Doctoral Dissertation

Study on Semiconductor Memory Devices for Low-Power Mobile Applications

Yoshimitsu Yamauchi

January 2014

Graduate School of Engineering
Osaka University

Abstract

This thesis is study of semiconductor memory devices for low power mobile applications and is organized into 7 chapters.

Chapter 1 introduces the thesis. It provides an historical review of semiconductor memory technologies, explains the background and purpose of the study, and describes the organization of the chapters.

Chapter 2 describes a scalable multi-level-cell (MLC) virtual-ground flash memory cell technology that has potential for continued scaling of fast-access NOR flash memory. The virtual-ground cell size is scaled by eliminating a contact from the standard NOR flash memory cell. However, the cell-to-cell interference caused by the virtual-ground architecture makes it difficult to achieve MLC. In order to solve this, the mechanism of the interference was investigated and clarified by using neighboring cell leakage and capacitive coupling models, and a new cell structure and cell-to-cell interference canceling technique were proposed and successfully implemented into 128 Mbit MLC test chips. The proposed scalable virtual-ground MLC technology allows the CPU to execute a program directly without going through volatile memory, leading to a reduction in dynamic random access memory (DRAM) refresh power consumption.

Chapter 3 describes a source-side injection single-polysilicon split-gate NOR (S4-NOR) flash memory cell compatible with a CMOS logic process. For the proposed cell, the access gate and the floating gate were patterned on a channel between its source and drain with a small gap length using conventional photolithography technology beyond the 100 nm generation. The hot-electrons generated by the access gate transistor at the source side are injected to the floating gate at the drain side. The mechanism of the source-side channel hot-electron injection in a split-gate transistor was investigated and clarified by a modified lucky-electron model. The source-side injection efficiency is higher by four orders of magnitude as compared with that of the standard NOR cell and is insensitive to process parameters except the gap length in the split-gate transistor. Furthermore, due to the simple structure of the S4-NOR cell, it does not require any specialized process steps. Thus, the proposed flash memory is a promising candidate for next-generation embedded applications that require low power consumption and high performance without an increase in the cost of the logic chip.

Chapter 4 describes a nonvolatile DRAM cell combining a DRAM cell with a floating-gate MOSFET, which stores all the data in its floating gate in the event of

sudden power failure and can then recall all the data to DRAM quickly at power-up. An innovative store concept, which uses highly efficient Fowler-Nordheim (FN) tunneling of electrons between the DRAM storage-node and the floating gate of a MOSFET, was proposed and successfully implemented into 1 Mbit test chips. The experimental results of store characteristics show good agreement with the calculation results based on FN tunneling model. These technologies allow the CPU to store the data being processed in the event of a sudden power failure, thus overcoming the main drawbacks of silicon MOSFET-based volatile memories

Chapter 5 describes a novel voltage-loss-compensation (VLC) memory in pixel that was developed for an advanced dual-mode LCDs, and which can display still image using an ultra-low-power-consumption memory mode in addition to the normal display mode. In normal mode, the panel specification is identical to a standard transmissive display. In memory mode, the VLC circuit in each pixel enables simultaneous refresh with a very small change in pixel voltage, resulting in a two-orders-of-magnitude reduction in panel power for a 64-color image (4 levels per 1 bit memory) display. This technology will become increasingly important in displays for future mobile information devices such as e-books and personal navigation systems.

Chapter 6 describes a floating-gate MOSTFT oxide semiconductor (FLOTOS) memory that consists of wide-band-gap indium-gallium-zinc oxide (IGZO) MOSFETs and a storage capacitor. The threshold voltage variation of IGZO-TFT fabricated at a low temperature causes an increase in power consumption, which was solved by using a precharge-assisted-threshold compensation (PAC) writing technique. Furthermore, the charge stored on the capacitor is maintained for a long time without refreshing by keeping the IGZO-TFT in the off-state, similar to a conventional ROM with the floating gate completely surrounded by a dielectric. This is because the IGZO bandgap is as large as the effective barrier height for electrons at the Si/SiO_2 interface. It has been demonstrated using simulation and experimental results that an infinite number of read/write operations and excellent data retention are obtained for the proposed FLOTOS memory. It is expected that FLOTOS will become a vital technology to meet the demands for ever thinner and more lightweight mobile devices.

Finally, Chapter 7 concludes the study.

Acknowledgments

First of all I would like to thank Associate Professor Toshimasa Matsuoka of Osaka University for all his guidance and for the kindness he has shown me during the course of this study. I would also like to thank Professor Yutaka Ohmori, Professor Heiji Watanabe, Professor Koji Nakamae, and Associate Professor Yoshinari Kamakura of Osaka University for valuable guidance and for reviewing my thesis and giving me much helpful advices as my thesis committee members. I must also thank Professor Toshimichi Ito, Professor Yusuke Mori, Professor Mitsuhiro Katayama, Professor Masanori Ozaki, Professor Toshiaki Suhara, Professor Masahiko Kondow, and Professor Tetsuya Yagi of Osaka University for serving as my thesis committee members.

I would especially like to express my gratitude to Professor Emeritus Kenji Taniguchi of Osaka University for providing me with this great opportunity to study in the Integrated Functional Devices Laboratory at Osaka University and for his encouragement in helping me complete my thesis.

I would like to thank Professor Yukiharu Uraoka and Professor Yasuhiko Nakamura of Nara Institute of Science and Technology, Professor Mutsumi Kimura of Ryukoku University, and Mr. Kazuya Fujimoto of Silvaco Japan for their fruitful advice and discussions about the work described in Chapter 6.

This thesis includes the results of research conducted at Sharp Corporation. The work described in Chapters 4 and 5 was mainly accomplished at Sharp Corporation. I have enormous appreciation for the useful advice and active discussions from my colleagues at Sharp. In particular I would like to acknowledge my gratitude to Dr. Naoki Ueda who helped me to complete all the work described in this thesis. I must also appreciate Mr. Toshio Mimoto, Mr. Nobuhiko Ito, Mr. Yoshiji Ohta and Mr. Katsumi Fukumoto for helping with the design development. My appreciation also goes to Mr. Hidehito Kitakado for his fruitful discussions about TFT technologies.

I would like to express my gratitude to Dr. Seiji Kimura, Dr. Setufumi Kamuro, and Dr. Hiromitsu Takeuchi for providing me with opportunity to study the semiconductor memory technologies at Sharp Corporation. I am also grateful to Mr. Ryuichiro Miyake, Dr. Keizo Sakiyama, Mr. Kenichi Tanaka, and Mr. Jin Taketani who always encouraged me to develop new technologies for future LSI when I was in the VLSI Research Laboratories of Sharp Corporation. I would also like to thank Mr. Yoshifumi Kishida,

Dr. Akira Ishihama and Mr. Kazuhiro Moritani for their helpful advices when I was in the Advanced Technology Development Laboratories at Sharp Corporation. I would also like to thank Dr. Masatsugu Teragawa, Dr. Shigeki Imai, and Mr. Hideo Komiya for their helpful advices when I was in the Display Device Laboratories at Sharp Corporation. I would also like to express my gratitude to Mr. Mikio Okada, Mr. Kazunori Nakahara, and Mr. Graham Barker for their helpful suggestions.

Finally, I would like to thank the many researchers who participated in lively discussion with me at international conferences and symposiums.

Contents

Chapter 1 Introduction 1
I. Semiconductor Memory Technologies 1
II. Background of This Study 5
III. Purpose of This Study 7
IV. Chapter Organization 8
References 11

Chapter 2 Scalable Multi-Level-Cell Virtual-Ground Flash Memory 15
I. Introduction 15
II. Cell Structure and Process Technology 17
III. Test Chip 19
IV. Cell Performance and Reliability 21
V. Cell-to-Cell Capacitive Coupling Interference 24
VI. Multi-Level-Cell Program Algorithm 27
VII. Enhancement of Program Efficiency 37
VIII. Conclusions 39
References 40

Chapter 3 Logic Compatible Split-Gate Flash Memory 43
I. Introduction 43
II. Cell Structure and Process Technology 44
III. Operation Concept 48
IV. Study of Source-Side Hot-Electron Injection 50
V. Evaluation Result 56
VI. Conclusions 69
References 69

Chapter 4 Floating-Gate MOSFFT Nonvolatile DRAM 73
I. Introduction 73
II. Cell Structure and Process Technology 74
III. Principle of Operations 79
IV. Store Characteristics 84
V. Reliability of Floating-Gate MOSFET 87
VI. Scalable Memory Cell 89

VII. Conclusions 91
References 91

Chapter 5 Dual-Mode Liquid-Crystal-Display with Voltage-Loss-Compensation Memory in Pixel 93
I. Introduction 93
II. Pixel Structure 95
III. Refreshing Scheme 96
IV. Polarity Inversion Scheme 100
V. Dual-Mode LCD Panel 101
VI. Gray-Scale Levels 106
VII. Conclusions 107
References 108

Chapter 6 Floating-Gate MOSTFT Oxide Semiconductor Memory 109
I. Introduction 109
II. IGZO MOSTFT 111
III. Memory Cell Structure 112
IV. Principle of Cell Operation 114
V. Cell Performance 116
VI. Reliability 120
VII. Data Retention Characteristics 126
VIII. Conclusions 128
References 129

Chapter 7 Conclusions 133

Lists of Publications, Conferences and Patents 137

CHAPTER 1

INTRODUCTION

I. Semiconductor Memory Technologies 1
A. Historical Review of Semiconductor Memory Technologies 1
B. Comparison of Memories 4
II. Background of This Study 5
A. System Power 5
B. LCD Panel Power 6
C. Battery Capacity 6
III. Purpose of This Study 7
IV. Chapter Organization 8
References 11

I. SEMICONDUCTOR MEMORY TECHNOLOGIES

A. Historical Review of Semiconductor Memory Technologies

The metal-oxide-semiconductor field-effect transistor (MOSFET) is essential for today's high-density integrated circuits such as microprocessors and semiconductor memories [1.1]. The principle of the surface field-effect transistor was first proposed in the early 1930s by Lilienfeld [1.2] and Heil [1.3] and it was subsequently investigated by the Bell Laboratories group in the late 1940s [1.4]. The first working MOSFET using a thermally grown oxidized silicon structure is, which paved the way for the planar process [1.6], was demonstrated by Kahng and Atalla in 1960 [1.5]. In 1963, Hofstein and Heiman published a paper on basic device characteristics of silicon MOSFET using an aluminum material as a gate electrode [1.7]. In the same year the first complementary MOS (CMOS) circuit was proposed by Wanlass [1.8]. In 1969, polycrystalline silicon (polysilicon) gate technology, which allows the source/drain to be self-aligned to the

gate, was proposed by Kerwin [1.9]. By the end of the 1960s, MOSFETs were finally on their way to establishing their current illustrious position in the commercial semiconductor industry.

The use of MOSFETs in semiconductor memories became especially intensive in 1969 after the introduction of 256 bit static random access memory (SRAM), which used a 6 transistor based flip-flop circuit in each memory cell. The concept of charge-storage-on-capacitor memory was invented by Dennard [1.10] in 1968 and was introduced into a commercial 1 Kbit dynamic random access memory (DRAM) in the early 1970s by Intel Corporation [1.11]. This was the beginning of the semiconductor memory revolution. Magnetic core memories were used as main memory for computers in the 1960s, but were replaced by DRAM because the density and performance of DRAM are superior to magnetic core memories. Today's DRAM cell, composed of an n-channel MOSFET and a storage capacitor, was first introduced in a 4 Kbit DRAM in 1974 [1.12]. Since then, continuous technology scaling has yielded DRAM with ever higher storage capacities, scaling which continues up to this day.

The main drawback of MOSFET-based memory is its volatility due to inherent MOSFET leakage. As a result, these random access memories (RAMs) require power supplies to retain information. DRAM is used in the main system memory due to its higher density and lower bit-cost as compared with SRAM, but they need frequent read and write operations (refresh) in addition to a power supply while retaining the information.

The first sound solutions to the problems inherent in RAMs, with applicability beyond the mere read only memory (ROM) function, were the concept of floating gate and charge-trapping memory devices. The concept of using a floating gate to realize a programmable nonvolatile memory device was suggested in 1967 by Kahng and Sze [1.13]. This was the first time that the possibility of a programmable nonvolatile MOS memory device had been recognized. Writing and erasing are performed by the direct tunneling of electrons through very thin oxide layers. The metal-nitride-oxide-semiconductor (MNOS) memory device based on the concept of charge-trapping was proposed by Wegner et al. [1.14], almost simultaneously with the floating-gate memory device. In 1977, today's electrically programmable ROM (EPROM) device based on channel hot-electron injection was proposed [1.15]. Furthermore, in 1980, 16 Kbit electrically erasable and programmable ROM (EEPROM) that used the charging-trapping concept [1.16] and floating-gate concept [1.17] were reported. Many analysts projected that EEPROMs would grow into a high volume market and gradually even replace EPROM as the standard program storage

medium in microprocessor-controlled systems. However, the densities of EEPROM products have always lagged EPROM densities by one to two generations because of the more complicated structure of EEPROM cells, which prevented the growth of the EEPROM market.

To solve these issues in EEPROM, the new concept of flash EEPROM (FLASH) was proposed by Masuoka [1.18] in 1984. The term "flash" refers to the fact that the information in the whole memory can be erased very quickly at one time. The first products were merely used as substitutes for EPROM. However, by the end of the 1980s, FLASH had become the largest market in nonvolatile technology due to the highly competitive tradeoff between functionality and cost per bit. Intel Corporation introduced the first commercially available NOR FLASH chip in 1988. And then, starting in the early 2000s, Toshiba Corporation introduced the first commercial NAND FLASH chip, which enabled a lower bit cost and higher density at the expense of performance as compared with that of NOR FLASH chips. The FLASH market continues to grow rapidly as storage memory needed for smartphone, computers, music players and other consumer products. FLASH can store operating system information, the basic input/output system (BIOS), and application software. However, it cannot compete with DRAM, because its write and erase times are far too slow for normal system operation and the number of rewrite cycles is limited, as will be described in Section I.B.

Some applications require the fast read and write times afforded by SRAMs or DRAMs, but also require the memory to be nonvolatile (the information is retained when power is turned off). One such alternative is nonvolatile SRAM (NV-SRAM), which combines SRAM with EEPROM [1.19].

New material based universal memories such as ferro-electric RAM, phase change memory, resistive RAM, and magnetic RAM [1.20] – [1.26] have been studied for a long time for next generation applications, since it offers the performance of RAM with the non-volatility of ROM. However, these new memories have not been implemented into high-density storage applications, because they will require more time to improve yields in mass production due to the use of the new materials. NV-SRAM, which combines SRAM with emerging memories using new materials [1.27] – [1.29], has also been studied intensively to reduce system power.

Table 1.1

Comparison of Semiconductor Memories.

Classification		Nonvolatility	Electrical Erasability	Infinite Rewrite	Write Speed	Read Speed	Density
RAM	SRAM	No	Yes (Byte)	Yes	Very Fast	Very Fast	Low
	DRAM	No	Yes (Byte)	Yes	Fast	Fast	High
ROM	MROM	Yes	No (Circuit)	No	---	Fast	High
	EPROM	Yes	No (UV-Erase)	No	---	Fast	High
	EEPROM	Yes	Yes (Byte)	No	Slow	Fast	Low
	NOR FLASH	Yes	Yes (Block)	No	Slow	Fast	High
	NAND FLASH	Yes	Yes (Block)	No	Slow	Slow	Very High
NV-SRAM		Yes	Yes (Byte)	Yes	Very Fast	Very Fast	Low
Universal Memory		Yes	Yes (Byte)	Yes	Fast	Fast	High

B. Comparison of Semiconductor Memories

Table 1.1 is a comparison of the semiconductor memories mentioned above. They are mainly classified into volatile RAM with an infinite number of fast read/write operations and nonvolatile ROM. DRAM is used as main working memory. SRAM is used as internal processor memory to compensate for the speed gap between the main working memory and the processor, which requires smaller density and faster speed compared with main memory.

ROM is generally categorized by the erase methods. Mask programmable ROM (MROM) is a truly read-only memory, and the users cannot modify the information in the MROM because it is programmed when the integrated circuit is manufactured. The main advantage of MROM is that is significantly cheaper than any other kind of semiconductor memory, due to its very simplest cell structure and operation. However, the one-time mask cost is high and programming requires a long turn-around time. EPROM allows electrical programming. However, erasing typically requires about 20

minutes of UV light exposure. In this case, the whole memory circuit has to be re-programmed byte by byte, even if the information of only a single byte has to be changed. These drawbacks obviated in EEPROM. In this type of nonvolatile memory circuit, all operations can be performed in a byte-addressable way. Although FLASH combines the electrical programming capability and high density of EPROM with the electrical erasability of EEPROM, it cannot be erased selectively. The term "flash" refers to the fact that all the information in the memory can be erased very quickly at one time. FLASH can be further subdivided into NOR and NAND FLASH. NOR FLASH can offer fast random access similar to RAM and nonvolatility. On the other hand, although NAND FLASH cannot provide the fast random access, it allows greater storage density and lower cost per bit than NOR FLASH because of its simple cell structure.

NV-SRAM integrates SRAM and ROM technologies on the same chip. In normal operation, the CPU will read and write data to the SRAM at high speed. However, if the NV-SRAM detects that a power failure is beginning, special circuits on the chip quickly (a few milliseconds) copy the data from the SRAM section to the ROM section of the chip, thus preserving the data. When power is restored, the data is copied from the ROM back to the SRAM, and operations can continue as if there had been no interruption.

Universal memory is based on new material concepts that can offer an infinite number of fast read/writes and nonvolatility.

II. BACKGROUND OF THIS STUDY

With the dynamic growth in wireless broadband networking infrastructure, the processing speeds and memory densities of mobile applications are rapidly increasing, and this results in an increase in the system power consumption. Furthermore, liquid crystal display (LCD) panel power is also becoming a serious issue due to the emergence of new functions such as e-books and personal navigation systems that require large LCD panels that display high-resolution still-image for long periods. On the other hand, battery capacity is severely restricted by the demands for ever thinner and more lightweight mobile devices.

A. System Power

Most current mobile devices use DRAM as main working memory. DRAM allows the processor to execute operating system and application programs quickly and to process the data at high speed. The processor loads information from a storage ROM

chip into a DRAM chip and has to perform high-volume processing there, because the access speed of storage ROM is very slow. In addition, DRAM needs frequent refreshing even in standby due to the inherent leakage of silicon MOSFETs, as described in Section I.A. Furthermore, to store data being processed in the event of sudden power failure, it must be continuously transferred from the DRAM chip to a ROM chip. However, the current approach suffers from slow word-by-word serial data transfer between the DRAM and the storage ROM. Hence, the DRAM standby power and the chip-to-chip communication power between the DRAM chip and the storage ROM chip occupy a significant portion of total system power consumption and also cause an increase in latency.

B. LCD Panel Power

Each LCD pixel consists of a switching transistor in series with a storage capacitor, similarly to DRAM. Frequent refresh operations are required to hold pixel data because of the off-state leakage of the switching transistor. This refresh power occupies a significant portion of total LCD power even for still-image display.

C. Battery Capacity

Recently, a requirement for wearable devices, shown in Fig. 1.1, is increasing from the view point of mobile convenience, which can give a new added value to the current mobile devices as well as a mobile convenience. A flexible ultra-lightweight solar cell [1.30] and touch sensor [1.31] using organic TFT were proposed in 2012 and 2013, respectively. And, low temperature amorphous In-Ga-Zn Oxide (IGZO) semiconductor TFT was also proposed in 2004 [1.32]. The performance of IGZO semiconductors is significantly improved compared with that of organic TFTs and is suitable for enabling a true wearable device including all circuit elements such as the logic and memory circuits.

This wearable device must be driven by small capacity batteries such solar cells. Thus, power consumption is becoming the limiting factor for mobile device functionality, and reducing power is the key challenge for future wearable device applications.

