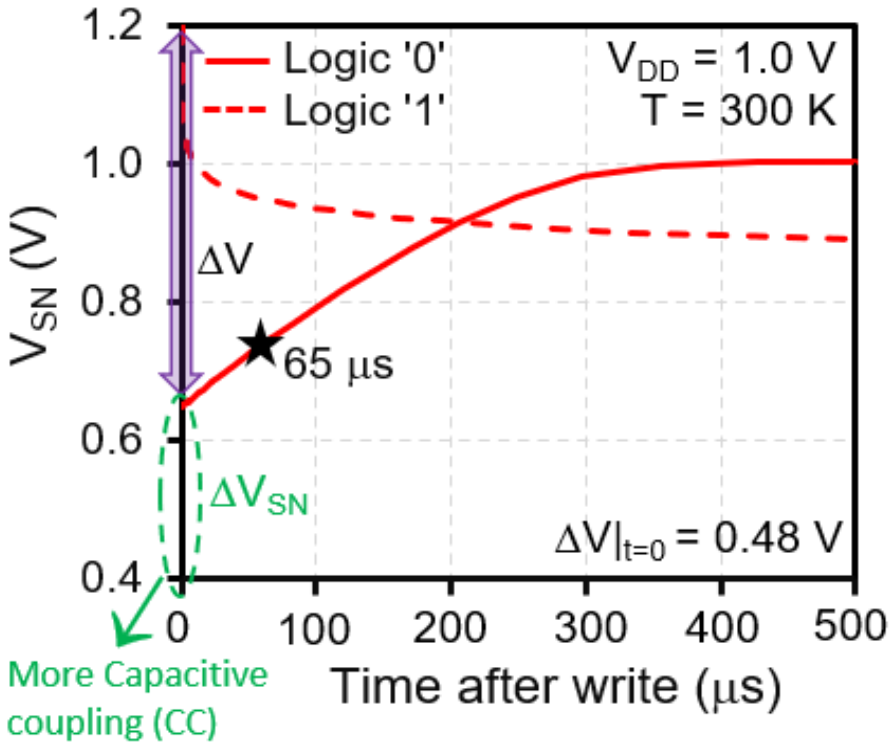


Fig. 3.16 Pre optimized V_{SN} degradation of DG 2T pMOS-only GC following write operation.

It is important to note that when V_{WWL1} = -0.4 V and V_{WWL2} = -0.4 V, then also the problem exists. Although this transition from -0.4 V to +1 V creates a lower voltage swing of 1.4 V, but now it has the coupling on SN from both of the gates. Using both gates (WWL_1 and WWL_2) simultaneously during writing adds more coupling paths [62]. Hence, V_{WWL2} is fixed at 0 V (grounded) and control MW using only WWL_1. The mathematical models for capacitive coupling are shown in equations 3.8 and 3.9, which conclude that a larger change in write word line biases (such

as ΔV_{WWL1} and ΔV_{WWL2}) results in a larger change in storage node voltage (ΔV_{SN}) and minimizes ΔV.

$$\Delta V_{SN} \cong \frac{C1}{C1+C2+C3+C4} \Delta V_{WWL1} \quad (3.8)$$

$$\Delta V_{SN} \cong \frac{C2}{C1+C2+C3+C4} \Delta V_{WWL2} \quad (3.9)$$

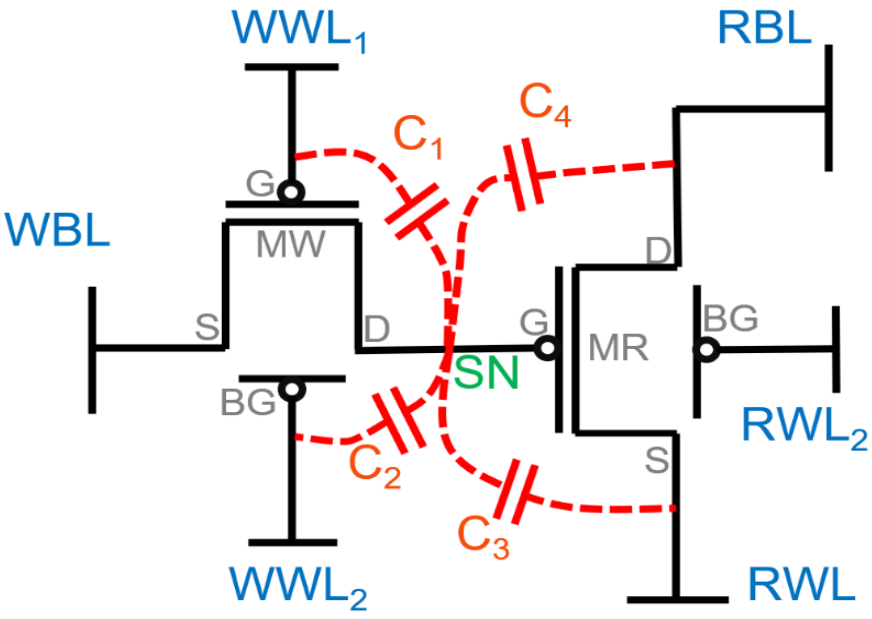


Fig. 3.17 Schematic of a DG 2T pMOS-only GC with the coupling capacitances between SN, WWL_1, WWL_2, RBL, and RWL.

Moving forward with the second requirement, i.e., reducing CC and improving DRT, the voltage swing on WWL1 need to be decreased. From the threshold voltage curve (Fig. 3.15), it is clear that MW turns ON at around V_{WWL1} = -0.7 V while keepingV_{WWL2} = 0 V. Therefore, instead of using -1 V, set V_{WWL1} = -0.7 V, which is just enough to write data while still turning off with a smaller voltage swing (from -0.7 V to +1 V = 1.7 V swing, instead of 2 V). The result is shown in Fig. 3.18, where the SN now experiences less voltage change due to reduced CC. Here, ΔV increases from 0.48 V to 0.57 V, giving more sensing margin, and DRT improves to 102 µs from 65 µs.

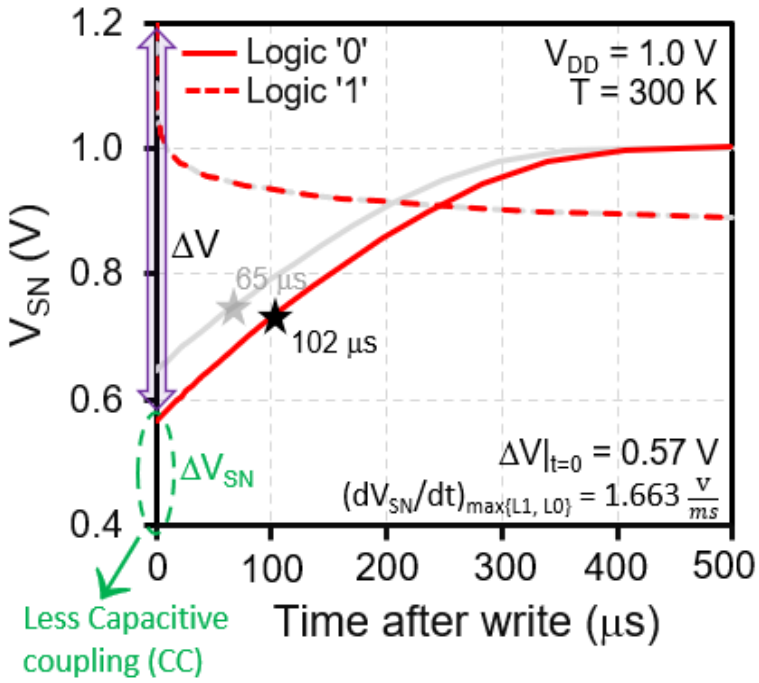


Fig. 3.18 V_{SN} degradation with optimized WWL node bias.