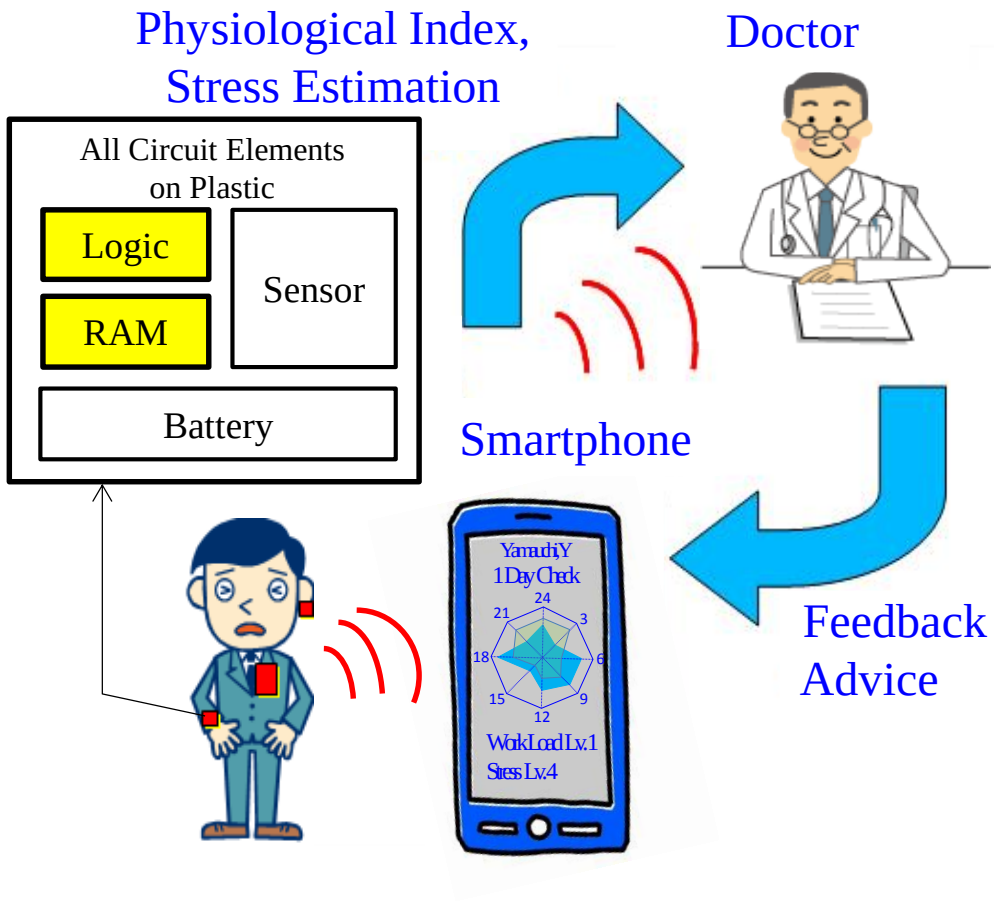


Fig. 1.1. Wearable mobile application.

III. PURPOSE OF THIS STUDY

The inherent leakage of silicon MOSFET causes a significant increase in the power consumption of the system and the LCD panel in mobile applications, and as described in Section II, this is serious issue for such devices. The purpose of this thesis is to investigate various types of semiconductor memory device and to discuss their basic operation and design concept from the perspective of solving these issues. The power consumption of main working DRAM contributes to around 50% of total power consumption in a typical internet mobile device [1.33], [1.34], and it is estimated, that main memory power can be reduced to one tenth while enhancing mobile device performance, by using the technologies described in this thesis.

IV. CHAPTER ORGANIZATION

An outline of how the chapters are organized is illustrated in Fig. 1.2.

Chapter 2 describes a scalable multi-level-cell (MLC) virtual-ground flash memory cell technology [1.35] that has potential for continued scaling of fast-access NOR flash memory. The virtual-ground cell allows the CPU to execute instructions directly without going through DRAM, that leading to reduction in DRAM refresh power consumption. The virtual-ground cell size is scaled by eliminating a contact from the standard NOR flash memory cell. However, both the low gate-coupling-ratio (GCR) and the cell-to-cell interference caused by the virtual-ground architecture make it difficult to achieve the MLC. In order to solve this, the mechanism of this interference is investigated and revealed by using neighboring cell leakage and capacitive coupling models, and a new cell structure and cell-to-cell interference canceling technique are proposed, the evaluations of these were conducted on 128 Mbit MLC test chips. A contact-less virtual-ground cell with an auxiliary gate is also proposed [1.36] to enhance the program efficiency, and this has been introduced into embedded applications as will be described in Chapter 3.

Chapter 3 describes Chapter 3 describes a source-side injection single-polysilicon split-gate NOR (S4-NOR) flash memory cell with a standard CMOS logic process [1.37] for embedded applications that require low power consumption and high performance without an increase in the cost of the logic chip, and which can reduce power consumption and latency caused by chip-to-chip communication. For the proposed cell, the access gate and the floating gate are patterned on a channel between its source and drain with a small gap length using conventional photolithography technology beyond the 100 nm generation. The hot-electrons generated by the access gate transistor at the source side are injected to the floating gate at the drain side. The mechanism of source-side channel hot-electron injection in a split-gate transistor is revealed by a modified lucky-electron model and is compared with the case of conventional channel hot-electron injection in a MOSFET. Reliability evaluation is conducted on 4 Kbit test chips. Furthermore, the effect of process scaling on performance and reliability for next generation logic devices is also discussed.

Chapter 4 describes a nonvolatile DRAM cell combining a DRAM cell with a floating-gate MOSFET [1.38], which can store all the data in DRAM to the floating gate of MOSFET when a sudden power failure occurs (power-off-store) and then recall

all the data to DRAM quickly (instant-on) without chip-to-chip communication at power-up. An innovative store concept is proposed, which uses highly efficient Fowler-Nordheim tunneling of electrons between the DRAM storage-node and the floating gate of a MOSFET. A recall concept is also proposed. The power-off-store/ instant-on operations are investigated using simulation and evaluation on 1 Mbit test chips. A scalable nonvolatile DRAM cell is also proposed for high-density applications.

Chapter 5 describes a novel voltage-loss-compensation (VLC) memory in pixel that is developed for an advanced dual-mode LCDs, and which can display still image using an ultra-low-power-consumption memory mode in addition to the normal display mode. In normal mode, the panel specification is identical to the standard transmissive display. In memory mode, VLC memory in pixel technology is proposed [1.39] to display high-resolution still-images with ultra-low-power-consumption. The proposed pixel memory technology makes it possible to refresh all the pixels at one time, which leads to reduction in power consumption because of a low refresh rate. The number of gray-scale levels is increased by using an analog gray-scale technique. The image quality and power consumption of the panel with the proposed memory are evaluated on a 3.17 inch transflective panels (320×480 pixels), and is compared with the case of a conventional panel.

Chapter 6 describes a floating-gate MOSTFT oxide semiconductor (FLOTOS) memory [1.40] that consists of wide-band-gap indium-gallium-zinc oxide (IGZO) MOSTFTs and a storage capacitor. The FLOTOS memory technology has the potential to replace all the memories except storage ROM. The write operating principle of FLOTOS memory is based on charging and discharging a storage capacitor, which can realize an infinite number of fast write/read operations. In addition, the charge stored on the capacitor can be maintained for a long time without refreshing by keeping IGZO MOSTFT in the off-state, similar to a conventional ROM with the floating gate completely surrounded by a dielectric. This is because the IGZO bandgap is as large as the effective barrier height for electrons at the Si/SiO_2 interface However, the threshold voltage deviation of IGZO MOSTFT fabricated at a low temperature causes an increase in power consumption, which increases in proportional to the square of the voltage. In order to solve this issue, the precharge-assisted-threshold compensation writing technique is proposed. The number of read/write operations and data retention are investigated using simulation and measurement.

Finally, in Chapter 7, the conclusions of this study are summarized.

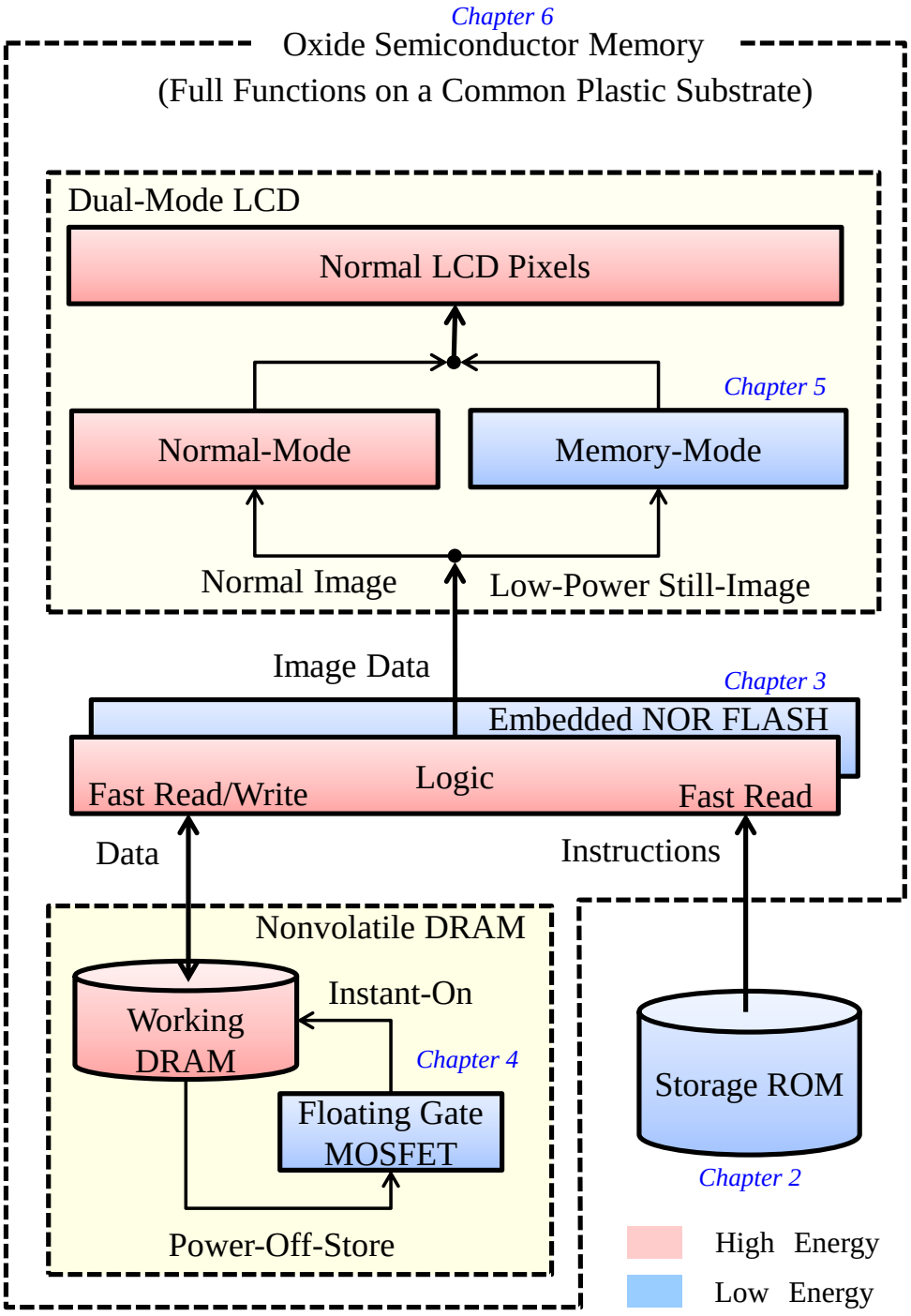


Fig. 1.2. Outline of chapter organization.

REFERENCES

[1.1] S. M. Sze, "MOSFETs", *Physics of Semiconductor Devices*, 3rd ed., Wiley, New York, 1981, pp. 293 – 294.

[1.2] J. E. Lilienfeld, United States Patent 1745175, 1930.

[1.3] O. Heil, British Patent 439457, 1935.

[1.4] W. Shockley, G. L. Pearson, "Modulation of Conductance of Thin Films of Semi-Conductors by Surface Charges," *Physical Review*, vol. 74, no. 2, pp. 232 – 233, Jun. 1948.

[1.5] D. Kahng and M. M. Atalla, "Silicon-Silicon Dioxide Field Induced Surface Devices", *IRE/AIEE Solid-State Device Research Conference*, 1960.

[1.6] J. A. Hoerni, "Planar Silicon Transistor and Diodes," *International Electron Devices Meeting*, 1960, p. 50.

[1.7] S. R. Hofstein and F. P. Heiman, "The Silicon Insulated-Gate Field-effect Transistor", *Proceedings of the IEEE*, vol. 51, no. 9, pp. 1190 – 1202, Sep. 1963.

[1.8] F. M. Wanlass and C. T. Sah, "Nanowatt Logic using Field-effect Metal-Oxide Semiconductor Triodes", *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 1963, pp. 32 – 33.

[1.9] R. E Kerwin, D. L. Klein, and J. C. Sarace, United States Patent 3 475 234, 1969.

[1.10] R. Dennard, United States Patent 3387286, 1968.

[1.11] W. M. Regitz and J. A. Karp "Three-Transistor-Cell 1024-bit 500-ns MOS RAM", *IEEE Journal Solid-State Circuits*, vol. 5, no. 5, pp.181 – 186, Oct. 1970.

[1.12] C. W. Lambrechtse , R. H. W. Salters and L. Boonstra, "A 4096 Bit One-Transistor-per-Bit RAM with Internal Timing and Low Dissipation", *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 1973, pp. 26 – 27

[1.13] D. Kahng and S. M. Sze, "A Floating Gate and its Application to Memory Devices", *Bell System Technical Journal*, vol. 46, pp. 1288 – 1295, Jul. /Aug. 1967.

[1.14] H. A. R. Wagner, A. J. Lincoln, H. C. Pao, M. R. O'Connell, and R. E. Oleksiak, "The Variable Threshold Transistor, A New Electrically Alterable, Non-Destructive Read-Only Storage Device," *International Electron Devices Meeting Technical Digest*, 1967, pp. 70.

[1.15] B. Rössler and R. Müller, "Electrically Erasable and Reprogrammable Read only Memory Using the N-channel SIMOS One-Transistor Cell," *IEEE Transaction on Electron Dev*ices, vol. 24, no. 5, pp. 606 − 610, May 1977.

[1.16] T. Hagiwara, Y. Yatsuda, R. Kondo, S. Minami, T. Aoto, and Y. Itoh, "A 16 Kbit Electrically Erasable PROM Using N-Channel Si-Gate MNOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 15, no. 3, pp. 346 – 353, Jun. 1980.

[1.17] W. Johnson, G. Perlrgos, A. Rennigner, G. Kuhn, and T. Ranganath, "A 16 Kbit Electrically Erasable Nonvolatile Memory," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 1980, pp. 152 – 153.

[1.18] F. Masuoka, M. Asano, H. Iwahashi, T. Komuro, N. Tozawa, and S. Tanaka, "A New Flash EEPROM Cell Using Triple Polysilicon Technology," *International Electron Devices Meeting Technical Digest*, 1984, pp. 464 − 467.

[1.19] E. Harari, L. Schmitz, B. Troutuman, and S. Wang, "A 256 Bit Non-Volatile Static RAM," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 1978, pp. 108 – 109

[1.20] S. Gerardin and A. Paccagnella, "Present and Future Non-Volatile Memories for Space," *IEEE Transactions Nuclear Science*, vol. 57, no. 6, pp. 3016 – 3039, Dec. 2010.

[1.21] H. S. P. Wong, S. B. Kim, B. Lee, M. A. Caldwell, J. Liang, Y. Wu, R. G. D. Jeyasingh, and S. Yu, "Recent Progress of Phase Change Memory (PCM) and Resistive Switching Random Access Memory (RRAM)," *IEEE International Conference Solid-State and Integrated Circuit Technology*, 2010, pp. 1055 – 1060.

[1.22] K. Prall, N. Ramaswamy, W. Kinney, K. Holtzclaw, X. Chen, J. Strand, R. Bez, "An Update on Emerging Memory: Progress to 2Xnm," *IEEE International Memory Workshop*, 2012.

[1.23] A. Makarov, V. Sverdlov, and S. Selberherr, "Emerging memory technologies: Trends, challenges, and modeling methods," *Microelectronics Reliability*, vol. 52. No. 4, pp. 628 – 634, Apr. 2012.

[1.24] Y. Fujisaki, "Review of Emerging New Solid-state Nonvolatile Memories," *Japanese Journal of Applied Physics*, vol. 52, pp. 040001-1 – 040001-11, 2013.

[1.25] A. Baumann, M. Jung, K. Huber, M. Arnold, C. Sichert, S. Schauer, and R. Brederlow, "A MCU Platform with Embedded FRAM Achieving 350nA Current Consumption in Real-Time Clock Mode with Full State Retention and 6.5μs System Wakeup Time," *Symposium on VLSI Circuits Digest of Technical Pap*ers, 2013, pp. C202 – C203.

[1.26] R. Annunziata, P. Zuliani, M. Borghi, G. Sandre, L. Scotti, C. Prelini, M. Tosi, I. Tortorelli, and F. Pellizzer, "Phase Change Memory Technology for

Embedded Nonvolatile Memory Applications for 90nm and beyond," *International Electron Devices Meeting Technical Digest*, 2009, pp. 97 – 100.

[1.27] T. Endoh, T. Ohsawa, H. Koike, T. Hanyu, and H. Ohno, "Restructuring of Memory in Coming System with Spintronics-Based Technologies," *Symposium on VLSI Technology Digest of Technical Papers*, 2012, pp. 89 – 90.

[1.28] T. Osawa, H. Koike, S. Miura, H. Honjo, K. Kinoshita, S. Ikeda, T. Hanyu, H. Ohno, and T. Endoh, "A 1 Mb Nonvolatile Embedded Memory Using 4T2MTJ Cell With 32 b Fine-Grained Power Gating Scheme," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 6, pp. 1511 – 1520, Jun. 2013.

[1.29] P. F. Chiu, M. F. Chang, C. W. Wu, C. H. Chuang, S. S. Sheu, Y. S. Chen, and M. J. Tsai "Low Store Energy, Low VDDmin, 8T2R Nonvolatile Latch and SRAM with Vertical-Stacked Resistive Memory (Memristor) Devices for Low Power Mobile Applications," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 6, pp.1483 –1496, Jun. 2012.

[1.30] M. Kaltenbrunner, T. Sekitani, J. Reeder, T. Yokota1, K. Kuribara1, T. Tokuhara, M. Drack, R. Schwodiauer, I. Graz, S. B. Gogonea, S. Bauer, and T. Someya, "An Ultra-Lightweight Design for Imperceptible Plastic Electronics," *Nature*, vol. 432, pp. 458 – 463, Jul. 2013.

[1.31] M. Kaltenbrunner, M. S. White, E. D. Glowacki, T. Sekitani , T. Someya, N. S. Sariciftci, and Siegfried Bauer, "Ultrathin and Lightweight Organic Solar Cells with High Flexibility," *Nature Communications*, vol. 3, pp. 1 – 7, Apr. 2012.

[1.32] K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano, and H. Hosono, "Room-Temperature Fabrication of Transparent Flexible Thin-Film Transistors Using Amorphous Oxide Semiconductors," *Nature*, vol. 432, pp. 488 – 492, Nov. 2004.

[1.33] T. Kawahara, H. Takahashi, and H. Ohno, "SPRAM (Spin-transfer torque RAM) Technology for Green IT World," *IEEE International Conference on Electron Devices and Solid-State Circuits*, 2008.

[1.34] T. Kawahara, "Scalable Spin-Transfer Torque RAM Technology for Normally-Off Computing," *IEEE Design and Test of Computers*, vol. 20, no. 1, pp. 52 – 63, Jan./Feb. 2011.

[1.35] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "Scalable Virtual-Ground Multi-Level-Cell Floating-Gate Flash Memory," *IEEE Transactions on Electron Devices*, vol. 60, no. 8, pp. 2518 − 2524, Aug. 2013.

[1.36] Y. Yamauchi, K. Tanaka, H. Shibayama, and R. Miyake, "A 5V-Only Virtual Ground Flash Cell with An Auxiliary Gate for High Density and High Speed Application," *International Electron Devices Meeting Technical Digest*, 1991, pp. 319 − 322.

[1.37] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "Source-side Injection Single-Polysilicon Split-gate Flash Memory," to be published in *Japanese Journal of Applied Physics*, vol. 43, Mar. 2014.

[1.38] Y. Yamauchi, K. Tanaka, K. Sakiyama, and R. Miyake, "A Versatile Stacked Storage Capacitor on FLOTOX Cell for Megabit NVRAMs," *IEEE Transactions Electron Devices*, vol. 39, no. 12, pp. 2791 – 2796, Dec. 1992.

[1.39] Y. Yamauchi, N. Ueda, Y. Ogawa, K. Tanaka, and K. Yamamoto, "A Novel Pixel Memory using Integrated Voltage-Loss-Compensation (VLC) Circuit for Ultra-Low-Power TFT-LCDs," *Journal of the SID*, vol. 19, pp. 57 – 62, Jan. 2011.

[1.40] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "Study of Novel Floating-Gate Oxide Semiconductor Memory Using Indium-Gallium-Zinc Oxide for Low-Power System-on-Panel Applications," *Japanese Journal of Applied Physics*, vol. 52, pp. 094101-1 – 094101-7, Aug. 2013.

Chapter 2

Scalable Multi-Level-Cell Virtual-Ground Flash Memory

I. Introduction 15
II. Cell Structure and Process Technology 17
III. Test Chip 19
A. Array Architecture 19
B. On-chip Circuits 21
IV. Cell Performance and Reliability 21
A. Cell Performance 21
B. Cell Reliability 21
V. Cell-to-Cell Capacitive Coupling Interference 24
VI. Multi-Level-Cell Program Algorithm 27
A. Program Operation 28
B. Cell-to-Cell Interference in BL Direction 28
C. Cell-to-Cell Interference in WL Direction 31
D. Experimental Results 35
E. Program Throughput 35
VII. Enhancement of Program Efficiency 37
A. Cell Structure 37
B. Principle of High Efficient Programming 38
C. Program Performance 38
VIII. Conclusion 39
References 40

I. INTRODUCTION

Mobile consumer-electronic device functionality is rapidly increasing with the substantial progress made in wireless broadband network infrastructures. However, battery capacity is severely restricted due to constraints on the size and weight of devices.