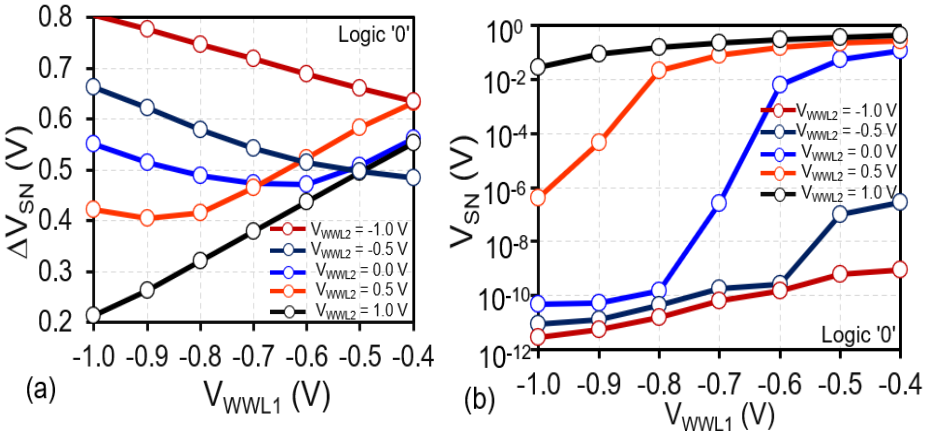


Fig. 3.19 (a) Change in storage node voltage at the beginning of hold operation under different bias, and (b) Written voltage level for logic '0' on storage node under different bias conditions.

Fig. 3.19 provides a trade-off perspective between capacitive coupling and write strength. The left plot (Fig. 3.19 (a)) shows the variation of ΔV_{SN} with V_{WWL1} for several fixed values of V_{WWL2}. A lower ΔV_{SN} is desirable because it implies less capacitive disturbance during the write-to-hold transition. When V_{WWL2} is high, such as 0.5 V or 1 V, ΔV_{SN} remains

low because the total coupling from MW is reduced due to reduced back gate participation. This suggests that higher WWL_2 bias helps to reduce CC, but this comes at a cost. The right plot (Fig. 3.19 (b)) shows how well logic '0' is written onto the SN for different V_{WWL1} values, again under five WWL_2 voltage conditions. A lower SN voltage after writing logic '0' indicates stronger writing, which is essential for full-swing memory operation. When V_{WWL2} is high (e.g., 1 V or 0.5 V), the SN does not reach a proper '0' level even when V_{WWL1} is very negative. This is because the back gate is pulling the channel in the opposite direction, increasing the threshold and weakening MW's conduction. The best trade-off occurs where V_{WWL2} = 0 V and V_{WWL1} = -0.7 V. This ensures a full-swing logic '0' write while maintaining moderate capacitive coupling. Therefore, V_{WWL1} = -0.7 V and V_{WWL2} = 0 V are selected as the optimized WWL bias for our 2T pMOS-only DG Gain Cell.

Using a higher negative WWL_1 bias (like -1 V) ensures strong writing but introduces high CC. On the other hand, using a less negative WWL_1 bias (like -0.7 V) slightly reduces write strength but significantly improves DRT by reducing CC. The Fig. 3.18 also shows data degradation slope of 1.663 V/ms which needs to be minimized to improve DRT further and for this WBL bias need to be optimized.

3.4.2 Optimizing WBL node bias

After optimizing the WWL node to reduce CC, the next important step in enhancing DRT is minimizing the leakage current that causes a slow degradation of the stored data over time. This can be effectively achieved by optimizing the write bit line bias. During the hold phase, the MW is OFF, and SN is left floating. However, there is a voltage difference between the WBL and SN (a non-zero V_{SD} across MW), resulting in leakage current flowing through the MW transistor. This causes a slow but continuous degradation of SN voltage, reducing DRT. When V_{SD} (= V_{WBL} - V_{SN}) is minimized by choosing the WBL voltage during Hold, the leakage can be

suppressed and improve retention further, which leads to best-case retention.

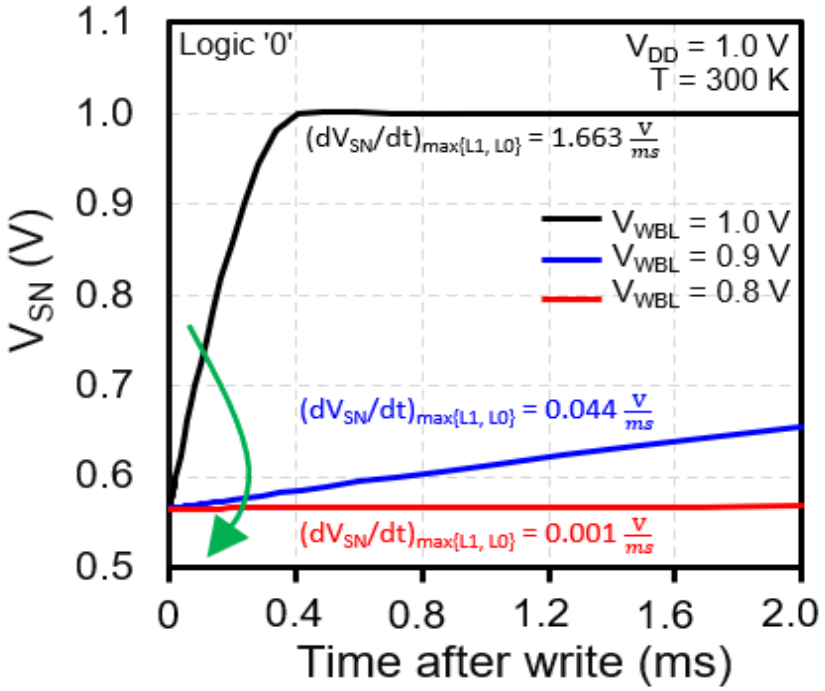


Fig. 3.20 V_{SN} degradation of logic ‘0’ with three fixed WBL node bias.

The idea is to fix the WBL voltage during hold to a level that minimizes the voltage difference between SN and WBL. This minimizes the leakage current through MW. However, this also requires careful balancing, because in a practical gain cell, one logic tends to degrade very quickly. For pMOS-only designs, logic '0' tends to degrade faster. Therefore, a focus on optimizing logic ‘0’ retention is needed. Different choices of WBL bias lead to either worst-case or best-case retention behavior. The worst-case scenario occurs when WBL is pulled opposite to SN logic, i.e. set to a voltage level that maximizes V_{SD}. For example, if SN stores logic ‘0’, set WBL to 1 V, and if SN stores logic ‘1’, set WBL to 0 V. This high potential difference results in significant subthreshold leakage through MW and causes rapid degradation of the SN, as shown by the steep slope in Fig. 3.20. The best-case scenario is achieved by fixing WBL at a level that minimizes V_{SD} for both logics.

To understand how fixed WBL bias affects leakage and DRT, simulation results are shown in Fig. 3.20. The WBL voltage during hold was swept across three fixed values: 1.0 V, 0.9 V, and 0.8 V. At V_{WBL} = 1.0 V, the voltage difference (V_{SD}) across MW is large, leading to high leakage and a fast SN degradation slope (dV_{SN}/dt) ≈ 1.663 V/ms. At V_{WBL} = 0.9 V, leakage reduces, and the slope decreases to 0.044 V/ms. At V_{WBL} = 0.8 V, leakage is minimal, and the slope drops to a very low value of 0.001 V/ms — an improvement of 3 orders of magnitude.

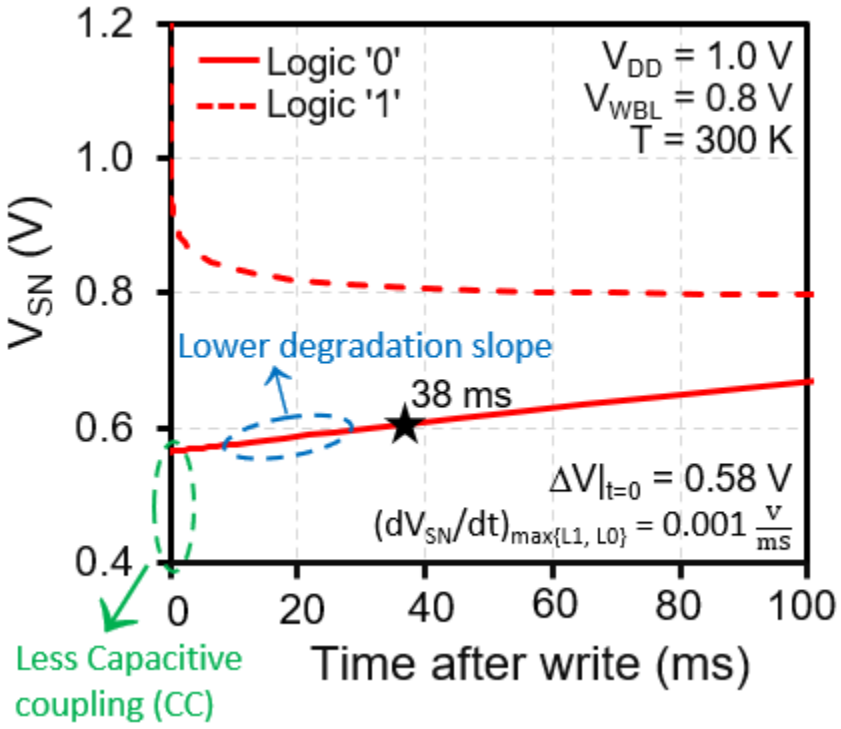


Fig. 3.21 Optimized V_{SN} degradation of DG 2T pMOS-only GC following write operation.

Fig. 3.21 shows that fixing WBL at 0.8 V during hold operation drastically reduces leakage current and minimizes the SN degradation slope, and DRT increases to 38 milliseconds. The analysis confirms that WBL biasing is a powerful tool to suppress leakage and enhance retention nearly by 1000x [40].

3.4.3 Optimizing RWL and RBL node biases

RWL and RBL node biases have less influence on DRT because these nodes are isolated from the SN by the read transistor. Hence, optimizing of RWL and RBL node bias is required to improve read access time (RAT).

During the write and hold operations, the read path should remain OFF. To ensure this, the drain to source voltage (V_{DS}) of MR is kept at 0 V, so that no leakage is introduced through the read transistor. For this we can consider two scenarios:

(1) Setting both RWL and RBL nodes at 0 V.

(2) Setting both RWL and RBL nodes at 1 V, again ensuring zero V_{DS}.

During read operation, RWL is pulled high (for the 1st scenario) or pulled low (for the 2nd scenario), which creates a change in RBL, and thus reading happens. Although both cases are acceptable, but the first option is preferable for pMOS read transistors as pMOS passes a strong '1' and a weak '0'. In read phase, when RWL is changed to 1 V from 0 V, it creates a change in RBL very quickly and thus improves read access time.

3.5 Benchmarking

The optimized results are compared with other reported works as shown in Fig. 3.22. Based on the proposed design and optimization using independent gate biasing, the final result shows a significant enhancement in DRT of the 2T topology. Estimated best case data retention time is 38 ms at 300 K and 300 µs at 358 K.

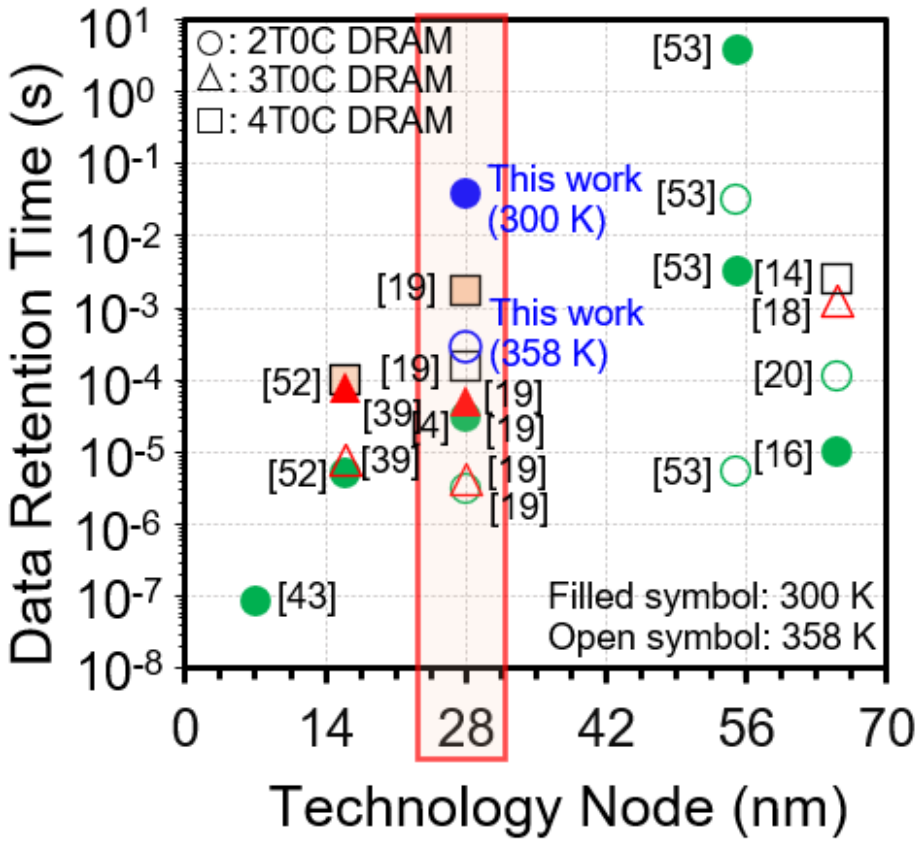


Fig. 3.22 Benchmarking of reported Si GC-eDRAM at different technology nodes.