Therefore, system power consumption must be reduced to extend the battery life. Memory is a significant contributor to the total system power consumption. Most high-end mobile devices use a combination of dynamic random access memory (DRAM) and serial-access NAND flash memory. This memory structure significantly increases power for DRAM refresh operations and degrades performance, because a program stored in the NAND flash memory must be transferred to DRAM before being executed [1.33], [1.34].

Emerging memories such as ferro-electric RAM, phase change memory, resistive RAM, and magnetic RAM [1.20] − [1.26] have been studied for next generation applications, since they offer the performance of RAM with the nonvolatility of read only memory (ROM). However, these nonvolatile RAMs have not been implemented into the mainstream market.

The NOR flash memory architecture continues to cope with the wireless mobile application requirements of faster execute-in-place performance and faster programming throughput, as well as higher density. Flash memory density has been increased through a combination of multi-level-cell (MLC) approaches [2.1] − [2.5] and physical scaling [2.6] − [2.8]. MLC storage in flash memory using the well established polycrystalline-silicon (polysilicon) floating gate (FG) transistor was first reported in 1995 [2.6] to realize double density with the same process technology, and is acceptable for any given process technology generation [2.7] − [2.10]. A virtual-ground flash memory architecture [2.9] has been proposed for significant scaling down of the cell size. However, scaling of the memory cell causes FG-to-FG coupling interference, which is well known as the critical scaling barrier for MLC NAND flash memory [2.10], [2.11].

This chapter describes a novel cell that employs a bowl-shaped FG structure for enhancing the gate coupling ratio (GCR), together with a new cell-to-cell interference canceling technique. These are combined with a two-step programming scheme to reduce FG-to-FG interference effect on threshold shift in the bit-line (BL) direction with the previously proposed channel hot electron injection threshold voltage (V_{th}) compensation scheme to reduce that in the word-line (WL) direction [1.35] , [2.12], for further scaling of a virtual-ground MLC FG memory. In Section II, the process technologies for the scaled virtual-ground cell with high GCR are proposed. In Section III, array architecture and on-chip circuits for an MLC test chip are described. In Section IV, memory cell performance and reliability are evaluated. In Section V, the effect of FG-to-FG coupling on the threshold shift is analyzed. In Section VI, the MLC program algorithm is discussed. Finally, conclusions are summarized in Section VII.

II. CELL STRUCTURE AND PROCESS TECHNOLOGY

Figure 2.1 shows a comparison of the cell structure between standard NOR flash memory and virtual-ground flash memory. A virtual-ground cell structure is simplified by avoiding the need for shallow trench isolation (STI) and replacing the drain contact and source rail across STI [2.6], [2.7] with buried BL diffusion. However, a virtual-ground flash memory cell suffers from a low GCR [2.9], [2.13], because it uses the short-side edge of the FG along its channel width (W_g) direction as the surface area of an interpoly capacitor, while a conventional NOR cell uses the long-side edge of FG along its channel length (L_g) direction. A high GCR is required for reduction of FG-to-FG coupling interference and for low voltage operation, as will be described later. The bowl-shaped FG structure is introduced to enhance the GCR without using a high-k dielectric material.

The key process steps for the novel cell with the bowl-shaped FG structure are shown in Fig. 2.2. The tunnel dielectric (TD) is grown first, followed by deposition of the thin film first level polysilicon (poly-1), high-temperature low pressure chemical vapor deposition oxide (HTO) and nitride. After patterning the long continuous poly-1 film, arsenic ions are implanted to form the diffusion BL (step-1). The poly-1 gap on the BL is filled with a high-density plasma CVD oxide (BL-Ox) and planarization is subsequently achieved by chemical mechanical polishing (CMP) to expose the nitride surface (step-2). The nitride acts as the CMP stopper. After removing the SiN and HTO on the poly-1 film, second level polysilicon (poly-2) deposition is followed by CMP planarization to form the bowl-shaped FG structure in a self-aligned manner (step-3). Finally, an interpoly oxide-nitride-oxide (ONO) stacked dielectric, and plysilicon/tungsten-silicide are deposited on the stacked floating gate and are patterned to form the self-aligned stacked MOS gate (step-4(a)). This process technology can avoid the remnant polysilicon stringer on the side-wall of BL-Ox over the field region between the neighboring FGs, due to the tapered profile of the BL-Ox film (step-4(b)).

A small cell size of 2*F* (BL direction) by 3*F* (WL direction) is achieved with a GCR of 0.6 using these technologies, where *F* is the minimum feature size. The cell size of the proposed scaled virtual-ground flash memory is approximately one half that of the conventional NOR flash memory.

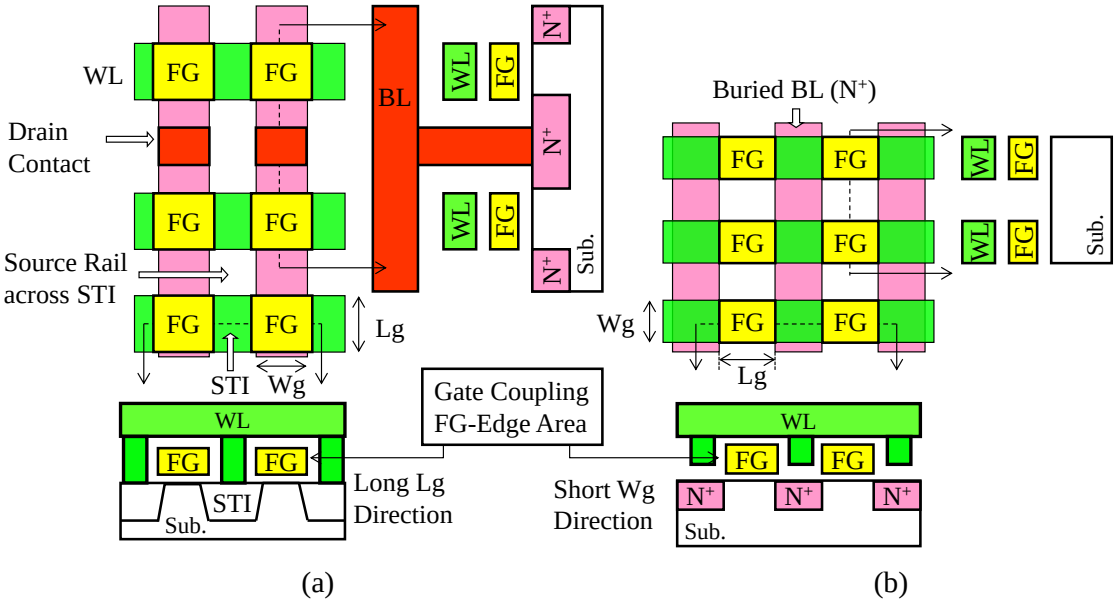


Fig. 2.1. Top view comparison of (a) standard NOR array and (b) virtual-ground NOR array with cross-sections parallel to the WL direction over the active area and perpendicular to the WL over the field region [1.35].

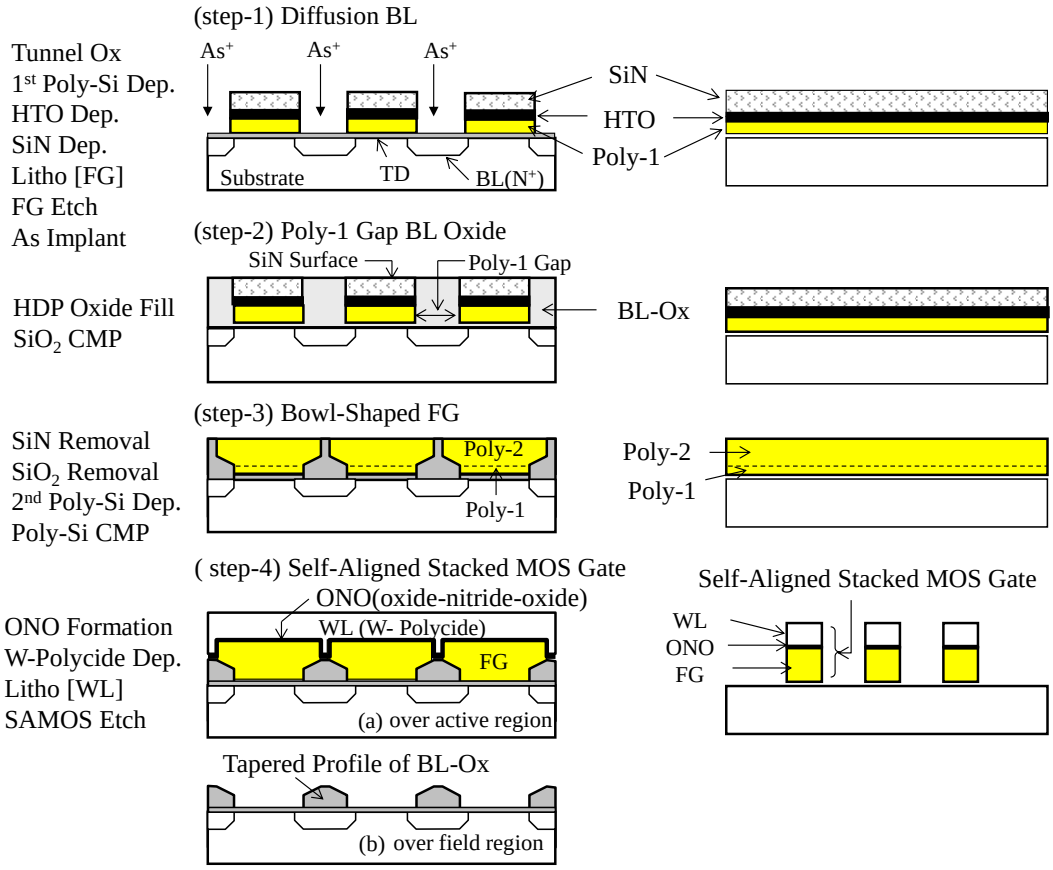


Fig. 2.2. Key process steps with cross-sectional views in the WL (left) and BL directions (right) for a virtual-ground flash memory array with a bowl-shaped FG structure [1.35].

III. TEST CHIP

A. Array Architecture

The proposed cell technology is implemented into a 128 Mbit MLC test chip using 0.13 μm rules. Figure 2.3 shows a photograph of the chip together with cross-sectional SEM images of the memory cell in the word-line (WL) and BL directions. The memory array in the chip is divided into 4 planes and each plane is divided into 32 blocks (1 Mb/block) in the BL direction. Each memory block is then divided into 64 sub-blocks (16 Kb/sub-block) in the WL direction. These blocks have individual read and program verification sensing amplifiers. Plural memory cells belonging to the different sub-blocks on the same WL can be programmed and sensed simultaneously to reduce the program time per bit, similarly to standard NOR flash memory.

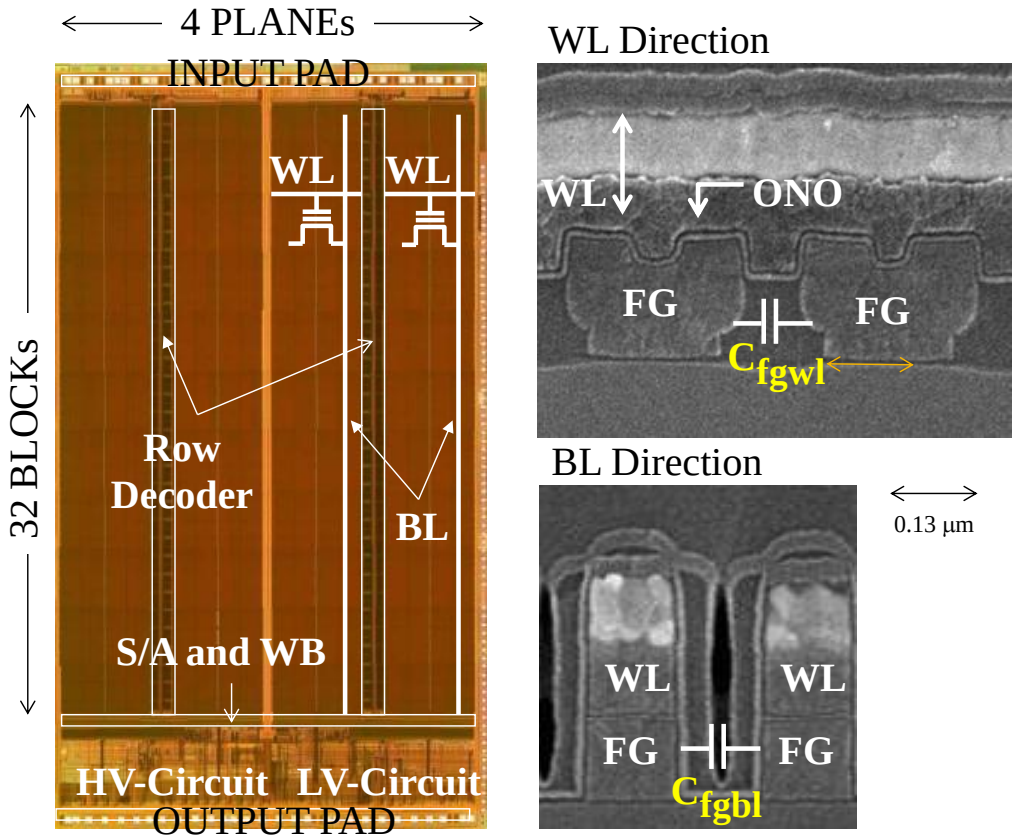


Fig. 2.3. Photograph of the test chip (left) and cross-sectional SEM images (right) of the memory cell over the active region in the WL direction and over the field region in the BL direction [1.35].

The array architecture for a sub-block is depicted in Fig. 2.4. One sub-block is further divided into 16 segments (1 Kb/segment). Furthermore, the neighboring BLs that belong to the different segments are electrically isolated by trench isolation in the WL direction.

A unit segment consists of 32 word-lines (WLs) by 17 diffusion BLs (DBLs). Each DBL has a BL contact (BC) with a metal sub-BL (SBL) every 32 WLs. One metal main BL (MBL) is shared by two SBLs of neighboring segments via two selection transistors (SG1 and SG2) to form a hierarchical structure suitable for physical scaling. One end of the BLs in each segment is connected to the source-line (SL), which is grounded in read and program operation modes.

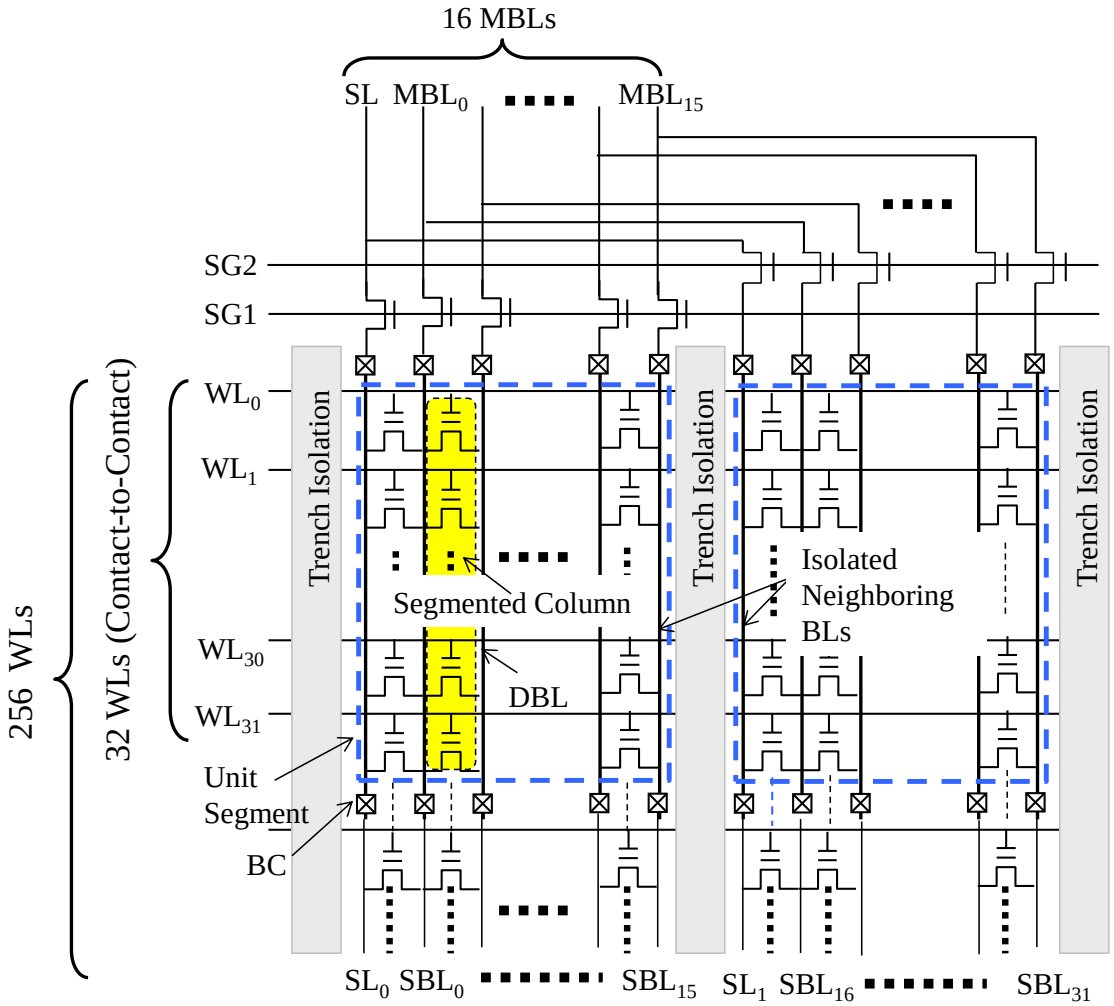


Fig. 2.4. Proposed virtual-ground architecture for a sub-block (16Kb) [1.35].

B. On-chip Circuits

This chip has charge-pump multipliers to generate high voltages from a low voltage supply (1.8 V) for programming, erasing, and reading. A write buffer to store program data (WB), a sense-amplifier (S/A) to verify program states, and a write state machine (WSM) to execute program sequences are also integrated on the test chip. These circuits are composed of high performance low-voltage (LV) transistors and high reliability high-voltage (HV) transistors.

IV. CELL PERFORMANCE AND RELIABILITY

A. Cell Performance

The program and erase characteristics were measured for a typical memory cell centered between contacts in the array at fixed bias voltages, and Fig. 2.5 shows the threshold (V_{th}) as a function of program time (t_p) and erase time (t_e). Channel hot electron (CHE) injection into the FG from the drain-edge of the channel is used to program the cell to the higher V_{th}, and erasing brings the cell to its lowest V_{th} by Fowler-Nordheim tunneling electron extraction from the FG to the substrate through the tunnel oxide. The typical values of the highest-threshold program level and the erase level are 7 V and 2.5 V, respectively. The bowl-shaped FG structure cell enables a program time of 1 μs to be achieved with voltages of V_d/V_{cg} = 5.5/9.5 V for the highest-V_{th} programmed state, and an erase time of 50 ms is achieved with voltages of $V_{sub} - V_{cg}$ = 18 V, where V_d is the drain voltage, V_{cg} is the control gate voltage and V_{sub} is the substrate voltage.

B. Cell Reliability

The program/erase endurance and data retention are the key reliability issues for a FG flash memory. This reliability test was carried out using 32K cells for each stress condition.

1) *Endurance*

The effects of cycling on program time (t_p), erase time (t_e), and read current (I_{read}) are shown in Fig. 2.6. t_p, t_e and I_{read} were measured after each cycle for a typical cell. No degradation in t_p was observed up to 10^4 cycles and t_p increased from 10^4 to 10^6 cycles. t_p is degraded to approximately 3 times the initial value after 10^5 cycles. No reduction of t_e was observed up to 10^6 cycles. In addition, I_{read} degradation was as low as 10%, even after 10^6 cycles.