3.6 Conclusion

The double gate 2T pMOS-only gain cell has been thoroughly studied in this chapter, with a focus on improving data retention time. Discussion started with the basic device behaviour of the double gate pMOSFET, focusing on its enhanced electrostatic control and lower leakage than single gate devices. The architecture of the DG 2T pMOS-only gain cell was introduced, followed by a step-by-step explanation of its write, hold, and read operations. Key physical phenomena such as capacitive coupling and subthreshold leakage were identified as critical barriers to achieving long retention times. To address these issues, a two-step

enhancement strategy, involving (1) minimizing capacitive disturbances through optimized word-line biases and (2) suppressing leakage currents by carefully fixing the write bit-line during the hold phase, has been analysed. Detailed simulation results confirmed that this approach significantly improves the DRT from 65 μs to 38 ms.

Chapter 4

Impact of Supply Voltage and Temperature

As technology scales down and low-power operation becomes more important, lowering the supply voltage has become a popular way to minimize dynamic power consumption. Voltage scaling in digital and embedded memory systems not only helps to lower the total power budget but also enables compatibility with energy-constrained platforms including wearable electronics, biomedical implants [28], portable devices and IoT sensors [48]. This chapter investigates the impact of downscaling the supply voltage and increasing temperature on the data retention time of the proposed double gate 2T pMOS-only gain cell.

4.1 Impact of V_{DD} scaling on DRT

Scaling V_{DD} is an effective method for reducing dynamic power, as the power (P) consumed is proportional to the square of V_{DD} i.e., $P \propto V_{DD}^2$ [5]. In energy-sensitive applications where battery life is critical, this method is especially helpful. However, voltage scaling has a negative impact on memory behavior, especially for gain cells or embedded DRAMs, as it deteriorates sensing margins, makes it harder to write full-swing data, and exacerbates leakage behavior [27].

One critical reason for the degradation in DRT is the increase in leakage in the gain cell during hold. It seems that if supply voltage is scaled down, then V_{SD} of MW is reduced and leakage current also decreases. But in this case, it is not true, because scaling down not only changes the horizontal electric field (source to drain) but also the vertical electric field

(gate to channel). As a result, lower positive voltages at the WWLs (gates of MW) terminal during the hold state prevent the MW from entering a hard-off state. In pMOS devices, entering a strong cut-off state requires V_{SG} to be sufficiently positive. When V_{DD} is reduced, the voltage difference between the gate and source terminals becomes smaller, which weakens the off-state behavior of the pMOS. As a result, subthreshold leakage increases, leading to faster charge loss from the storage node and significantly reducing the retention time.

Another critical reason for the degradation in DRT is the decrease in ΔV. As supply voltage is reduced, one of the immediate effects is the decrease in the full-swing write voltage. With V_{DD} = 1.0 V, the write bit line alternates between 0 V and 1 V, so enabling a complete swing of 1 V to be applied to the SN. As a result, the SN is written with an accurately distinct '0' and '1', ensuring a longer DRT. However, when V_{DD} is scaled down to 0.7 V, the WBL now operates between 0 V and 0.7 V, reducing the voltage swing to just 0.7 V. This restricts the distinguishability between logic '1' and logic '0', so lowering the ΔV and hence lowering the DRT.

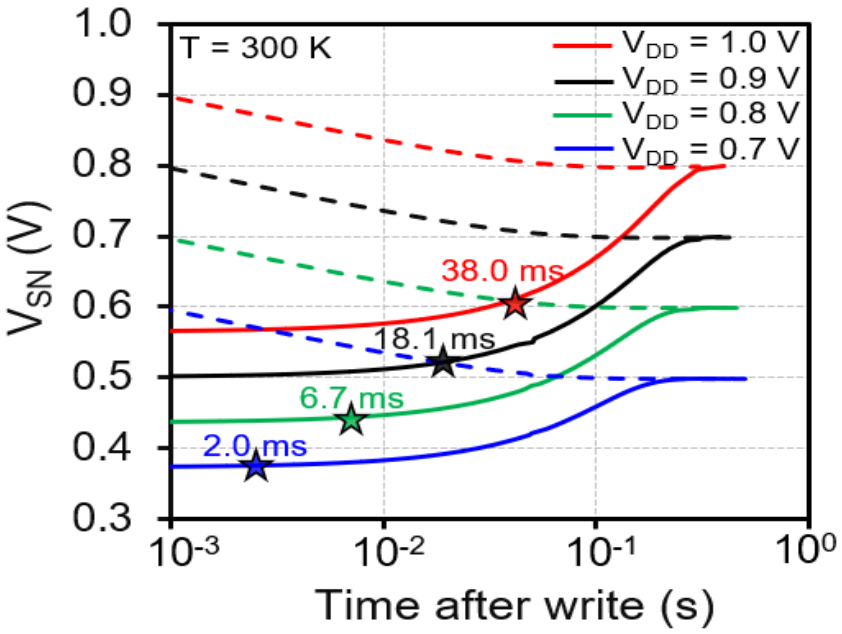


Fig. 4.1 V_{SN} degradation at four different supply voltages at 300 K.

The stored voltages degrade faster when the write swing is lowered and leakage is increased. Fig. 4.1 illustrates these issues, where the degradation of the storage node voltage is plotted for different supply voltages between 1.0 V and 0.7 V at room temperature (300 K). This plot clearly highlights how aggressively retention time reduces as V_{DD} is scaled. At V_{DD} = 1.0 V, the cell achieves the maximum DRT of 38 ms. At V_{DD} = 0.9 V, retention drops to 18.1 ms, which is roughly a 50% degradation. With further scaling to 0.8 V, DRT shrinks to 6.7 ms, a 5.7x reduction compared to 1.0 V. At V_{DD} = 0.7 V, the DRT drops drastically to just 2.0 ms, nearly a 20x reduction from the baseline.

Table 4.1 Impact of supply voltage (V_{DD}) scaling on DRT at 300 K.

Supply Voltage [V_{DD}] (V)	**Difference Window [$\Delta V\|_{t=0}$] (V)**	**WBL Bias (V)**	**DRT (ms)**
1.0	0.58	0.8	38.0
0.9	0.52	0.7	18.1
0.8	0.48	0.6	6.7
0.7	0.45	0.5	2.0

Supply voltage scaling poses a significant problem for memory retention even though it is advantageous from a power-saving point of view. From the analysis above, it is clear that downscaling V_{DD} in a gain cell memory directly reduces the data retention time for several reasons — reduced full-swing write, poor pMOS off-state, increased leakage, and lower sensing margin.

4.2 Impact of higher temperature on DRT

In practical case, electronic circuits including embedded memories do not run at perfect room temperature conditions. While most simulations and characterizations start at 300 K (27°C), actual conditions can differ. As a device heats up, its threshold voltage reduces, which makes the subthreshold leakage current rise. The leakage contributes to even more heat, and consequently, the temperature rises further. Because of this, memory arrays usually work at temperatures well above room temperature. Simulating and analyzing the behavior of our DG 2T pMOS-only gain cell at high temperatures is crucial to fairly estimate its DRT. Thus, the present work investigates performance criteria at T = 358 K (85°C), a reasonable upper limit for modern CMOS technology under continuous operation.

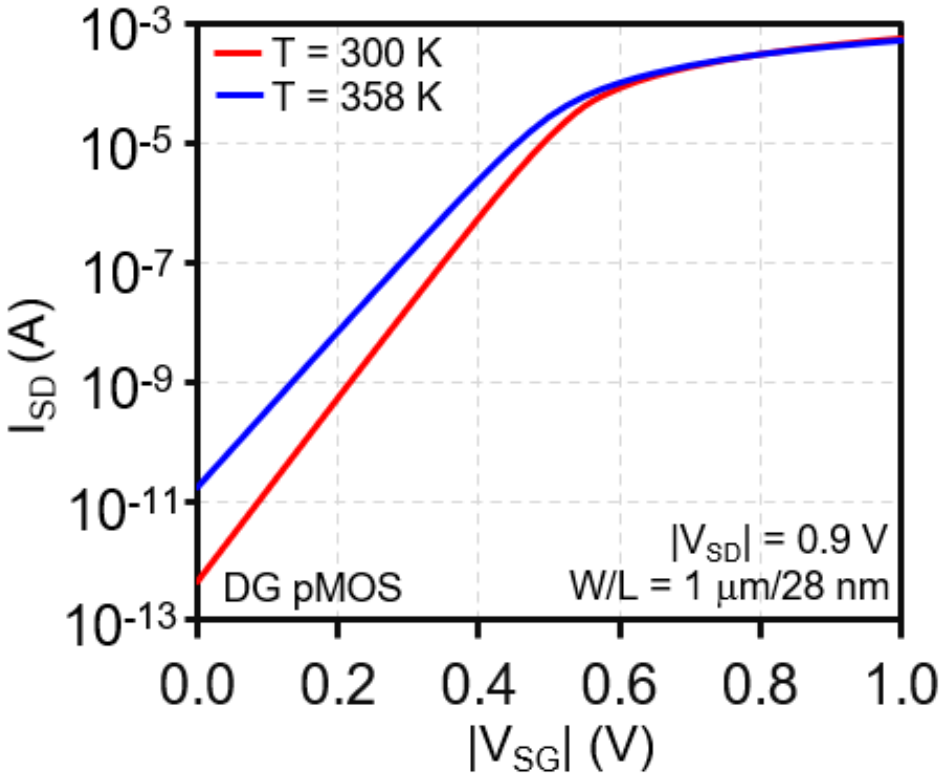


Fig. 4.2 Transfer characteristics of DG pMOS at different temperatures.

To understand the impact of temperature on memory performance, this study starts with the transfer characteristics (I_{SD} vs V_{SG}) of the pMOS write transistor. The simulated curves (I_{SD} vs V_{SG}) at 300 K and 358 K for a 28 nm double gate pMOSFET are plotted against each other in Fig. 4.2. The most noticeable change is the rise in I_{SD} at higher temperatures. As temperature rises, thermal excitation also increases, and thermal energy helps excite more intrinsic carriers. As a result, threshold voltage reduced and leakage currents increased. The results show OFF current is increased by a magnitude of almost two orders when temperature changes from 300 K to 358 K.

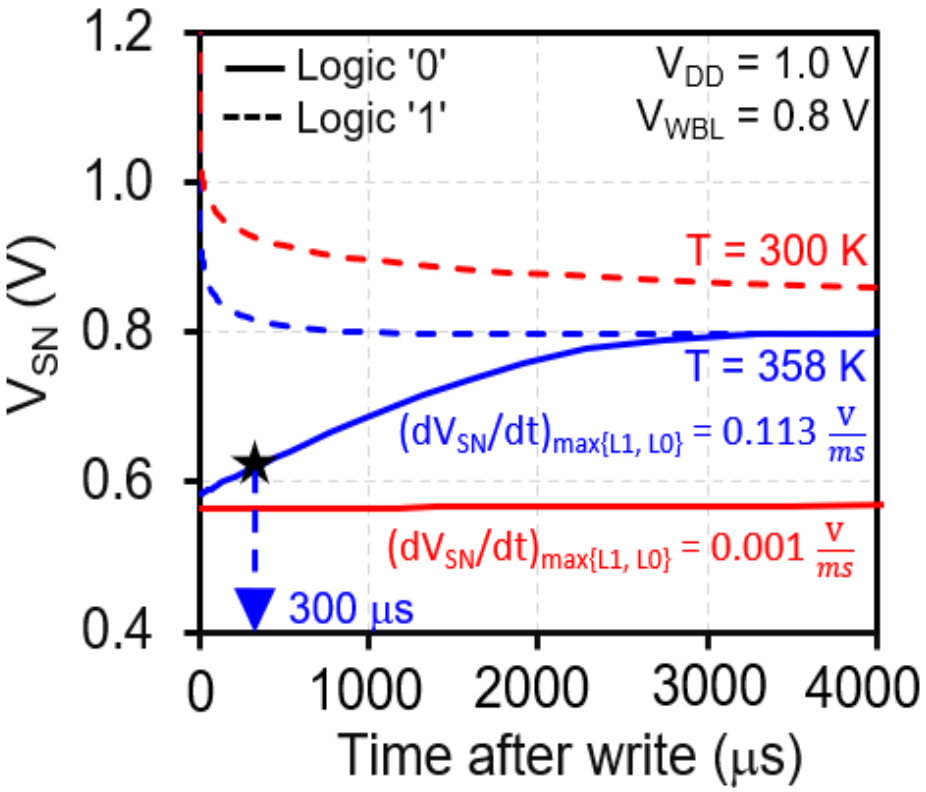


Fig. 4.3 Optimized V_{SN} degradation of DG 2T pMOS-only GC at two different temperatures.

The Fig. 4.3 represents SN voltage degradation at 300 K and 358 K for a 28 nm DG pMOS-only GC. The red curve corresponds to nominal room temperature operation where leakage is minimal. Hence, a relatively flat V_{SN} degradation slope of nearly 0.001 V/ms and a higher DRT of 38 ms. However, the degradation of V_{SN} is much steeper at T = 358 K, even under the same biasing conditions. The steeper V_{SN} degradation slope of 0.113 V/ms results sharp drop in the data retention time to just 300 µs.

Table 4.2 Impact of higher temperature on DRT at V_{DD} = 1 V.

Temperature [T] (K)	**Difference Window [$\Delta V|_{t=0}$] (V)**	**Degradation Slope [dV_{SN}/dt] (V/ms)**	**DRT (ms)**
300	0.58	0.001	38
358	0.49	0.113	0.3

One key observation is that even the logic '1' (dashed curve) begins to slope downward as compared to room temperature. Previously, even when the bias of several nodes changed, logic '1' was mostly immune to degradation at fixed room temperature. This highlights that, although logic '0' still degrades more rapidly due to the weaker strength of pMOS to write zero, both logic states suffer at higher temperatures. The second major noticeable point is that lowering the ΔV. Higher temperatures also worsen short-channel effects and parasitic coupling, which means a small voltage transition in adjacent control lines can now significantly disturb the SN voltage. The effective sensing voltage window level decreased to 0.49 V from 0.58 V as the temperature increased to 358 K from 300 K. As a combined effect, DRT reduces to 300 µs from 38 ms when the temperature rises to 358 K from 300 K.