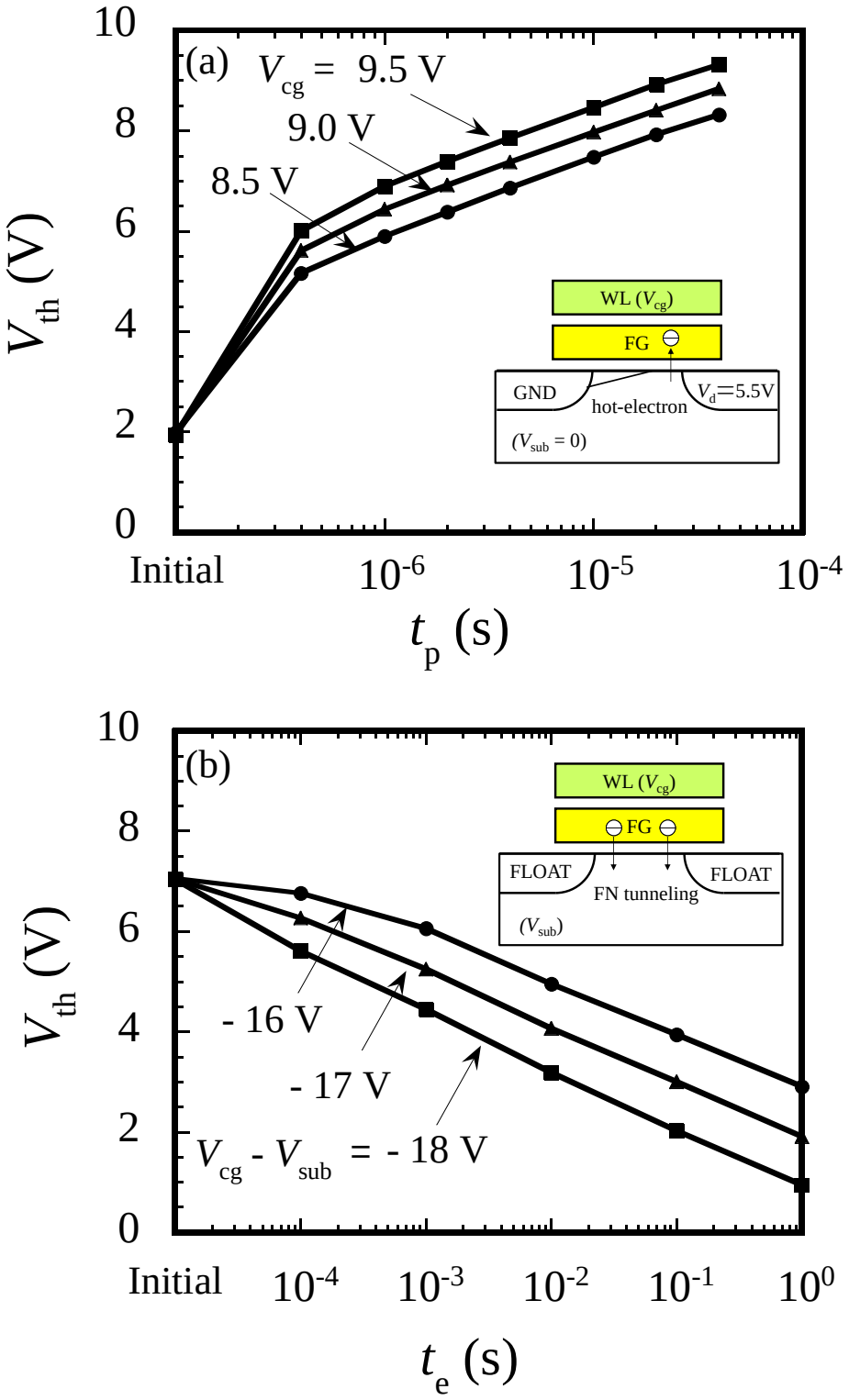


Fig. 2.5. Threshold voltage as a function of (a) program time (t_p) and (b) erase time (t_e) for a typical cell [1.35].

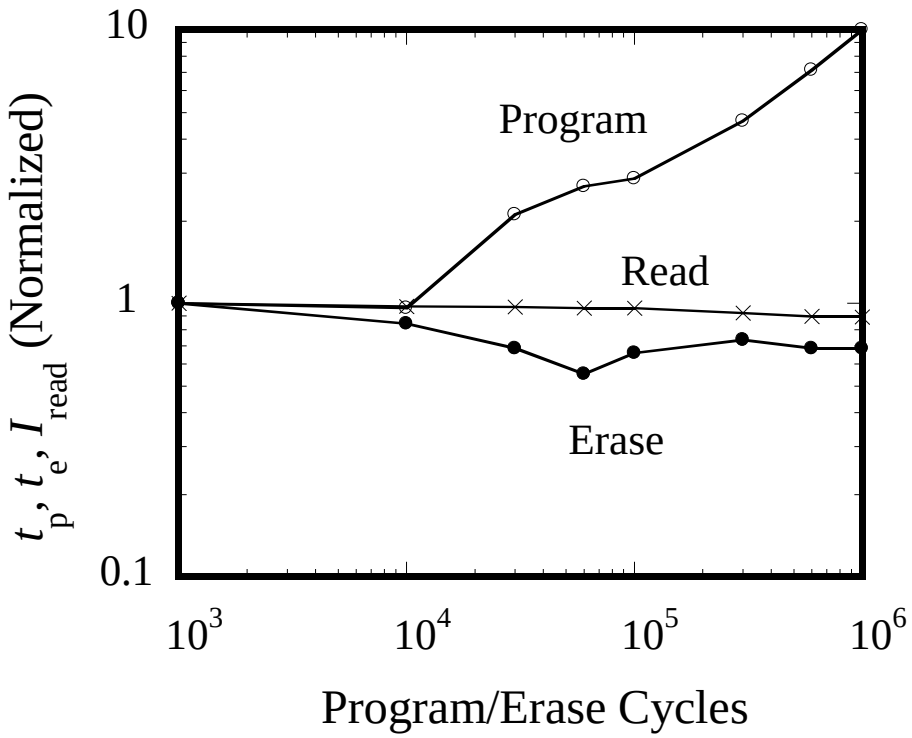


Fig. 2.6. Effect of program/erase cycle number on program time (t_p), erase time (t_e) and read current (Iread) for a typical cell at room temperature [1.35].

2) *Retention:*

Figure 2.7 shows the medium value and the standard deviation (σ) of the V_{th} distribution as a function of bake time at 250 °C for different numbers of program/erase cycles between 1 and 10^5. The charge loss rate increased with the cycle number. Charge loss involves two principal mechanisms. The charge loss rate in the initial stage (within 10 min) is ascribed to the detrapping of electrons trapped in the oxide layer or interface, which is strongly dependent on the cycle number. The charge loss rate in the second stage (after 100 min) is attributed to the charge loss from the FG, which is less dependent on the cycle number than that at the initial stage. The long-term charge loss behavior at 150 °C can be extrapolated from measurements at 250 °C for a typical cell using a single thermal activation energy of 1.4 eV [2.14]. Thus, a bake time of 100 h at 250 °C corresponds to a 10 year bake at 150 °C. Good data retention was achieved up to 100 k cycles.

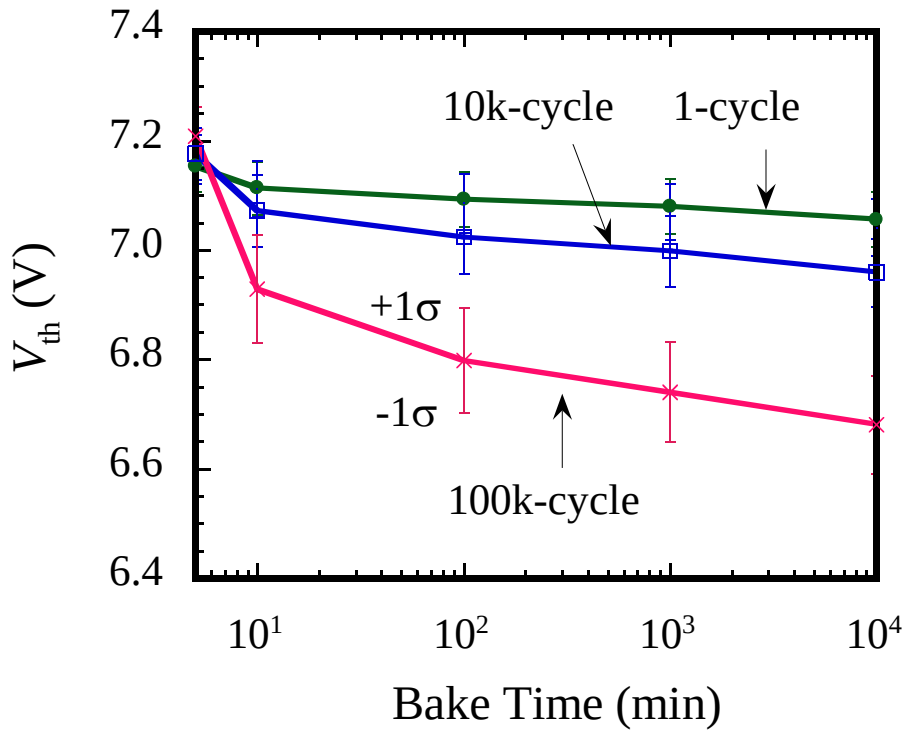


Fig. 2.7. Median value and standard deviation of the V_{th} distribution as a function of bake time at 250 °C after 1, 10k, and 100k cycles [1.35].

V. CELL-TO-CELL CAPACITIVE COUPLING INTERFERENCE

As the cell size of the flash memory is scaled down, the V_{th} shift caused by FG-to-FG interference through the capacitance of C_{fgwl} in the WL direction and C_{fgbl} in the BL direction (Fig. 2.3) becomes a serious problem. The V_{th} shift caused by cell-to-cell interference is measured using the test mode and the results are shown in Fig. 2.8, where BL_m is the BLof m-th column and WL_n is the word line of n-th row. V_{th} is defined as the WL voltage when the read current sent to the sensing amplifier (I_{sen}) is 10 μA. The influence of the neighboring cell leakage in a virtual-ground array [2.15] can be eliminated in test mode, because the two BLs, BL_{m+1} and BL_{m-2}, which are kept floating, are charged or discharged through the neighboring cell until the corresponding leakage currents ($I_{leak(d)}$ and $I_{leak(s)}$) become negligibly small compared to the current (I_{cell}) of the cell (m) being read, as shown in Fig. 2.9. This test-mode was prepared only for experimental characterization purposes.

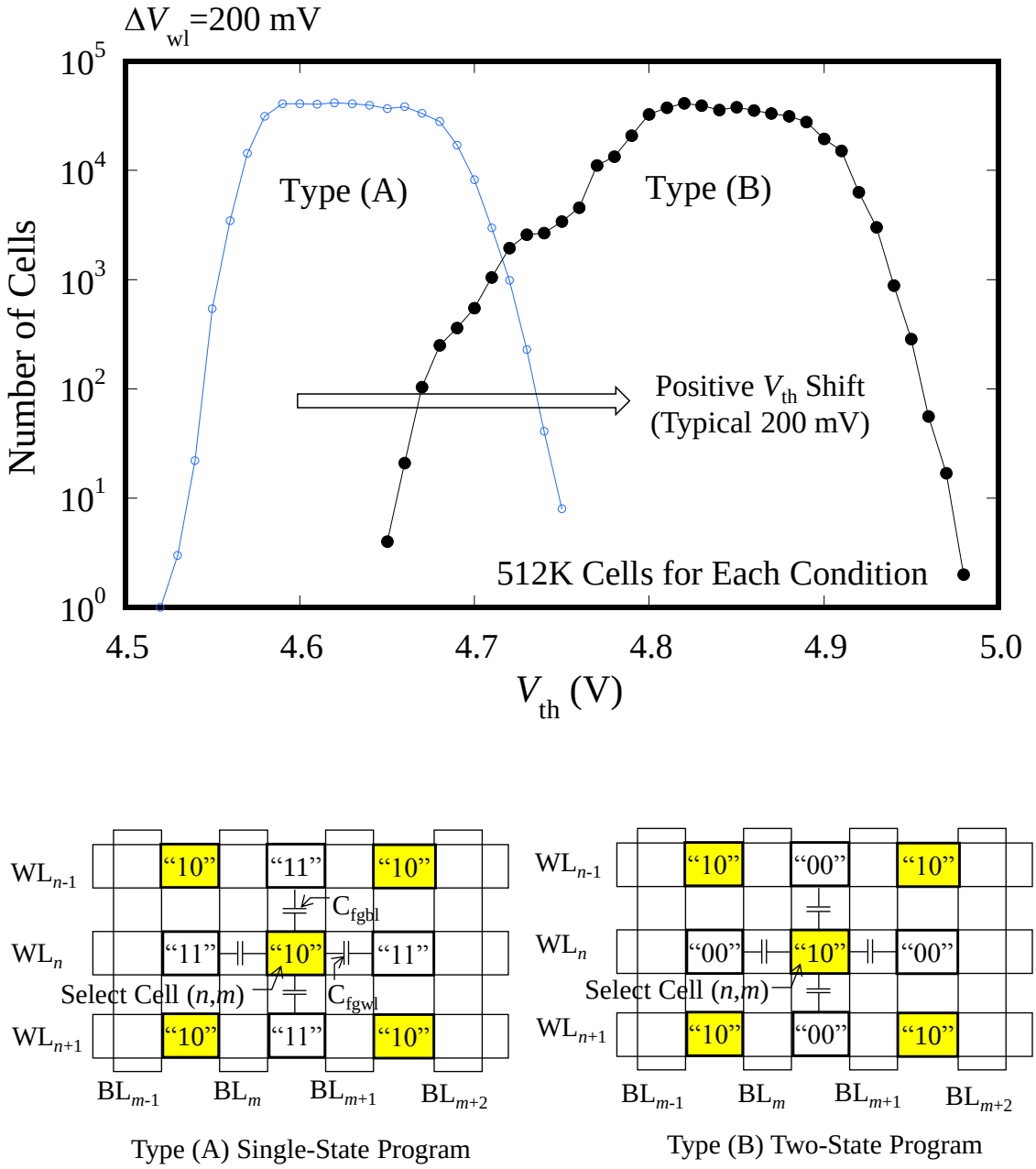


Fig. 2.8. Programmed V_{th} distribution of the selected cells ("10" state) for type (A) and type (B). Two-state programming with type (B) was carried out starting from the "10" state [1.35].

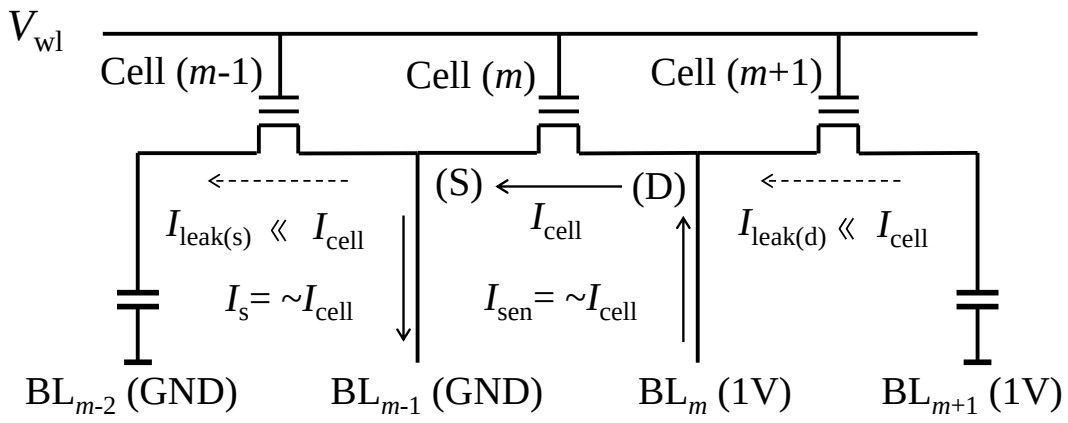


Fig. 2.9. Read operation for the selected cell (m) in test-mode [1.35].

After all memory cells in a block (512K cells) are erased to the "11" state (typical V_{th}=2.80 V), half of the memory cells are programmed to the "10" state (typical V_{th}=4.62 V) with the type (A) pattern. The other half of the memory cells are then programmed to the "00" state (typical V_{th}=7.08 V) with the type (B) pattern. Even if the selected cell is precisely programmed to the "10" state, the V_{th} of the selected cell shifts positively through the FG-to-FG coupling by programming the neighboring cells from the "11" state to the "00" state.

The V_{th} distribution of the previously programmed "10" state cells increases by 200 mV from the typical value and is widened by 100 mV after programming the neighboring cells from the "11" state to the "00" state. This typical value of V_{th} shift is in good agreement with that calculated (223 mV) from the total FG-to-FG coupling ratio of 5.2%, which breaks down into 3.2% through two sides of C_{fgbl} and 2.0% through two sides of C_{fgwl}. This typical V_{th} shift is estimated to be mainly due to the increase of FG-to-FG coupling interference.

VI. MULTI-LEVEL-CELL PROGRAM ALGORITHM

Figure 2.10 shows the target program-verify levels used in this work for 4-level MLC flash memory. The separation width between the erased state "11" and the lowest-threshold programmed state "10" is as large as 1 V in consideration of the V_{th} shift (less than 300 mV from Fig. 2.7) due to FG-to-FG coupling interference. However, the voltage width between the program verify levels is also 1V that includes the V_{th} distribution width and the separation width between neighboring program states. A reduction in widening of the V_{th} distribution of memory cells in the "10" state and "01" state due to FG-to-FG coupling interference is required to satisfy the target V_{th} distribution for a scalable virtual-ground MLC flash memory. A new algorithm is investigated using an MLC test chip.

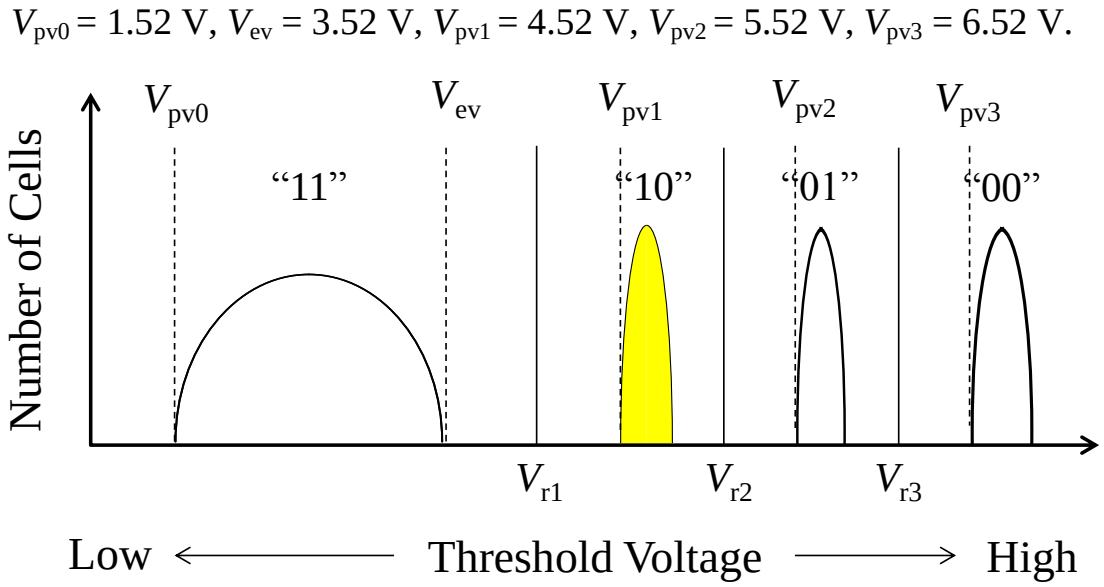


Fig. 2.10. Target V_{th} distributions for 4-level (states "11", "10", "01", and "00") MLC NOR flash memory. Reference voltages are used for erase-verify (V_{ev}), post-erase-program-verify (V_{pv0}), three levels of program-verify (lowest-threshold programmed level of V_{pv1}, middle-threshold programmed level of V_{pv2} and highest-threshold programmed level of V_{pv3}), and three levels of read-verify (V_{r1}, V_{r2} and V_{r3}) [1.35].

A. Program Operation

A buffer program scheme is used to improve the program throughput (PT). Firstly, program data is written into the on-chip write buffer and the buffer data is then programmed into the flash memory array. In this study, programming was performed in parallel on 8 memory cells (16 bits) and program-verification was performed in parallel on 64 memory cells (128 bits).

B. Cell-to-Cell Interference in BL Direction

The cell-to-cell interference in the BL direction is mainly due to the effect of C_{fgbl} coupling. A segment-by-segment recursive programming scheme is proposed to reduce the cell-to-cell interference effect, as shown in Figs. 2.11(a) and 12(a), which shows the program sequence for 32 cells in a segmented column (Fig. 2.4). Programming is carried out with a program gate voltage (V_{wl}) on each WL and a program drain voltage on the selected BL from the first row to the last row without verification, with subsequent verification carried out to determine whether the program is completed or not. Programming and verification are continued with an incremental gate voltage (ΔV_{wl}), only for those cells that have not reached the target program-verify level, until programming is completed [2.16]. Once a cell is programmed completely, no additional program pulse is applied to prevent the lower-threshold programmed state cells from being over-programmed with a high gate-voltage. In addition to the above-mentioned segment-by-segment recursive programming scheme, a two-step programming scheme is proposed for further reduction of the cell-to-cell interference effect, as shown in Fig. 2.13(a). The "00" state is firstly programmed with a large ΔV_{wl}, due to negligible influence of over-program and disturb issues. The "10" and "01" states are subsequently programmed recursively with a small ΔV_{wl} to control V_{th} accurately. These approaches are useful for the virtual-ground cell due to its large coupling interference as described in Section VI.E.