4.3 Impact of V_{DD} scaling at 358 K

This section investigates the combined impact of low supply voltage and high temperature on the performance of a 2T DG pMOS-only gain cell. Fig. 4.4 below shows the combined effect of V_{DD} scaling and high temperature (T = 358 K) on DRT. At V_{DD} = 1.0 V, the gain cell retains data for about 320 µs, but drops sharply to 30 µs when V_{DD} is reduced to 0.7 V.

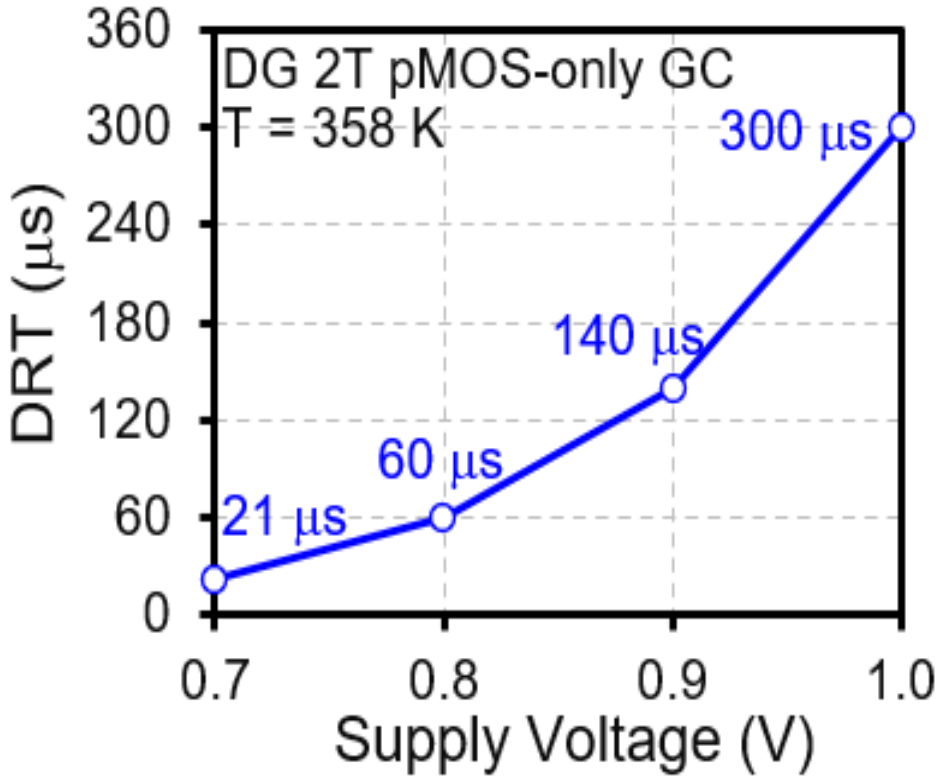


Fig. 4.4 Impact of supply voltage scaling on DRT at 358 K.

4.4 Conclusion

This chapter focused on exploring the robustness of the proposed DG 2T pMOS-only gain cell under two critical conditions — supply voltage scaling and temperature elevation. It begins with motivation for V_{DD} scaling, which is central to low-power design, especially in portable and embedded systems. Then the impact of high temperature is also discussed. Finally, combining both effects, this chapter demonstrated the challenges of low supply voltage and high temperature operation for DRAM as DRT falls below 60 μs in the worst-case situation.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

This thesis explores an effective approach to enhance DRT by using Double Gate (DG) pMOSFETs within a 2T pMOS-only GC structure at the 28 nm technology node. The work started with a basic understanding of gain cell operation and focused on addressing two major DRT limiting factors — capacitive coupling and subthreshold leakage.

A major highlight of this work is the use of independent-gate control to minimize capacitive coupling and subthreshold leakage. This work discussed a two-step DRT enhancement approach. First, minimizing capacitive coupling by independent gate bias operation, so allowing a larger voltage difference (ΔV) between two logic levels. Second, suppressing leakage by WBL bias tuning, so significantly lowering the data degradation rate. Simulation results confirmed that this approach significantly improves the DRT from 65 μs to 38 ms.

The work also analyzes the effects of higher temperature operation and supply voltage (V_{DD}) downscaling. After careful optimization, the obtained results exceed those reported in the literature. The work highlights the importance of GC optimization to improve DRT. Also, independent gate operation can utilized to further improve the performance of silicon based GCs at advanced technology nodes.

5.2 Future Work

Although the main goal of this thesis work was to enhance data retention time, the read access time is remains one important performance parameter needs to be improved. Read access time is directly related to the time required for the read transistor to charge or discharge the read bit line when it is turned on. To achieve this, future work can explore the optimization of read transistor.

REFERENCES

[1] P. Meinerzhagen, A. Teman, R. Giterman, N. Edri, A. Burg, and A. Fish (2018) Gain-Cell Embedded DRAMs for Low-Power VLSI Systems-on-Chip, Springer International Publishing AG, ISBN 978-3-319-60402-2.

[2] J. Constantin, A. Dogan, O. Andersson, P. Meinerzhagen, J. N. Rodrigues, D. Atienza, and A. Burg (2012) TamaRISC-CS: An ultra-low-power application-specific processor for compressed sensing, IEEE/IFIP International Conference on VLSI System-on-Chip (VLSI-SoC), 159-164.

[3] W. Zhang, K. C. Chun, and C. H. Kim (2010) Variation aware performance analysis of gain cell embedded DRAMs, In: Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED), 19-24.

[4] J. B. Kuo and J. H. Lou (1999) Low-Voltage CMOS VLSI Circuits, John Wiley & Sons, ISBN: 0-471-32105-2.

[5] J. M. Rabaey, A. Chandrakasan, and B. Nikolic (2003) Digital Integrated Circuits: A Design Perspective, 2nd edition, Prentice Hall, ISBN: 978-7-302-07968-2.

[6] A. Teman, and R. Visotsky (2015) A Fast Modular Method for True Variation-Aware Separatrix Tracing in Nanoscaled SRAMs, IEEE Transactions on Very Large Scale Integration (VLSI), 23, 2034-2042.

[7] B. H. Calhoun and A. P. Chandrakasan (2007) A 256-kb 65nm Sub-Threshold SRAM Design for Ultra-Low-Voltage Operation, IEEE J Solid state circuit, 42, 680-688.

[8] M. Qazi, M. E. Sinangil, and A. P. Chandrakasan (2011) Challenges and Directions for Low-Voltage SRAM, IEEE Design & Test of Computers, 28, 32-43.

[9] S. Kang, and Y. Leblebici (2003) CMOS Digital Integrated Circuits: Analysis and Design, 3rd edn. McGraw-Hill, ISBN: 0-07-119644-7.

[10] R. Giterman, R. Golman, and A. Teman (2019) Improving Energy-Efficiency in Dynamic Memories Through Retention Failure Detection, IEEE Access, 7, 27641-27649.

[11] H. A. Hussien Abdo (2020), Design of a 2Kb Cache Memory Array using GC-eDRAM Cell Implemented in 130nm Standard CMOS Technology. Degree Thesis, Universiti Teknologi Malaysia, Skudai.

[12] H. Kaeslin, and E. Zurich (2008) Digital Integrated Circuit Design From VLSI Architectures to CMOS Fabrication, 1st edn. Cambridge University Press, Cambridge, ISBN: 978-0-511-45537-7.

[13] A. Teman, P. Meinerzhagen, A. Burg, and A. Fish (2012) Review and Classification of Gain Cell eDRAM Implementations, IEEE 27-th Convention of Electrical and Electronics Engineers in Israel, 1-5.

[14] O. Maltabashi, H. Marinberg, R. Giterman, and A. Teman (2018) A 5-Transistor Ternary Gain-Cell eDRAM with Parallel Sensing, IEEE International Symposium on Circuits and Systema (ISCAS), 1-5.

[15] M. Ichihashi, H. Toda, Y. Itoh, and K. Ishibashi (2005) 0.5V Asymmetric Three-Tr. Cell (ATC) DRAM Using 90nm Generic CMOS Logic Process, IEEE Symposium on VLSI Circuits, Digest of Technical Papers, 23, 366-369.

[16] D. Somasekhar, Y. DaleYe, P. Aseron, S. Lu, M. M. Khellah, J. Howard, G. Ruhl, and T. Karnik (2009) 2 GHz 2 Mb 2T Gain Cell Memory Macro With 128 GBytes/sec Bandwidth in a 65 nm Logic Process Technology, IEEE Journal of Solid-State Circuits, 44, 174-185.

[17] Y. Lee, M. Chen, J. Park, D. Sylvester, and D. Blaauw (2010) A 5.42nW/kB Retention Power Logic-Compatible Embedded DRAM with 2T Dual-Vt Gain Cell for Low Power Sensing Applications, IEEE Asisn Solid-State Circuits (A-SSCC) Conference, Beijing, China, 1-4.

[18] K. C. Chun, P. Jain, J. H. Lee, and C. H. Kim (2011) A 3T Gain Cell Embedded DRAM Utilizing Preferential Boosting for High Density and Low Power On-Die Caches, IEEE Journal of Solid-State Circuts, 46, 1495-1505.

[19] R. Giterman, A. Fish, A. Burg, and A. Teman (2018) A 4-Transistor nMOS-Only Logic-Compatible GainCell Embedded DRAM With Over 1.6-ms Retention Time at 700 mV in 28-nm FD-SOI, IEEE Transactions on Circuits and Systems-I Regular Papers, 65, 1245-1256.

[20] R. Giterman, A. Fish, N. Geuli, E. Mentovich, A. Burg, and A. Teman (2017) An 800 Mhz Mixed-VT 4T Gain-Cell Embedded DRAM in 28 nm CMOS Bulk Process for Approximate Computing Applications, IEEE European Solid-State Circuits Conference (ESSCIRC) , 308-311.

[21] International technology roadmap for semiconductors (2023). https://irds.ieee.org/editions/2023

[22] T. N. Blalock and C. Jaeger (1990) An Experimental 2T Cell RAM with 7 ns Access Time at Low Temperature, IEEE Symposium on VLSI Circuits, 13-14.

[23] R. Giterman, A. Teman, P. Meinerzhagen, A. Burg, and A. Fish (2014) 4T Gain-Cell with Internal-Feedback for Ultra-Low Retention Power at Scaled CMOS Nodes, IEEE International Symposium on Circuits and Systems (ISCAS), 2177-2180.

[24] S. R. Soo, A. Hamzah, N. E. Alias, I. Kamisian, M. L. Tan, S. Isaak, and Z. Johari (2021) Design of Low Power Gain-Cell eDRAM for 4Kb Memory Array in 130nm CMOS, IEEE International Conference on Electrical Engineering and Informatics (ICEEI), 1-6.

[25] W. K. Luk, and R. H. Dennard (2005) A Novel Dynamic Memory Cell With Internal Voltage Gain, IEEE Journal of Solid-State Circuits, 40, 884-894.

[26] K. C. Chun, P. Jain, T. Kim, and C. H. Kim (2012) A 667 MHz Logic-Compatible Embedded DRAM Featuring an Asymmetric 2T Gain Cell for High Speed On-Die Caches, IEEE Journal of Solid-State Circuits, 47, 574-559.

[27] R. Iqbal, P. Meinerzhagen, and A. Burg (2012) Two-Port Low-Power Gain-Cell Storage Array: Voltage Scaling and Retention Time, IEEE International Symposium on Circuits, 2469-2472.

[28] P. Meinerzhagen, A. Teman, A. Mordakhay, A. Burg, and A. Fish (2012) A Sub-VT 2T Gain-Cell Memory for Biomedical Applications, IEEE Subthreshold Microelectronics Conference ($SubV_T$), 1-3.

[29] P. Meinerzhagen, A. Teman, R. Giterman, A. Burg, and A. Fish (2013) Exploration of Sub-VT and Near-VT 2T Gain-Cell Memories for Ultra-Low Power Applications under Technology Scaling, Journal of Low Power, 3, 54-72.

[30] D. Somasekhar, S. Lu, B. Bloechel, G. Dermer, K. Lai, S. Borkar and V. De (2005) A 10Mbit, 15GBytes/sec Bandwidth 1T DRAM Chip with Planar MOS Storage Capacitor in an Unmodified 150nm Logic Process for HighDensity On-Chip Memory Applications, In: Proceedings of the IEEE European Solid-State Circuits Conference (ESSCIRC), 355–358.

[31] R. Sailigram, S. Dutta, and A. Raychowdhury (2021) CryoMem: A 4–300-K 1.3-GHz Hybrid 2T-Gain-Cell-Based eDRAM Macro in 28-nm Logic Process for Cryogenic Applications, IEEE Solid-State Circuits Letters, 4, 194-197.

[32] R. Giterman, Y. Weizman, and A. Teman (2018) Gain-Cell Embedded DRAM-Based Physical Unclonable Function, IEEE Transactions on Circuit and Systems-I, 65, 4208-4218.

[33] M. Chang, P. Huang, and W. Hwang (2007) A 65nm Low Power 2T1D Embedded DRAM with Leakage Current Reduction, IEEE International SoC Conference, Hsinchu, Taiwan, 207-210.

[34] K. C. Chun, P. Jain, J. H. Lee, and C. H. Kim (2009) A Sub-0.9V Logic-compatible Embedded DRAM with Boosted 3T Gain Cell, Regulated Bit-line Write Scheme and PVT-tracking Read Reference Bias, IEEE symposium on VLSI circuits (VLSIC), 134–135.

[35] M. U. Khalid, P. Meinerzhagen, and A. Burg (2012) Replica Bit-Line Technique for Embedded Multilevel Gain-Cell DRAM, IEEE International NEWCAS Conference, 77–80.

[36] W. K. Luk and R. H. Dennard (2004) 2T1D Memory Cell with Voltage Gain, IEEE Symposium on VLSI Circuits (VLSIC), 184–187.

[37] P. Meinerzhagen, O. Andic, J. Treichler, and A. Burg (2011) Design and Failure Analysis of Logic-Compatible Multilevel Gain-Cell-Based DRAM for Fault-Tolerant VLSI Systems, IEEE/ACM Great Lakes Symposium on VLSI (GLSVLSI), 343–346.

[38] Y. S. Park, D. Blaauw, D. Sylvester, and Zhengya Zhang (2012) A 1.6-mm2 38-mW 1.5-Gb/s LDPC Decoder Enabled by Refresh-Free Embedded DRAM, IEEE Symposium on VLSI Circuits (VLSIC), 114–115.