The worst case pattern that causes the largest cell-to-cell interference differs according to the program algorithm, as shown in Figs. 2.12(b) and 2.13(b). If the neighboring cells [cell $(n-1)$ and cell $(n+1)$] are programmed from the "11" state to the "00" state after program completion of cell (n), then the amount of V_{th} shift of the neighboring cells are as large as approximately 4 V. However, the amount of V_{th} shift of the neighboring cells is reduced to 2 V with the segment-by-segment recursive programming scheme, because the neighboring cells are temporary programmed close to the "10" state when programming of cell (n) is completed. Furthermore, the amount of V_{th} shift of the neighboring cells is reduced to 1 V by combining the two-step programming scheme with the segment-by-segment recursive programming scheme, because there is no need to

program the “00” state neighboring cells in the second step. As a result, the effect of cell-to-cell interference on the V_{th} shift in the BL direction can be reduced to 25% using the new two-step programming scheme.

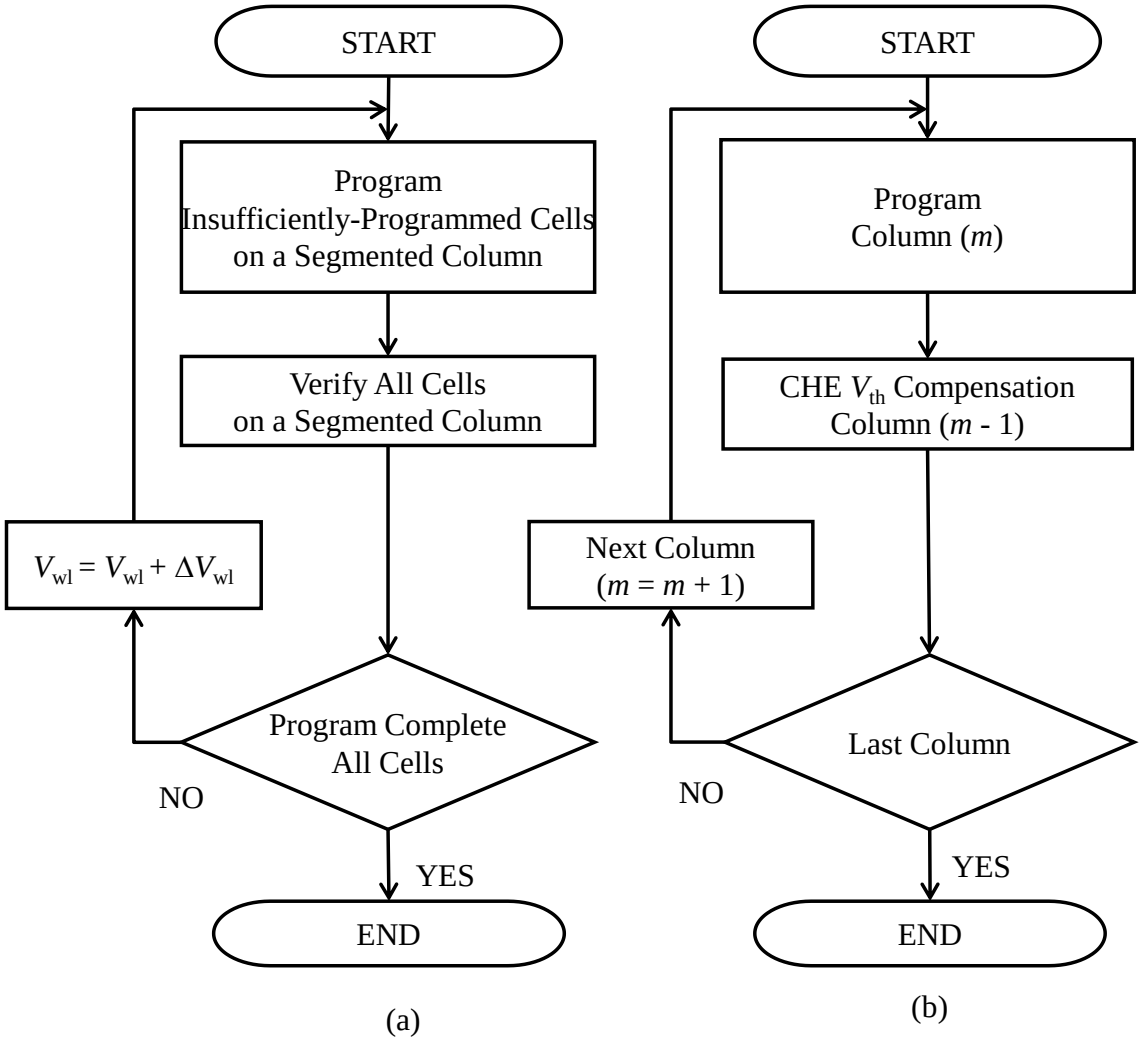


Fig. 2.11. Program and CHE V_{th} compensation sequence for (a) segment-by-segment recursive programming in the BL direction and (b) sequential programming with CHE V_{th} compensation in the WL direction [1.35].

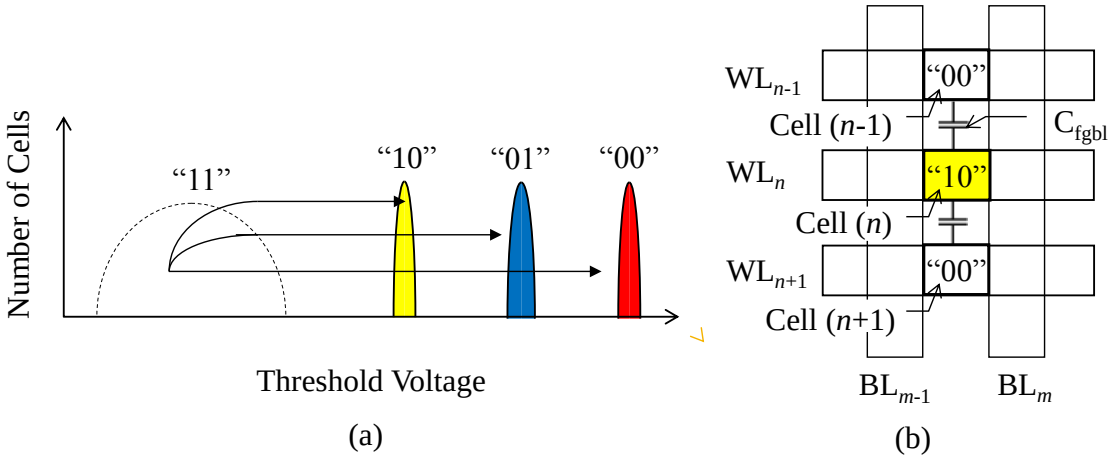


Fig. 2.12. (a) Segment-by-segment recursive programming in the BL direction. (b) Worst case cell-to-cell interference pattern [1.35].

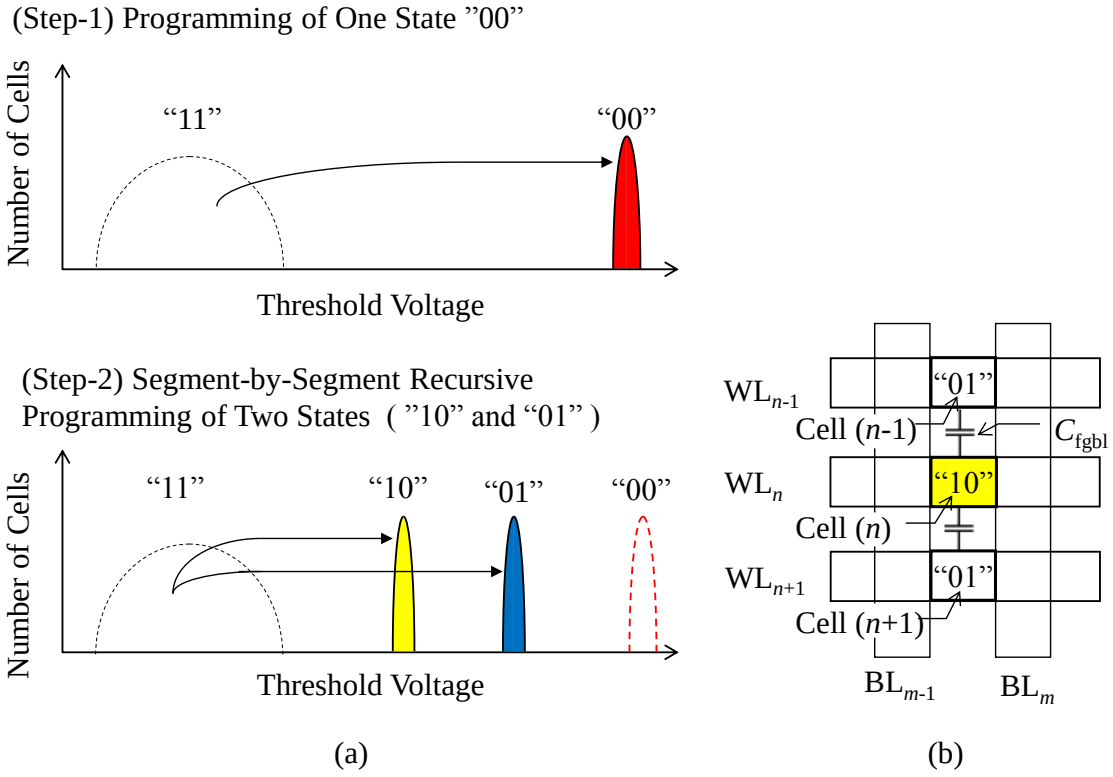


Fig. 2.13. (a) New two-step programming in the BL direction. (b) Worst case cell-to-cell interference pattern [1.35].

C. Cell-to-Cell Interference in WL Direction

The cell-to-cell interference in the WL direction is mainly due to the effect of C_{fgwl} coupling and the neighboring cell leakage ($I_{leak(d)}$ and $I_{leak(s)}$). The V_{th} shift caused by interference from the source-side and drain-side neighboring cells was evaluated under normal operating conditions using the test circuit shown in Fig. 2.14. To realize fast access, the drain-side neighboring BL (BL_{m+1}) is biased with a voltage similar to that of the selected drain BL (BL_m). The V_{th} shift due to C_{fgwl} coupling was obtained using the test-mode and the V_{th} shift due to the neighboring cell leakage was calculated from the difference in the V_{th} shift between the normal-mode and test-mode.

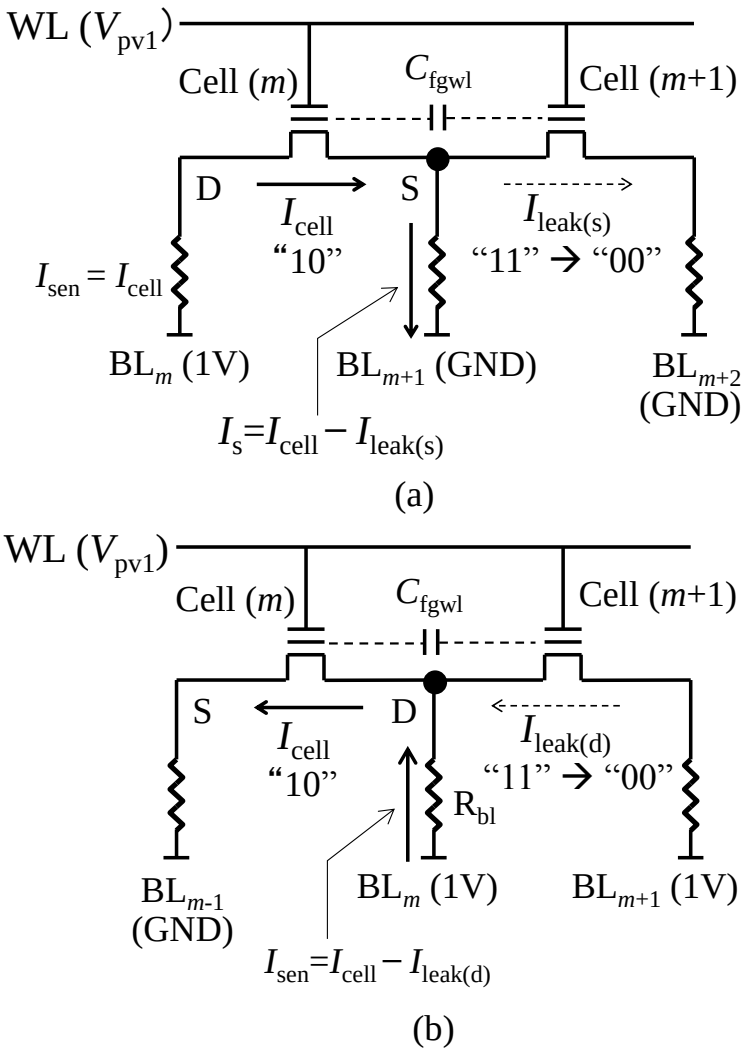


Fig. 2.14. Equivalent circuits for the test device with bias condition to evaluate cell-to-cell interference by programming the (a) source-side and (b) drain-side neighboring memory cells (m + 1) [1.35].

1) *Interference from the Source-Side Neighboring Cell*

The V_{th} of the previously programmed cell is affected by the V_{th} change of the source-side neighboring cell. Figure 2.15(a) shows ΔV_{th} as a function of $V_{th,src} - V_{pv1}$, where ΔV_{th} is the V_{th} shift of cell (m) and $V_{th,src}$ is the V_{th} of the source-side neighboring cell ($m + 1$).

The positive ΔV_{th} in normal-mode increases up to 75 mV with the increase of $V_{th,src}$ from $V_{pv1} - 2$ V to $V_{pv1} + 3$ V, which becomes significantly large with further scaling. The total amount of positive ΔV_{th} breaks down into 25 mV due to the $I_{leak(s)}$ effect and 50 mV by the C_{fgwl} coupling effect. Programming of the source-side neighboring cell ($m + 1$) causes an increase in the voltage (V_s) at the source terminal (S), because I_s increases due to decrease of $I_{leak(s)}$, as shown in Fig. 2.14(a). The increase in V_s directly increases the V_{th} of cell (m) which is defined as the WL voltage to obtain the read current of 10 μA, and the body-effect caused by the V_s increases it further.

2) *Interference from the Drain-Side Neighboring Cell*

The V_{th} of the previously programmed cell is affected by the V_{th} change of the drain-side neighboring cell, because V_{th} is defined as V_{wl} required to obtain the desired value of the read current of I_{sen} ($= I_{cell} - I_{leak(d)}$) sent to the sensing amplifier in normal mode, as shown in Fig. 2.14(b). Figure 2.15(b) shows ΔV_{th} as a function of $V_{th,drn} - V_{pv1}$, where ΔV_{th} is the V_{th} shift value of cell (m) and $V_{th,drn}$ is the V_{th} of the drain-side neighboring cell ($m + 1$). The ΔV_{th} by the $I_{leak(d)}$ effect increases in the negative direction with the increase of $V_{th,drn}$ and is saturated with the $V_{th,drn}$ above $V_{pv1} + 2$ V, because $I_{leak(d)}$ decreases with the increase of $V_{th,drn}$ and then becomes negligibly small compared with I_{cell}. $I_{leak(d)}$ is generated from BL_{m+1} to the drain (D) through the drain-side neighboring cell ($m + 1$), due to voltage drop at D by I_{sen} across the resistance (R_{bl}) of BL_m. On the other hand, the positive ΔV_{th} due to the C_{fgwl} coupling effect increases gradually with the increase of $V_{th,drn}$ that is lower than that by the $I_{leak(d)}$ effect. As a result, ΔV_{th} in normal-mode can be controlled to be less than 200 mV in the negative direction when $V_{th,drn}$ changes from $V_{pv1} - 2$ V to $V_{pv1} + 3$ V. The negative ΔV_{th} can be compensated by CHE programming and subsequent verification. The additional V_{th} shift of the cell previously programmed by CHE V_{th} compensation is estimated to be $\alpha_{fgwl,side} \times 200$ mV = 2 mV, where $\alpha_{fgwl,side}$ is FG-to-FG coupling ratio (0.01) between a memory cell and the one side of neighboring memory cell in the WL direction.

3) *CHE Compensation Technique*

A CHE compensation technique is used to tighten the V_{th} distribution by reducing the cell-to-cell interference between neighboring cells for scalable virtual-ground MLC NOR flash memory, as shown in Fig. 2.11(b). The compensation programming is carried out

sequentially from left to right, with reference to the scheme given in Fig. 2.4, to eliminate the interference from the source-side neighboring cell. In addition, a programmed cell (*m* – 1) is reprogrammed using CHE after programming of the drain-side neighboring cell (*m*) to compensate the interference from the drain-side neighboring cell.

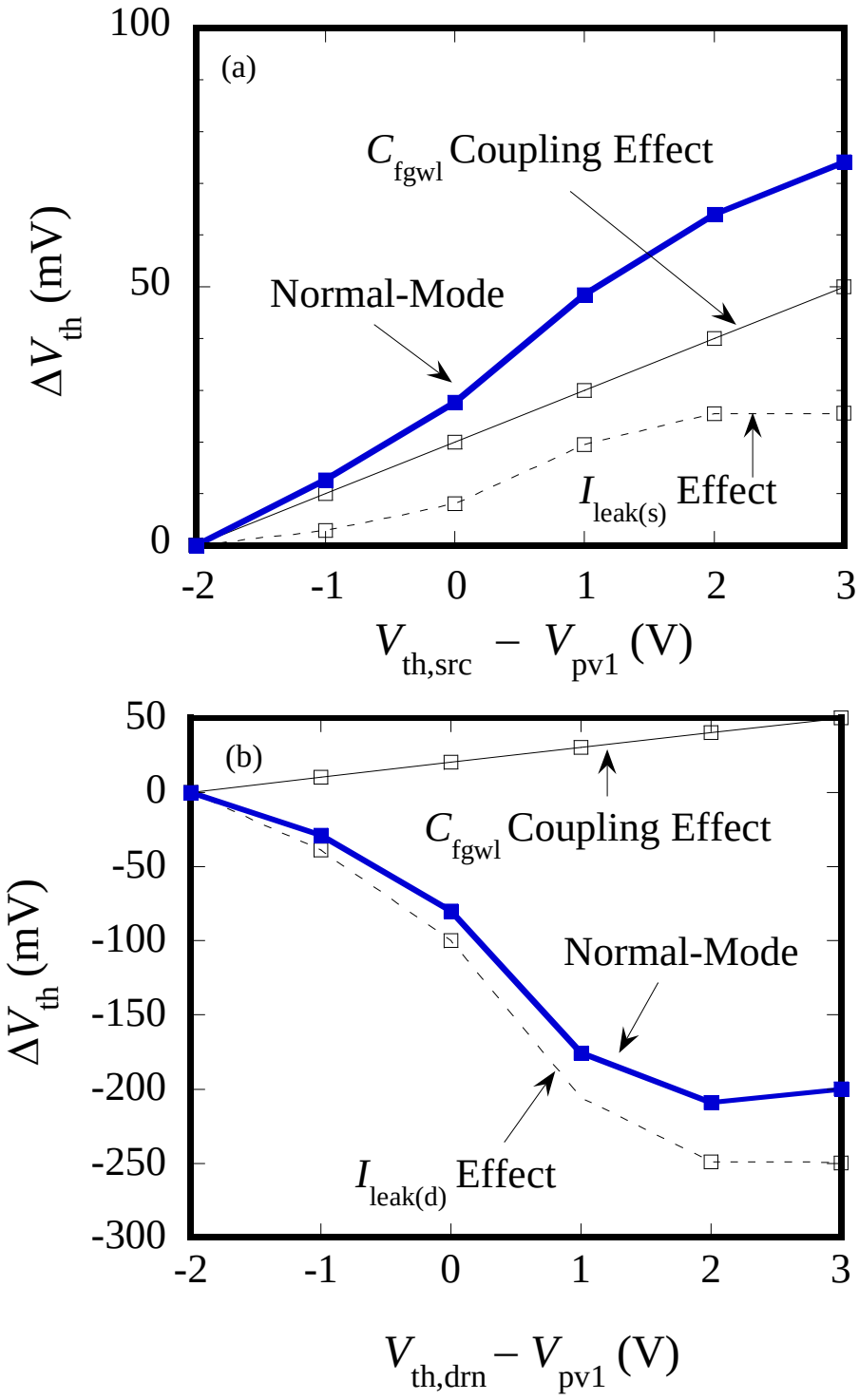


Fig. 2.15. ΔV_{th} for the previously programmed cell (*m*) as a function of (a) $V_{th,src} - V_{pv1}$ and (b) $V_{th,drn} - V_{pv1}$ [1.35].

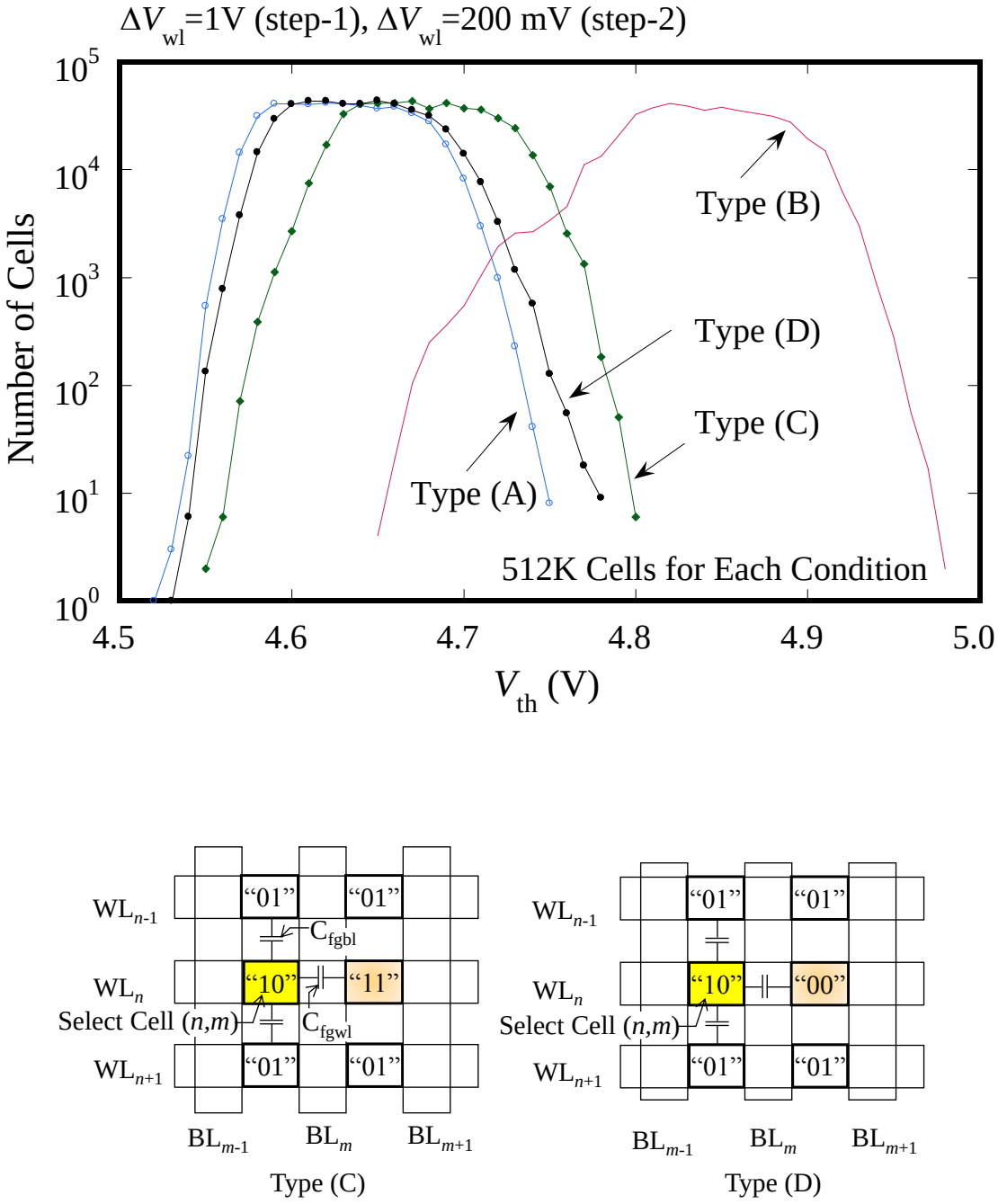


Fig. 2.16. Measured V_{th} distributions for the selected cells ("10" state) after two-step programming with V_{th} compensation for type (C) and type (D) [1.35].

D. Experimental Result

The proposed two-step programming using the CHE V_{th} compensation technique was performed with an incremental gate voltage of 200 mV for the two worst case patterns, as shown in Fig. 2.16. The worst case pattern that causes the largest cell-to-cell interference differs according to whether cell (n, m + 1) is programmed or not. Type (C) is the worst case pattern when cell (n, m + 1) is not programmed and type (D) is the worst case pattern when cell (n, m + 1) is programmed. Tight programmed V_{th} distribution widths of 280 mV and 230 mV are obtained for type (C) and type (D) patterns, respectively. The programmed V_{th} distribution width for type (D) is close to that of a single state programming without cell-to-cell interference, due to CHE compensation. These results are acceptable for an MLC approach.

E. Program Throughput

PT is compared for the proposed virtual-ground and standard NOR flash memories using the same 0.13 μm process technology, and is described by:

$$\mathrm{PT} \propto \frac{N_{\mathrm{bit}}}{T_{\mathrm{c}} N_{\mathrm{p}}}, \tag{2.1}$$

where N_{bit} is the number of memory cells programmed in parallel. $T_c = T_p + T_i$ ($T_p \gg T_i$), where T_p is the time width of a program pulse applied to the BL and T_i is the time interval between two successive pulses. N_p is the number of program pulses. N_{bit} and T_p are mainly determined by the program current and the program time required for a single memory cell, respectively. The same values of N_{bit} and T_c are used for the two types of NOR flash memories because they show little difference in program characteristic for a single cell. N_p is given by:

$$N_{\mathrm{p}} = \frac{\Delta V_{\mathrm{th,pg}}}{\Delta V_{\mathrm{wl}}}, \tag{2.2}$$

$$\Delta V_{\mathrm{wl}} = \Delta V_{\mathrm{th,target}} - \Delta V_{\mathrm{th,fg}}, \tag{2.3}$$

$$\Delta V_{\mathrm{th,fg}} = \alpha_{\mathrm{fg}} \Delta V_{\mathrm{th,pg}}, \tag{2.4}$$

where $\Delta V_{th,pg}$ is the change in V_{th} (4 V) required for programming, $\Delta V_{th,target}$ is the target V_{th} distribution width (300 mV) and $\Delta V_{th,fg}$ is the V_{th} shift due to FG-to-FG coupling interference through the two sides of C_{fgbl} and two sides of C_{fgwl} from the neighboring cells, and α_{fg} is FG-to-FG coupling ratio between a memory cell and the four sides of neighboring memory cells (α_{fg} = 0.052 for a virtual-ground and 0.010 for a standard NOR). From Eqs. (2.2) to (2.4), N_p is 44 cycles with $\Delta V_{th,fg}$ = 208 mV for virtual-ground, and 16 cycles with $\Delta V_{th,fg}$ = 40 mV for standard NOR.

On the other hand, N_p can be obtained from the following equations for the proposed two-step programming with CHE compensation that requires an additional one program

pulse for plural memory cells programmed in parallel.

$$N_{\mathrm{p}} = N_{\mathrm{p1}} + N_{\mathrm{p2}}, \tag{2.5}$$

$$N_{\mathrm{p1}} = \frac{\Delta V_{\mathrm{th,pg1}}}{\Delta V_{\mathrm{wl,step1}}}, \tag{2.6}$$

$$N_{\mathrm{p2}} = \frac{\Delta V_{\mathrm{th,pg2}}}{\Delta V_{\mathrm{wl,step2}}} + 1, \tag{2.7}$$

where N_{p1} and N_{p2} are the number of required program pulse in the first step and in the second step, respectively. $\Delta V_{\mathrm{th,pg1}}$ and $\Delta V_{\mathrm{th,pg2}}$ are the change in V_{th} required for programming in the first step and the second step, respectively ($\Delta V_{\mathrm{th,pg1}}$ = 4V and $\Delta V_{\mathrm{th,pg2}}$ = 3 V). $\Delta V_{\mathrm{wl,step1}}$ is the incremental gate voltage in the first step ($\Delta V_{\mathrm{wl,step1}} \geq 1$ V). $\Delta V_{\mathrm{wl,step2}}$ is the incremental gate voltage in the second step, and given by

$$\Delta V_{\mathrm{wl,step2}} = \Delta V_{\mathrm{th,target}} - \Delta V_{\mathrm{th,fgbl}}, \tag{2.8}$$

$$\Delta V_{\mathrm{th,fgbl}} = 25\% \times \alpha_{\mathrm{fgbl}}\, \Delta V_{\mathrm{th,pg2,}} \tag{2.9}$$

where $\Delta V_{\mathrm{th,fgbl}}$ is the V_{th} shift due to FG-to-FG coupling interference through the two sides of C_{fgbl} from the neighboring cells, and α_{fgbl} is FG-to-FG coupling ratio between a memory cell and the two sides of neighboring cells in the BL direction (α_{fgbl} = 0.032 for a virtual-ground). N_{p} is reduced from 44 to 16 cycles for virtual-ground memory, from Eqs. (2.5) to (2.9). Thus, the virtual-ground flash memory can achieve PT close to the standard NOR flash memory.

VII. ENHANCEMENT OF PROGRAM EFFICIENCY

The above-mentioned CHE injection requires high power consumption because of its low injection efficiency. A split-gate virtual-ground cell [1.36] has been proposed to enhance the program efficiency.

A. Cell Structure

Figure 2.17 shows the schematic cross-sectional view of the split-gate virtual-ground cell, together with an SEM cross sectional view. This cell consists of a memory transistor and an auxiliary transistor. The sidewall FG is formed on the drain side of an auxiliary gate (AG) and is self-aligned to the word line (WL). The inter-poly-dielectric (IPD) is formed on the vertical surface of a sidewall FG that can enable a large gate coupling ratio without sacrificing cell area. A cell area as small as $6F^2$ can be achieved.

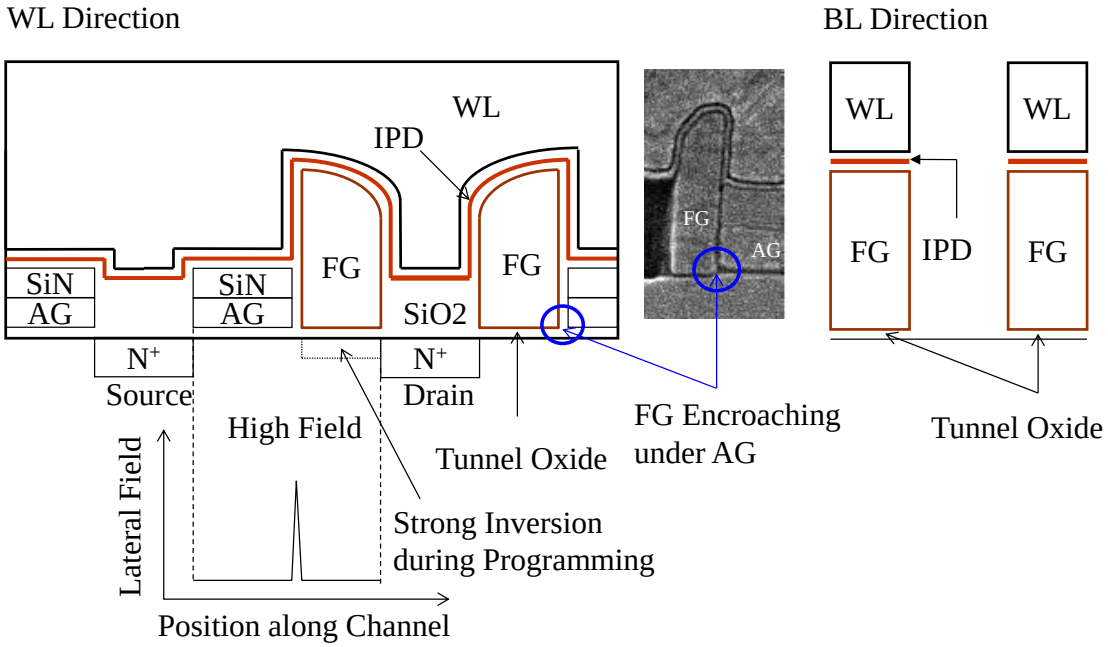


Fig. 2.17. Cross-sectional view of split-gate virtual-ground cell [1.36].

B. Principle of High Efficient Programming

The split-gate virtual-ground cell is programmed by hot electron injection into the FG through the auxiliary transistor from the source (source-side injection) due to the high lateral field (E_x) at the gap region between AG and FG as shown in Fig. 2.17. The E_x strongly depends on the voltage (V_{ag}) on the auxiliary gate and increases with a decrease of V_{ag}. The channel under the FG merely works as the drain extension during programming, since the high voltage is applied to the word line to induce the strong inversion layer. Gate oxide formation is the heart of the split-gate flash memory technology. A sharp corner at the bottom edge of the FG encroaching under the AG in the vicinity of the gap region causes an increase in the leakage between the FG and the AG.

C. Program Performance

The auxiliary gate voltage ($V_{ag,eff}$) effect on programming is shown in Fig. 2.18, where $V_{ag,eff}$ is defined as $V_{ag} - V_{th}$. $V_{th,ag}$ is the threshold voltage of the auxiliary transistor. The cell V_{th} after programming strongly depends on V_{ag}. The enhanced hot electron injection of the cell is observed at a V_{ag} slightly above the threshold voltage (1.2V) of the auxiliary transistor.

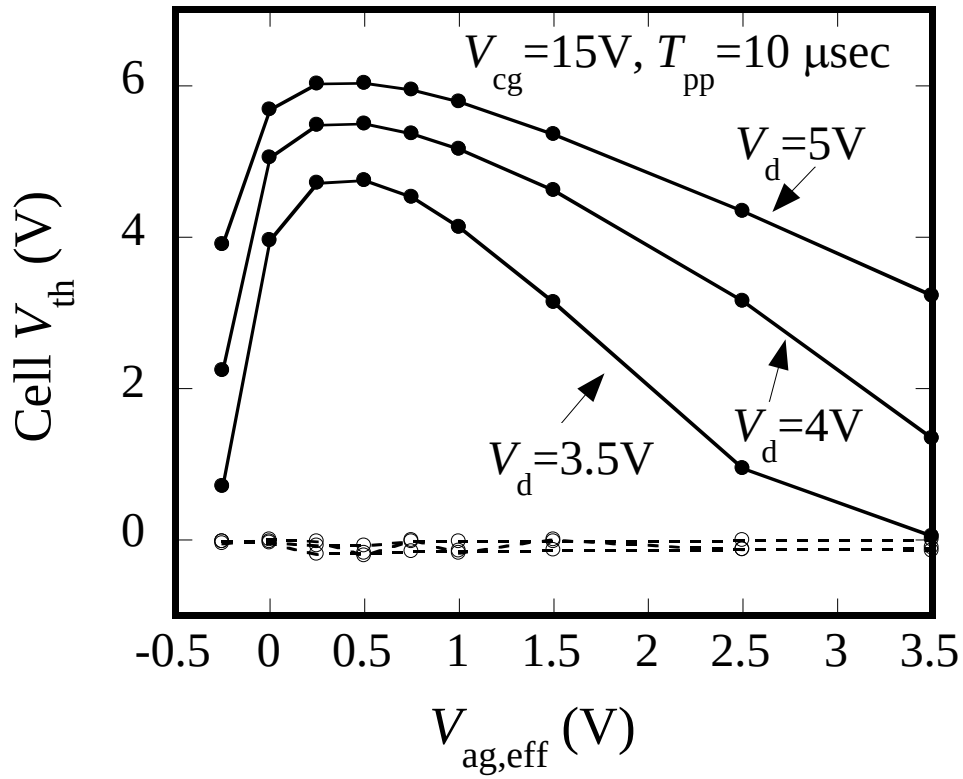


Fig. 2.18. Program characteristics [1.36].

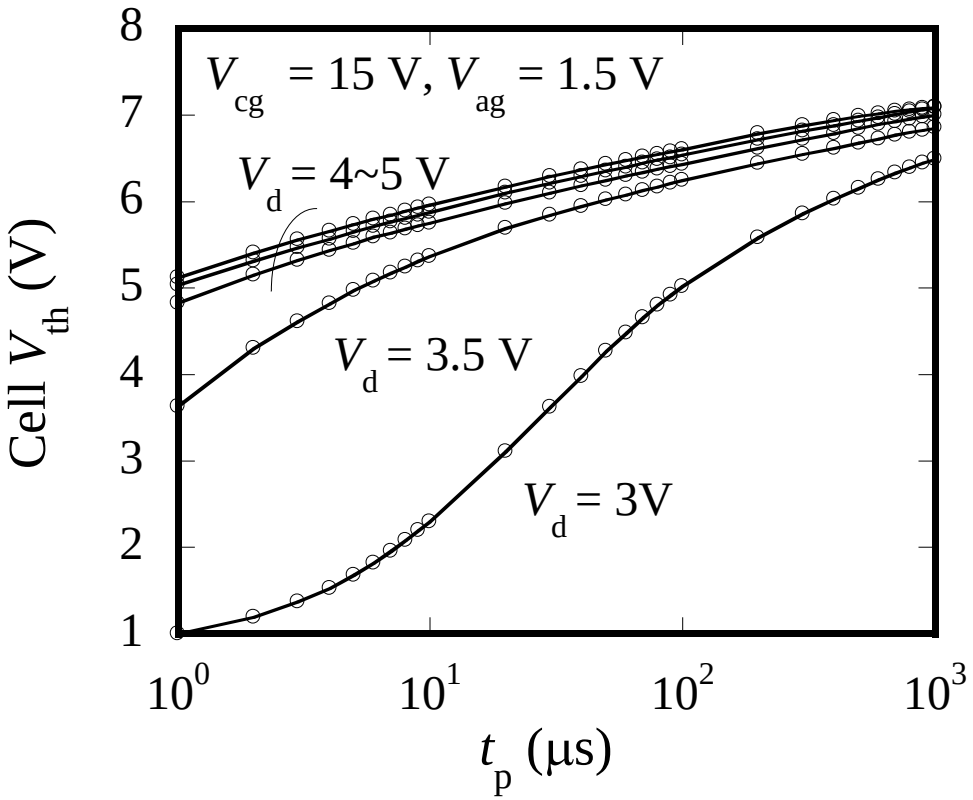


Fig. 2.19. Program characteristics [1.36].

The fast programming characteristics are obtained as shown in Fig. 2.19. The cell is programmed to 5V in less than 1 μs with a drain voltage higher than 4 V. The program speed of the auxiliary gate cell is faster by more than one order of magnitude than that achievable in a same generation standard NOR Flash EEPROM cell. The programming current can be also reduced because of a low V_{ag} during programming. As a result, high efficiency programming can be realized.

VIII. CONCLUSIONS

A scalable MLC virtual-ground NOR flash memory cell technology is developed, which has potential for continued scaling of fast-access NOR flash memory because of its contactless structure. However, the cell-to-cell interference caused by virtual-ground architecture makes it difficult to achieve MLC, the mechanisms of which can be explained by both the neighboring cell leakage and capacitive coupling models. An adequate GCR and compensation for cell-to-cell interference are required to enable further scaling of virtual-ground MLC FG flash memory. A high GCR of 0.6 is obtained using a novel bowl-shaped FG structure cell technology without sacrificing cell size. Increasing the GCR is important for reducing FG-to-FG coupling interference and achieving low voltage operation. A novel array segmented virtual-ground architecture with BL isolation between neighboring segments and two-step programming with the CHE injection threshold voltage compensation technique are proposed to reduce the number of neighboring cells that are programmed after programming a given cell, as well as the amount of V_{th} shift of the neighboring cells. Adoption of this programming approach realizes a reduction in the V_{th} shift caused by the FG-to-FG coupling interference to less than 25% in the BL direction, compared with conventional programming approaches, and the V_{th} shift is almost completely eliminated in the word-line direction without sacrificing program throughput. The proposed virtual-ground MLC FG cell, which is as small as $3F^2$/bit (F is the minimum feature size) is successfully implemented into a test chip. Good reliability, up to 10^5 cycles, is also obtained. The proposed scalable MLC virtual-ground technology allows the CPU to execute a program directly without going through volatile memory, that leading to reduction in DRAM refresh power consumption. Thus, the technology is suitable for internet mobile applications that require execute-in-place performance and high density.

The split-gate virtual-ground cell is also studied, which can enhance program speed with lower power consumption without sacrificing cell area, as compared with the case

of the above-mentioned single-gate virtual-ground cell. A technology that overcomes the issue in gap oxide formation, the single polysilicon split-gate cell, and which has been adopted in embedded applications will be described in Chapter 3.

REFERENCES

[2.1] M. Bauer, R. Alexis, G. Atwood, B. Baltar, A. Fazio, K. Frary, M. Hensel, M. Ishac, J. Javanifard, M. Landgraf, D. Leak, K. Loe, D. Mills, P. Ruby, R. Rozman, S. Sweha, S. Talreja, and K. Wojciechowski,"A Multilevel-Cell 32-Mb Flash Memory," *IEEE International Solid-State Circuits Conference Digest of Technical Papers, 1995*, pp. 132 – 133.