[39] R. Giterman, A. Shalom, A. Burg, A. Fish, and A. Teman (2020) A 1-Mbit Fully Logic-Compatible 3T Gain-Cell Embedded DRAM in 16-nm FinFET, IEEE Solid-State Circuits Letters, 3, 110-113.

[40] R. Giterman, A. Teman, P. Meinerzhagen, L. Atias, A. Burg, and A. Fish (2016) Single-Supply 3T Gain-Cell for Low-Voltage Low-Power Applications, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 24, 358-362.

[41] J. Narinx, R. Giterman, A. Bonetti, N. Frigerio, C. Aprile, A. Burg, and Y. Leblebici (2019) A 24 kb Single-Well Mixed 3T Gain-Cell eDRAM with Body-Bias in 28 nm FD-SOI for Refresh-Free DSP Applications, IEEE Asian Solid-State Conference, 219-222.

[42] W. K. Luk, J. Cai, R. H. Dennard, M. J. Immediato, and S. V. Kosonocky (2006) A 3-Transistor DRAM Cell with Gated Diode for Enhanced Speed and Retention Time, IEEE Symposium on VLSI circuits (VLSIC), 184–185.

[43] B. S. Sany, and B. Ebrahimi (2021) A 1-GHz GC-eDRAM in 7-nm FinFET with static retention time at 700 mV for ultra-low power on-chip memory applications, International Journal of Circuit Theory and Applications, 50, 417-426.

[44] R. Giterman, A. Teman, and A. Fish (2017) A 11.5pW/bit 400mV 5T Gain-Cell eDRAM for ULP Applications in 28nm FD-SOI, IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S), Burlingame, CA, USA, 1-3.

[45] A. Bonetti, R. Golman, R. Giterman, A. Teman, and A. Burg (2020) Gain-Cell Embedded DRAMs: Modeling and Design Space, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 28, 646-659.

[46] K. C. Chun, P. Jain, T. Kim, and C. H. Kim (2010) A 1.1V, 667MHz Random Cycle, Asymmetric 2T Gain Cell Embedded DRAM with a 99.9 Percentile Retention Time of 110μsec, IEEE Symposium on VLSI Circuits, 191-192.

[47] P. Meinerzhagen, A. Teman, R. Giterman, N. Edri, A. Burg, and A. Fish (2018) Gain-Cell Embedded DRAMs for Low-Power VLSI Systems-on-Chip, Springer International Publishing AG, Chapter 2, 13-26.

[48] S. Kim, and J. Park (2022) Pseudo-Static Gain Cell of Embedded DRAM for Processing-in-Memory in Intelligent IoT Sensor Nodes, Sensors 22, 1-15.

[49] W. Khwa, P. Wu, J. Su, C. Cheng, J. Hsu, Y. Chen, L. Hsieh, J. Bai, Y. Kao, T. Lou, A. S. Lele, J. Wu, J. Tien, C. Lo, R. Liu, C. Hsieh, K. Tang, and M. Chang (2025) A 16nm 216kb, 188.4TOPS/W and 133.5TFLOPS/W Microscaling Multi-Mode Gain-Cell CIM Macro Edge-AI Devices, IEEE International Solid-State Circuits Conference (ISSCC), 252-254.

[50] P. Meinerzhagen, A. Teman, R. Giterman, N. Edri, A. Burg, and A. Fish (2018) Gain-Cell Embedded DRAMs for Low-Power VLSI Systems-on-Chip, Springer International Publishing AG, Chapter 5, 61-90.

[51] P. Meinerzhagen, A. Teman, A. Fish, and A. Burg (2013) Impact of Body Biasing on the Retention Time of Gain-Cell Memories, The Journal of Engineering, 2013, 19-22.

[52] A. Shalom, R. Giterman, and A. Teman (2018) High Density GC-eDRAM Design in 16nm FinFET, IEEE International Conference on Electronics, Circuit and Systems (ICECS), 585-588.

[53] K. Wang, P. Hao, F. Zhang, L. Zhang, Q. Huang, and R. Huang (2025) Logic-Compatible Asymmetrical FET for Gain Cell eDRAM With Long Retention and Fast Access Speed, IEEE Journal of the Electron Device Society, 13, 237-244.

[54] M. Alioto (2012) Ultra-Low Power VLSI Circuit Design Demystified and Explained: A Tutorial, IEEE Transactions. Circuits and Systems I: Regular Papers, 59, 3–29.

[55] B. Yu and M. Meyyappan (2006) Nanotechnology: Role in Emerging Nanoelectronics, Solid State Electronics, 50, 536–544.

[56] R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz, and M. Radosavljevic (2005) Benchmarking Nanotechnology for High-Performance and Low-Lower Logic Transistor Applications, IEEE Transactions on Nanotechnology, 4, 153-158.

[57] Q. Xie, Z. Wang, and Y. Taur (2017) Analysis of Short-Channel Effects in Junctionless DG MOSFETs, IEEE Transactions Electron Devices, 64, 3511–3514.

[58] A. Sarkar, A. K. Das, S. De, and C. K. Sarkar (2012) Effect of Gate Engineering in Double-Gate MOSFETs for Analog/RF Applications, Microelectronics Journal, 43, 873-882.

[59] A. Tsormpatzoglou, C. A. Dimitriadis, R. Clerc, Q. Rafhay, G. Pananakakis, and G. Ghibaudo (2007) Semi-Analytical Modeling of Short-Channel Effects in Si and Ge Symmetrical Double-Gate MOSFETs, IEEE Transactions Electron Devices, 54, 1943-1952.

[60] ATLAS User Manual, (2016) Silvaco, Santa Clara, CA, USA.

[61] A. Beckers, F. Jazaeria, H. Bohuslavskyib, L. Hutinb, S. Franceschib, and C. Enz (2019) Characterization and Modeling of 28nm FDSOI CMOS Technology Down to Cryogenic Temperatures, Solid-State Electronics, 159, 106-115.

[62] O. Phadke, S. G. Kirtania, D. Chakraborty, S. Datta, and S. Yu (2024) Suppressed Capacitive Coupling in 2TGC with Oxide Channel and Split Gate, IEEE Transactions on Electron Devices, 71, 6749-6755.

[63] K. Toprasertpong, S. Liu, J. Chen, S. Wahid, K. Jana, W.-C. Chen, S. Li, H. S. P. Wong, and E. Pop (2023) Co-designed Capacitive Coupling-Immune Sensing Scheme for Indium-Tin-Oxide 2TGC Operating at Positive Voltage below 2V, IEEE Symposium on VLSI Technology and Circuits.

[64] W. K. Henson, N. Yang, S. Kubicek, E. Vogel, J. Wortman, K. Meyer, and A. Naem (2000) Analysis of leakage currents and impact on off-state power consumption for CMOS technology in the 100-nm regime, IEEE Transactions on Electron Devices, 1393-1400.

[65] C. Neau, and K. Roy (2023) Optimal Body Bias Selection for Leakage Improvement and Process Compensation Over Different Technology Generations, International Symposium on Low Power Electronics and Design, ISLPED '03, 116-121.

[66] L. Dobrescu, M. Petrov, D. Dobrescu, and C. Ravariu (2000) Threshold Voltage Extraction Methods for MOS Transistors, IEEE International Semiconductor Conference, 371-374.