[2.2] G. Campardo, R. Micheloni, S. Commodaro, E. Yero, M. Zammattio, S. Mognoni, A. Sacco, M. Picca, A. Manstretta, M. Scotti, I. Motta, C. Golla, A. Pierin, R. Bez, A. Grossi, A. Modelli, A. Visconti, O. Khouri, and G. Torelli, "40 mm2 3-V-Only 50-MHz 64-Mb 2-b/cell CHE NOR Flash Memory," *IEEE Journal Solid-State Circuits*, vol. 35, no. 11, pp. 1655 – 1667, Nov. 2000.

[2.3] D. Elmhurst, R. Bains, T. Bressie, C. Bueb, E. Carrieri, B. Chauha, N. Chrisman, M. Dayley, R. De Luna, K. Fan, M. Goldma, P. Govindu, A. Huq, M. Khandaker, J. Kreifels, S. Krishnamacha, P. Lavapie, K. Loe, T. Ly, F. Marvin, R. Melcher, S. Mona. Nguyen, B. Pathak, A. Proescholdt, T. Rahman, B. Srinivasan, R. Sundaram, P. Walimbe, D. Ward, Dave, R. Zeng, and H. Zhang, "A 1.8V 128Mb 125MHz Multi-level Cell Flash Memory with Flexible Read While Write," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2003, pp. 54 – 55.

[2.4] M. Taub, R. Bains, G. Barkley, H. Castro, G. Christensen, S. Eilert, R. Fackenthal, H. Giduturi, M. Goldman, C. Haid, R. Haque, K. Parat, S. Peterson, A. Proescholdt, K. Ramamurthi, P. Ruby, B. Sivakumar, A. Smidt, B. Srinivasan, M. Szwarc, K. Tedrow and D. Young, "A 90nm 512Mb 166MHz Multi-level Cell Flash Memory with 1.5MByte/s Programming," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2005, pp. 54 – 55.

[2.5] C. Villa, D. Vimercati, S. Schippers, S. Polizzi, A. Scavuzzo, M. Perroni, M. Gaibotti, and M. L. Sali, "A 65 nm 1Gb 2b/Cell NOR Flash with 2.25MB/s Program Throughput and 400MB/s DDR Interface," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2007, pp. 476 – 477.

[2.6] S. Keeney, "A 130nm Generation High Density Etox^{TM} Flash Memory Technology", *International Electron Devices Meeting Technical Digest*, 2001, pp. 251 – 254.

[2.7] G. Atwood, "Future Directions and Challenges for ETox Flash Memory Scaling," *IEEE Transactions on Device and Materials Reliability*, vol. 4, no. 3, pp. 301 – 305, Sept. 2004.

[2.8] T. Ogura, M. Mihara, Y. Kawajiri, K. Kobayashi, T. Sakaniwa, K. Nishikawa, S. Shimizu, S. Shukuri, N. Ajika and M. Nakashima," A Fast Rewritable 90nm 512Mb NOR "B4-Flash" Memory with $8F^2$ Cell Size," *Symposium on VLSI Circuits Digest of Technical Papers*, 2011, pp. 198 – 199.

[2.9] R. Koval, V. Bhachawat, C. Chang, M. Hajra, D. Kencke, Y. Kim, C. Kuo, T. Parent, M. Wei, B.J. Woo, and A. Fazio, "Flash ETOX Virtual Ground Architecture: A Future Scaling Direction," *Symposium on VLSI Technology Digest of Technical Papers*, 2005, pp. 204 – 205.

[2.10] J. D. Lee, S. H. Hur, and J. D. Choi, "Effects of Floating-Gate Interference on NAND Flash Cell Operation," *IEEE Electron Device Letters*, vol. 33, no. 5, pp. 264 – 266, May 2002.

[2.11] K. T. Park, M. Kang, D. Kim, S. Hwang, B.Y. Choi, Y.T. Lee, C. Kim, and K. Kim, "A Zeroing Cell-to-Cell Interference Page Architecture with Temporary LSB Storing and Parallel MSB Program Scheme for MLC NAND Flash Memories," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 4, pp. 919 – 928, May 2008.

[2.12] N. Ito, Y. Yamauchi, N. Ueda, K. Yamamoto, Y. Sugita, T. Mineyama, A. Ishihama, and K. Moritani, "A Novel Program and Read Architecture for Contact-Less Virtual Ground NOR Flash Memory for High Density Application," *Symposium on VLSI Circuits Digest of Technical Papers*, 2006, pp. 204 – 205.

[2.13] Y. Yamauchi, M. Yoshimi, S. Sato, H. Tabuchi, N. Takenaka, and K. Sakiyam, "A New Cell Structure for Sub-quarter Micron High Density Flash Memory", *International Electron Devices Meeting Technical Digest*, 1995, pp. 267 – 270.

[2.14] R. E. Shiner, J. M. Caywood, and B. L. Euzent, "Data Retention in EEPROMs," *IEEE International Reliability Physics Symposium*, 1980, pp. 238 – 243.

[2.15] B. Q. Le, M. Achter, C.G. Chng, X. Guo, L. Cleveland, P.L. Chen, M.V. Buskirk, and R.W. Dutton, "Virtual-Ground Sensing Techniques for a 49-ns/200MHz Acess Time 1.8-V 256-Mb 2-Bit-per-Cell Flash Memory," *IEEE J. Solid-state Circuits*, vol. 39, no. 11, pp. 2014 – 2022, Nov. 2004.

[2.16] A. Modelli, A. Manstretta, and G. Torelli, "Basic Feasibility Constraints for Multilevel CHE-Programmed Flash Memories," *IEEE Transactions Electron Devices*, vol. 48, no. 9, pp. 2032 – 2042, Sept. 2001.

Chapter 3

Logic Compatible Split-Gate Flash Memory

I. Introduction 43
II. Cell Structure and Process Technology 44
III. Operation Concept 48
A. Program 49
B. Erase 50
C. Read 50
IV. Study of Source-Side Hot-Electron Injection 50
A. Role of Access Gate Transistor 51
B. Injection Efficiency 52
C. Injection Mechanism 53
V. Evaluation Results 56
A. Performance 56
B. Endurance and Data Retention 61
C. Program Disturb Immunity 64
VI. Conclusions 69
References 69

I. Introduction

The NAND flash memory market continues to show large and rapid growth with the increase in density of the data storage needed for smartphones, computers, cameras, music players, and other consumer products. Equally important is the embedded nonvolatile memory market, which is growing steadily with the increase in the micro-controller-unit market for industrial, automotive, and consumer applications [3.1] – [3.3].

A large variety of different cell concepts for embedded nonvolatile memories can be found in the market depending on specific application requirements. However, most of

them use the NOR array configuration because of the requirements for fast programming and fast random-access read. The channel hot-electron (CHE) injection programming method is used for standard NOR (STD-NOR) flash memory, which can offer fast programming at the expense of high power consumption because of its low injection efficiency. On the other hand, the Fowler-Nordheim (FN) tunneling programming method can offer low power consumption with high injection efficiency, while its low tunneling current results in slow programming. To overcome these problems, the enhanced source-side channel hot-electron (SSHE) injection programming method was proposed [3.4], [3.5].

Many studies of split-gate flash memory using SSHE have been reported [3.5] – [3.14]. The split-gate flash memory is good candidate for embedded nonvolatile memories because of its advantages of fast programming and low power consumption. However, conventional split-gate memory cells require significant modification to the standard logic process because of their process complexity.

Emerging memories using new materials, such as the ferroelectric random access memory (RAM), phase change memory, resistive RAM, and magnetic RAM [1.20] – [1.29], have been studied for next-generation embedded applications, because they can offer the performance of RAM with nonvolatility of read only memory (ROM). However, these new nonvolatile RAMs have not become the mainstream for embedded nonvolatile memories that require expensive special processing steps.

This chapter describes a source-side injection single-polycrystalline-silicon (single-polysilicon) split-gate NOR (S4-NOR) flash memory technology, which can satisfy high performance and low power without using any special process. In Section II, the memory cell structure and process technology are described. In Section III, the memory cell operation concept is explained. In Section IV, the SSHE injection mechanism is studied. In Section V, the memory cell performance and reliability are evaluated. Finally, conclusions are given in Section VI.

II. CELL STRUCTURE AND PROCESS TECHNOLOGY

A schematic layout and cross-sectional views of the S4-NOR cell are shown in Fig. 3.1. For the S4-NOR cell, the access gate (AG) and the floating gate (FG) are patterned on a channel between its source (S) and drain (D) with a small gap length (L_{gap}) using conventional photolithography technology beyond the 100 nm generation. The FG is capacitively coupled to the n-well memory gate (MG) through the gate oxide film with a

capacitance of C_{mg}. C_{mg} is designed to achieve a gate coupling ratio (R_{mg}) of 0.7, where R_{mg} is defined as the ratio of C_{mg} to the total capacitance of the FG. This technology can improve the compatibility with a single-polysilicon CMOS logic process.

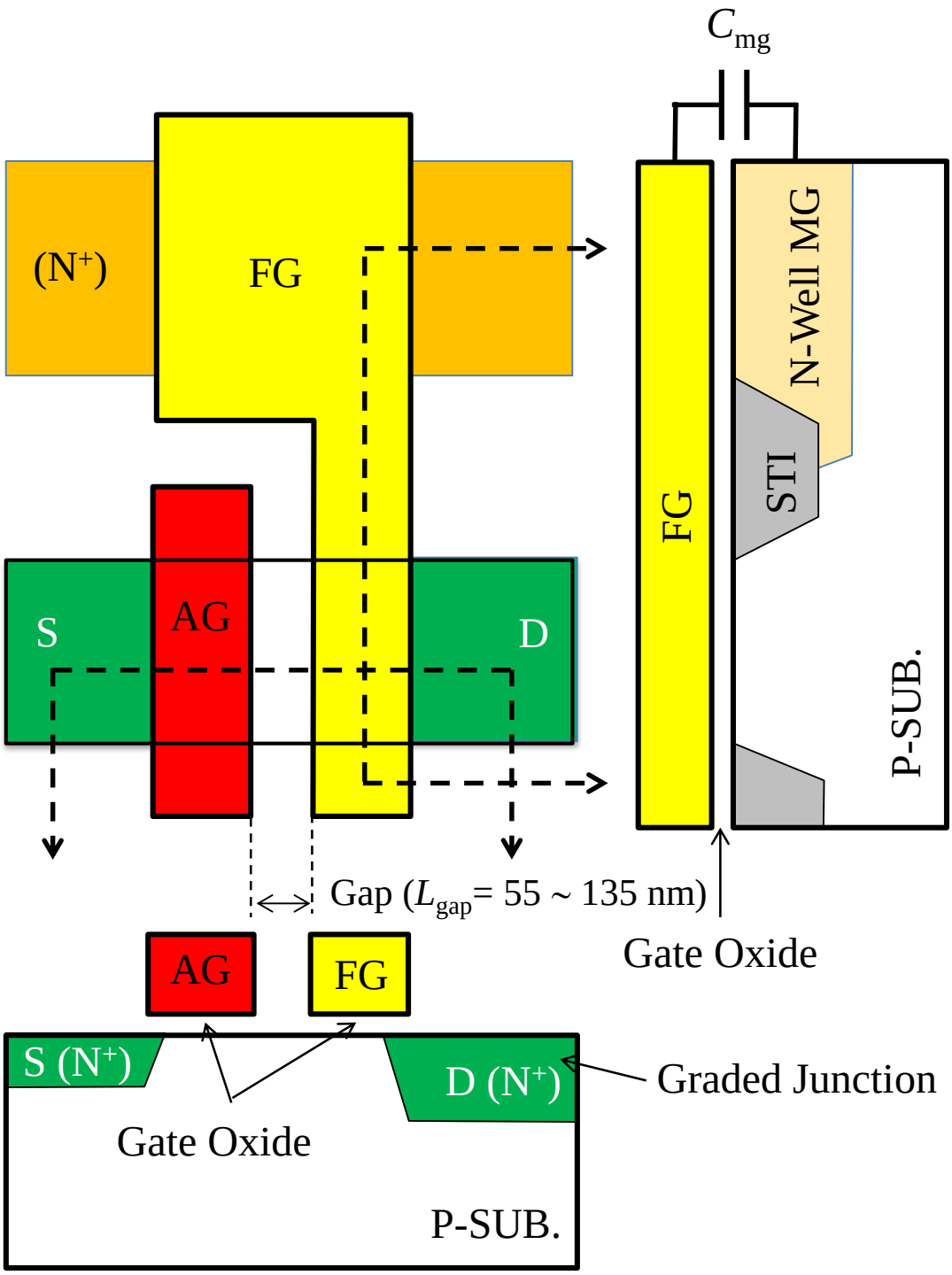


Fig. 3.1. Schematic layout and cross-sectional views of the S4-NOR cell [1.37].

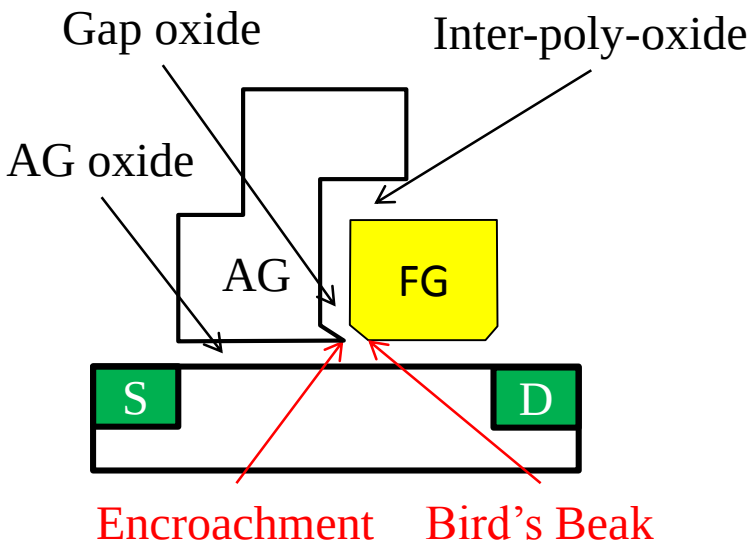


Fig. 3.2. Conventional double-polysilicon split-gate structure [1.33].

The graded drain junction is formed by the implantation of phosphorus ions (dosage: 1×10^{15} cm^{-2}) and arsenic ions (dosage: 1×10^{15} cm^{-2}), which can be optimized only for erase operation. This is because the program efficiency is mainly determined by L_{gap} and is not dependent on the junction profiles as will be discussed in Section IV.C. The AG and FG transistors have a gate length and a gate width of 0.18 μm, a gate oxide thickness of 10 nm, and L_{gap} = 55 nm, unless otherwise mentioned.

Gap oxide formation is the heart of the split-gate flash memory cell technology. In the conventional split-gate structure shown in Fig. 3.2, the inter-poly-oxide, gap oxide, and AG oxide are formed simultaneously, after defining the FG. During these process steps, a sharp corner at the bottom edge of the AG encroaching under the FG in the vicinity of the gap region is formed [3.15]. These process issues cause the leakage between the FG and the AG in the split-gate flash memory, and studies of gap oxide technologies have been performed to prevent them [3.15] – [3.17].

Figure 3.3 shows the key process steps for the S4-NOR cell. First, after gate oxidation, the polysilicon film is deposited and then patterned to form the AG and the FG using the photoresist (PR) mask (step 1). Next, after patterning the polysilicon film, arsenic ions are implanted into the source and drain regions and phosphorus ions are subsequently implant using the PR (step 2). After that, the CVD oxide is deposited to fill the gap region between the FG and the AG (step 3). In the proposed S4-NOR, it is possible to reduce L_{gap} without suffering from the above-mentioned process issues in

conventional split-gate structures, because both the AG and the FG are patterned simultaneously.

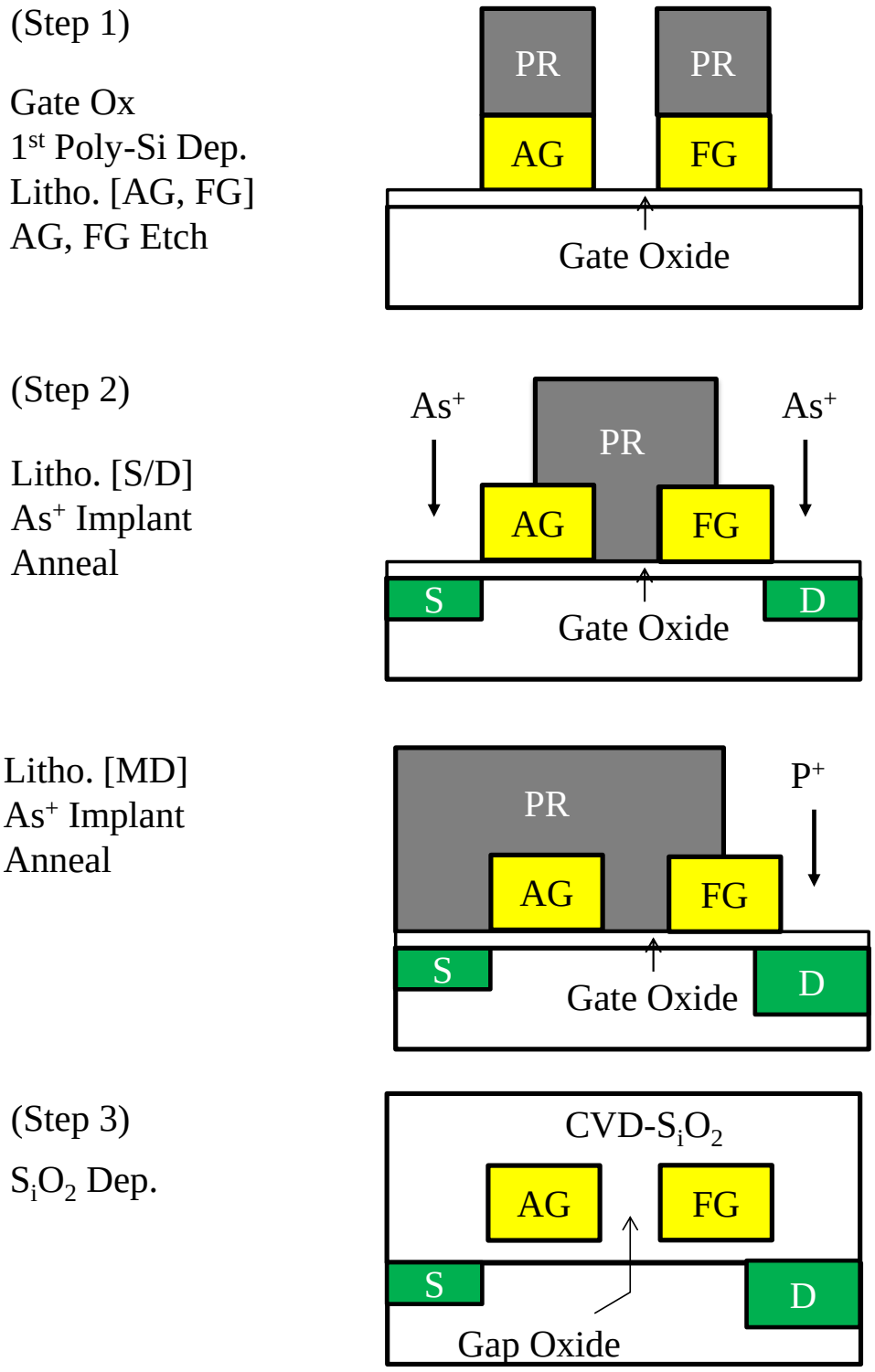


Fig. 3.3. Key process steps for the gap oxide formation in the S4-NOR cell [1.37].

III. OPERATION CONCEPT

Figure 3.4 shows the operation concepts of program, erase, and read for the S4-NOR cell. Typical bias conditions (in volts) are shown in Table 3.1.

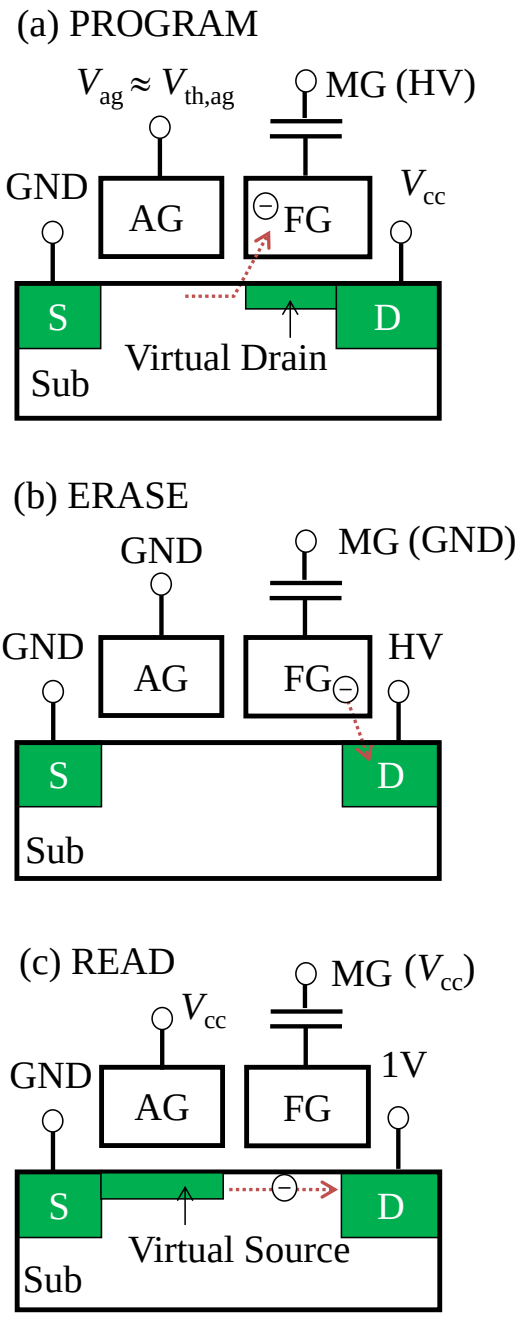


Fig. 3.4. Operation concepts of (a) program, (b) erase, and (c) read for the S4-NOR cell [1.37].

Table 3.1

Typical Bias Conditions for Program, Erase, and Read Operations [1.37].

Mode	Applied Voltages (V)			
	AG	D	MG	S
PROGRAM	1.8	V_{cc}	11.5	0
ERASE	0	10.5	0	0
READ	V_{cc}	1.0	V_{cc}	0

A. Program

The main disadvantage of the CHE programming method is its low injection efficiency and consequently its high power consumption, as mentioned in Section I. The low injection efficiency of CHE is caused by an inherent incompatibility of having a high lateral field and a high vertical field, favorable for electron injection. Indeed, the lateral field increases with a decrease in the FG voltage (V_{fg}), while the vertical field increases with the increase in V_{fg}. This issue can be solved by implementing an additional AG transistor, as will be studied in detail in Section IV.

Figure 3.4(a) shows the concept of the SSHE injection programming [3.18] – [3.20]. The high voltage (HV) is applied to the MG on the drain side of the channel to obtain a high vertical field across the gate oxide between the FG and the substrate (Sub) while a supply voltage (V_{cc} = 5 V) is applied to the drain with the grounded source. Here, the channel of the FG transistor acts as an extension of the drain (virtual drain). On the other hand, the AG on the source side of the channel is biased under the conditions for enhancing the lateral field induced in the gap region. The maximum hot-electron generation can be obtained at an AG voltage (V_{ag}) very close to the threshold voltage ($V_{th,ag}$) of the AG transistor independent of the vertical field. This approach can satisfy both the high hot-electron generation rate in the lateral direction and the high collection efficiency of injected electrons in the vertical direction.

B. Erase

Erase is achieved by applying an HV to the graded drain junction with the grounded MG, AG, and source, and then extracting electrons from the FG to the drain through the gate-to-drain overlap region as shown in Fig. 3.4(b). In the case of STD-NOR, a complicated erase algorithm is required to eliminate the leakage caused by single bit overerase, which is not required for S4-NOR because of its no-overerase structure. Therefore, this approach can simplify the circuit organization as compared with the erase schemes used for STD-NOR [3.21] – [3.23]. Furthermore, since program and erase occur over different areas of the memory cell, the stresses of program and erase are spread over different regions, resulting in improved reliability, as will be evaluated in Section V.B.

C. Read

Read-out is achieved by applying V_{cc} to both the AG and the MG and a low voltage of 1 V to the drain with the grounded source, as shown in Fig. 3.4(c). It is conducted from the drain side as well as program operation to simplify the read-path circuit organization. The read current is controlled completely within the FG channel, while the AG channel acts as an extension of the source (virtual source).

When reading a particular erased memory cell, the memory cell threshold voltage ($V_{th,mg}$) may be increased by the hot-electron injection into the FG even with a low lateral field, and thus compromise the 10-year retention period required for non-volatile memory. This unwanted programming is referred to as soft write. The soft write during reading is one of the problems in achieving fast programming with a low drain voltage (V_d) in STD-NOR because both program and read operations only differ in applied voltage. However, for S4-NOR, the injection efficiency can be markedly suppressed during the read-out operation. This is because program efficiency is enhanced by the AG channel, while it merely acts as the extension of the source during read-out operation.

IV. STUDY OF SOURCE-SIDE HOT ELECTRON INJECTION

The injection efficiency of the source-side injection single-polysilicon split-gate transistor is evaluated. Then, its injection efficiency is compared with that of an NMOS transistor with a single arsenic ion implanted in the source/drain region. The dependences of injection efficiency on V_d and L_{gap} are also evaluated to verify the validity of the source-side injection model described in a previous study [3.18].

A. Role of Access Gate Transistor

To study the role of the AG transistor, the effects of V_{ag} on the current injected into the gate on the drain side of channel (I_{inj}) and the drain current (I_d) were measured on a split-gate transistor. Figure 3.5 shows I_{inj}, I_d, and the injection efficiency (I_{inj}/I_d) as a function of V_{ag} at a gate voltage on the drain side of the channel (V_{dg}) of 8 V and V_d = 5 V. I_d increases exponentially to approximately $V_{th,ag}$ = 1.2 V and then results in saturation. Here, $V_{th,ag}$ is defined as V_{ag} when I_d = 1 μA at V_{fg} = 8 V. On the other hand, I_{inj}/I_d decreases exponentially with an increase in V_{ag}. As a result, the maximum I_{inj} = 600 pA is obtained at approximately $V_{ag} = V_{th,ag}$ with a low I_d = 10 μA, while I_{inj} is reduced by a factor of 600 at V_{ag} = 5 V as compared with its maximum at approximately $V_{ag} = V_{th,ag}$. This result indicates that SSHE injection can satisfy both high injection efficiency and high soft write immunity.

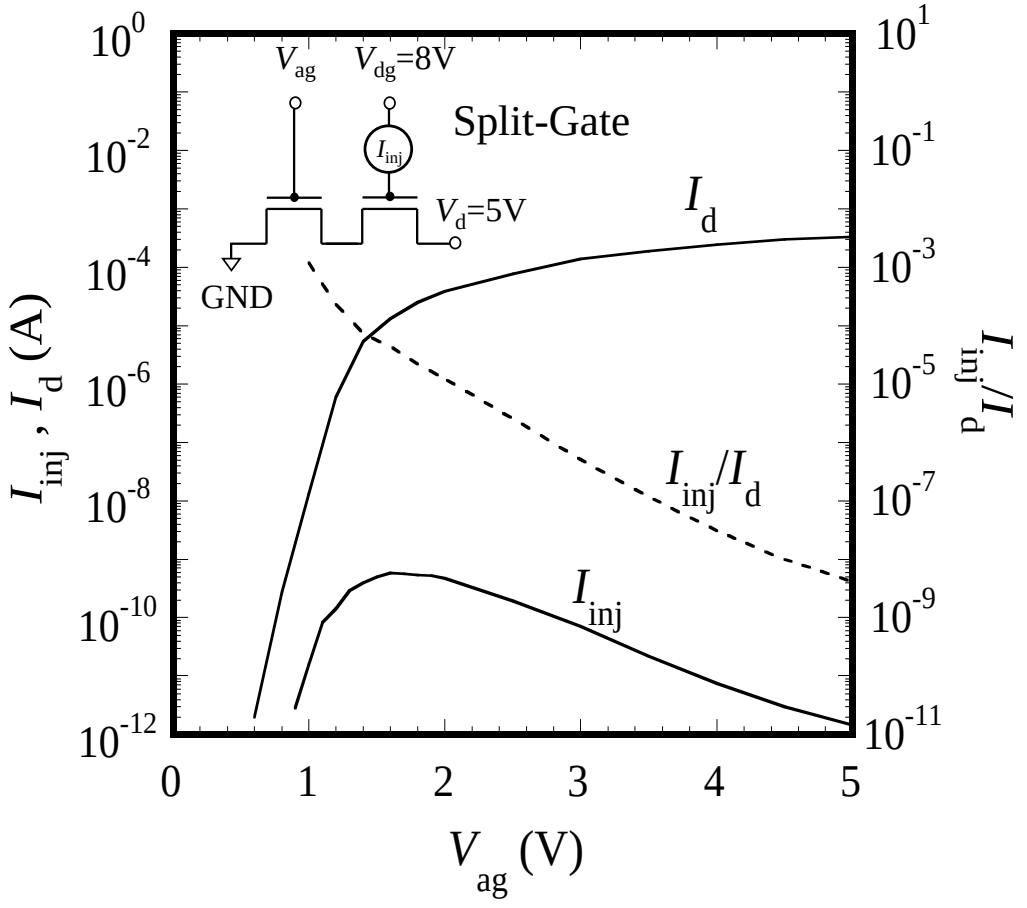


Fig. 3.5. I_{inj}, I_d, and I_{inj}/I_d at V_{dg}= 8 V and V_d = 5V as a function of V_{ag} [1.37].

B. Injection Efficiency

To compare the injection efficiency between SSHE and CHE, the effects of single-gate voltage (V_g) on I_{inj} and I_d were measured at V_d = 5 V on a single-gate transistor, as shown in Fig. 3.6, which is compared with the case of SSHE injection in a split-gate transistor (Fig. 3.4). The split-gate and single-gate transistors are fabricated with the same process technologies and have almost the same effective channel length and width under programming operation. The threshold voltage of the single gate transistor (V_{th}) is close to $V_{th,ag}$. In the CHE case, the maximum I_{inj} = 5 pA is obtained at approximately $V_g = V_d - 1$ V with a high I_d = 1 mA, indicating a low injection efficiency. This is because both V_g and V_d are kept high as a compromise between the incompatibility of the high lateral field and the high vertical field, as discussed above.

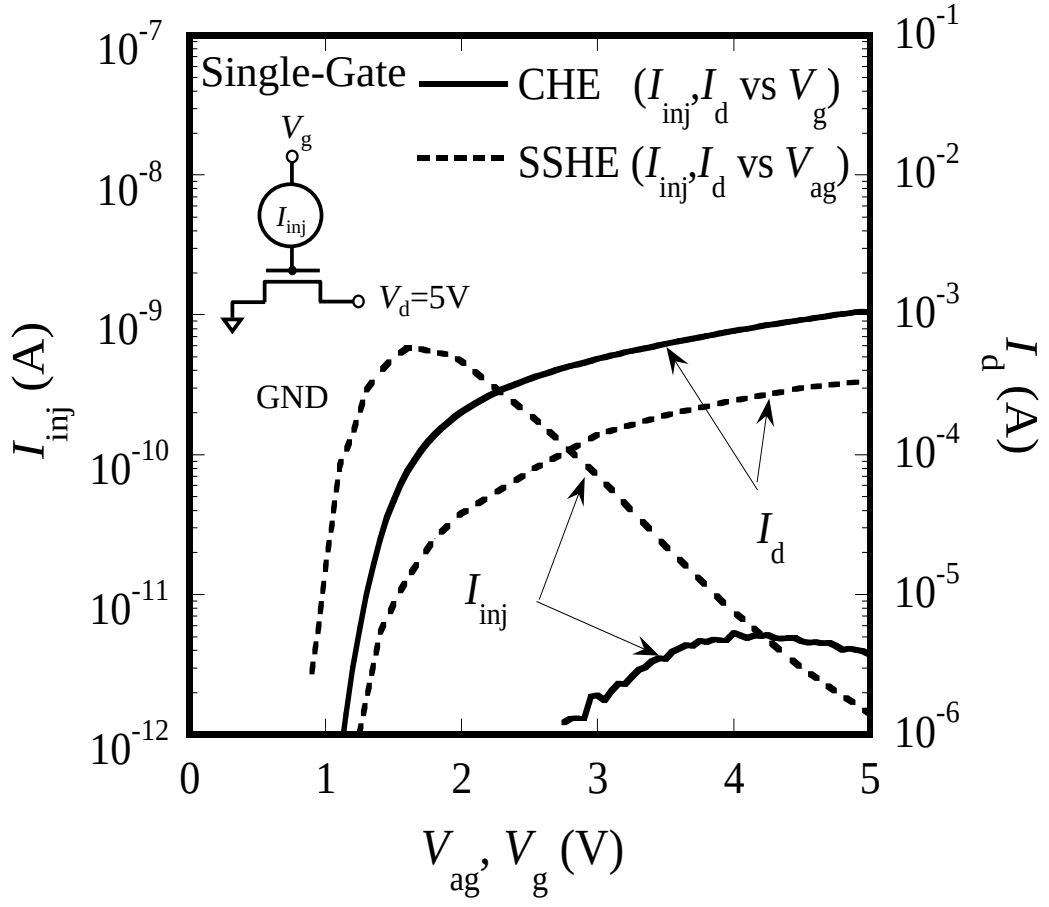


Fig. 3.6. Comparison of I_{inj} and I_d for the SSHE injection (V_{dg} = 8 V) and CHE injection, both measured at the same V_d = 5 V [1.37].

It is clear that SSHE injection provides a gate current higher by two orders of magnitude than that in the case of CHE injection. At the same time, I_d at the maximum I_{inj} in the SSHE case is also reduced by two orders of magnitude with respect to that in the CHE case. From these results, SSHE injection provides a higher injection efficiency by four orders of magnitude than CHE injection, similarly to previous studies [3.4].

C. Injection Mechanism

The injection efficiency of CHE in a single-gate NMOS transistor is given by the following expression based on the lucky-electron model [3.24] – [3.31]:

$$\frac{I_{inj}}{I_d} \propto \exp\left(-\frac{\phi_b}{q\lambda E_m}\right), \tag{3.1}$$

where ϕ_b is the effective barrier height for electrons at the Si/SiO_2 interface (~3.2 eV), q is the electron charge, λ is the hot-electron mean free path, and E_m is the maximum lateral electric field at the drain end of the channel. In the case of the split-gate NMOS transistor, most of the source-to-drain voltage is applied to the gap region instead of in the velocity saturation region in a MOS transistor, as shown in previous studies [3.18], [3.32]. Therefore, E_m can be empirically approximated by

$$E_m = \frac{V_d - V_{dsat}}{L_{gap}}, \tag{3.2}$$

where V_{dsat} is V_d at which carrier velocity saturates and $V_d - V_{dsat}$ corresponds to the voltage drop in the pinch-off region. In the case of CHE injection, $V_{dsat} = V_g - V_{th} \approx 2.8$ V (Fig. 3.6), and it is strongly dependent on the effective channel length in previous studies [3.31]. [3.32]. However, in the case of SSHE injection, $V_d - V_{dsat}$ can be approximated using V_d because of $V_{dsat} = V_{ag} - V_{th,ag} << V_d$ ($V_{ag} \approx V_{th,ag}$), as studied in Section IV.A. As a result, E_m in Eq. (3.2) can be rewritten as

$$E_m = \frac{V_d}{L_{gap}}. \tag{3.3}$$

By substituting Eq. (3.3) into Eq. (3.1), I_{inj}/I_d is rewritten as

$$\frac{I_{inj}}{I_d} \propto \exp\left\{-\left(\frac{\phi_b}{q\lambda}\right)\left(\frac{L_{gap}}{V_d}\right)\right\}. \tag{3.4}$$

This model indicates that the injection efficiency of SSHE is insensitive to the process parameters except L_{gap}, while the injection efficiency of CHE is strongly dependent on junction depth, gate oxide thickness, and substrate doping concentration as well as on effective channel length [3.24] – [3.28].

To verify the validity of the above-mentioned model, the injection efficiency at the maximum I_{inj} was measured at V_{dg} = 9 V and V_d = 5 V for different L_{gap} values, as shown in Fig. 3.7(a), and also at V_{dg} = 8 V for different V_d values on the split-gate transistor with L_{gap} = 55 nm, as shown in Fig. 3.7(b).

The AG is biased under the condition for the maximum I_{inj}. I_{inj}/I_d changes exponentially with L_{gap} and $1/V_d$. The values of λ calculated from the slopes of I_{inj}/I_d vs L_{gap} and I_{inj}/I_d vs $1/V_d$ are 18 and 12 nm, respectively. These values are considered to be reasonable from the results of previous studies [3.29] – [3.31]. These results support the above-mentioned model for SSHE injection.

From these discussions, it can be concluded that the S4-NOR cell can enhance the hot-electron injection efficiency and is also suitable for an embedded flash memory from the viewpoint of process control.

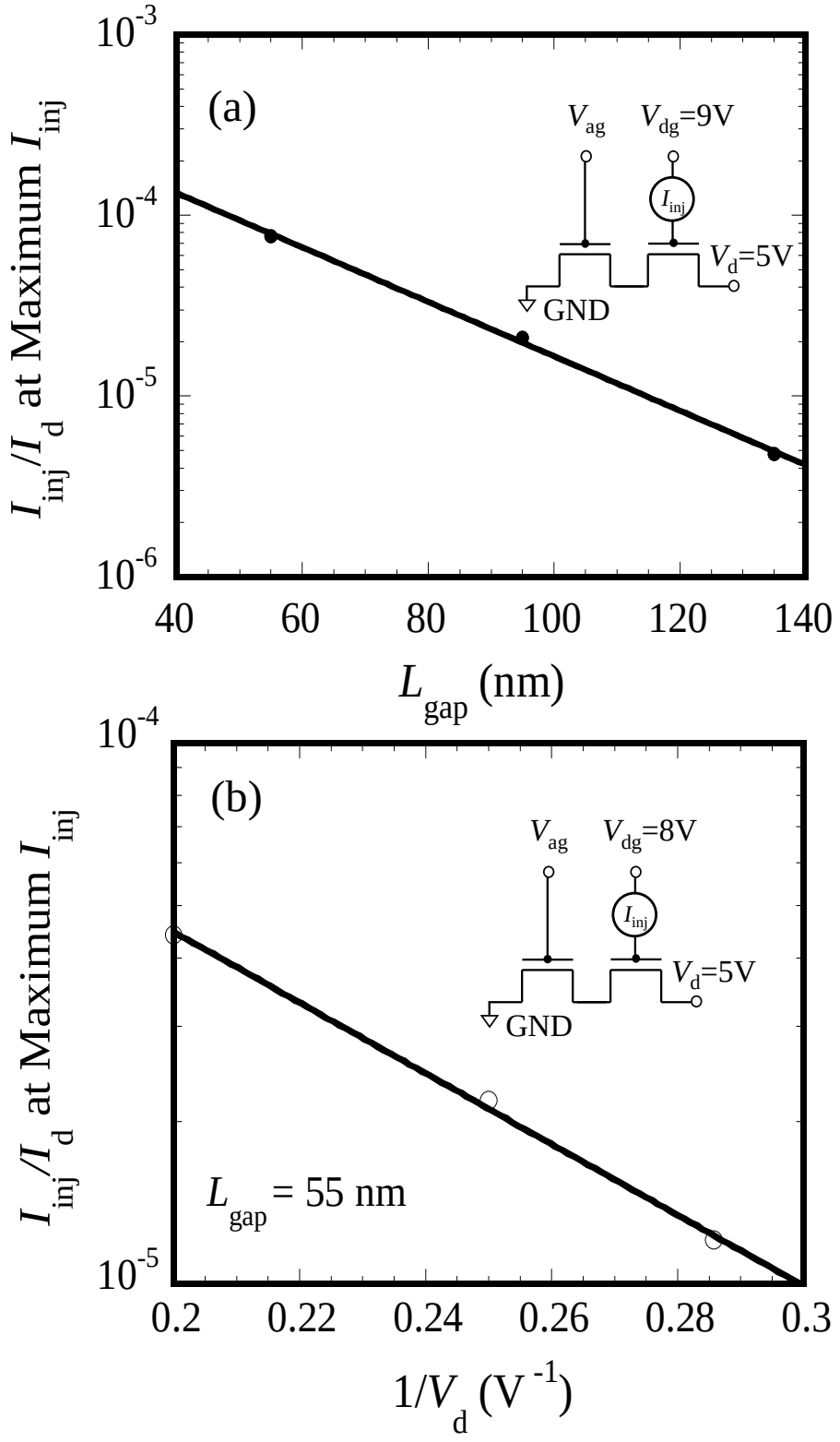


Fig. 3.7. (a) I_{inj}/I_d as a function of L_{gap} at V_{dg} = 9 V and V_d = 5 V and (b) I_{inj}/I_d as a function of $1/V_d$ at V_{dg} = 8 V, where the AG is biased at the maximum I_{inj} [1.37].

V. EVALUATION RESULTS

A. Performance

Program, erase, and read characteristic evaluations were conducted on typical single memory cells.

1) Program

Figure 3.8 shows the change in the memory cell threshold voltage ($\Delta V_{th,mg}$) as a function of program time at V_{mg} = 11.5 V and V_d = 5 V with V_{ag} as the parameter, which was measured on ultraviolet (UV) erased cells. The memory cell threshold voltage ($V_{th,mg}$) is defined as the memory gate voltage (V_{mg}) when I_d = 1 μA at V_{ag} = 5 V and V_d = 1 V. Enhanced programming is observed at V_{ag} between 1.6 and 2.0 V, which is close to $V_{th,ag}$ as studied in Section IV.A.

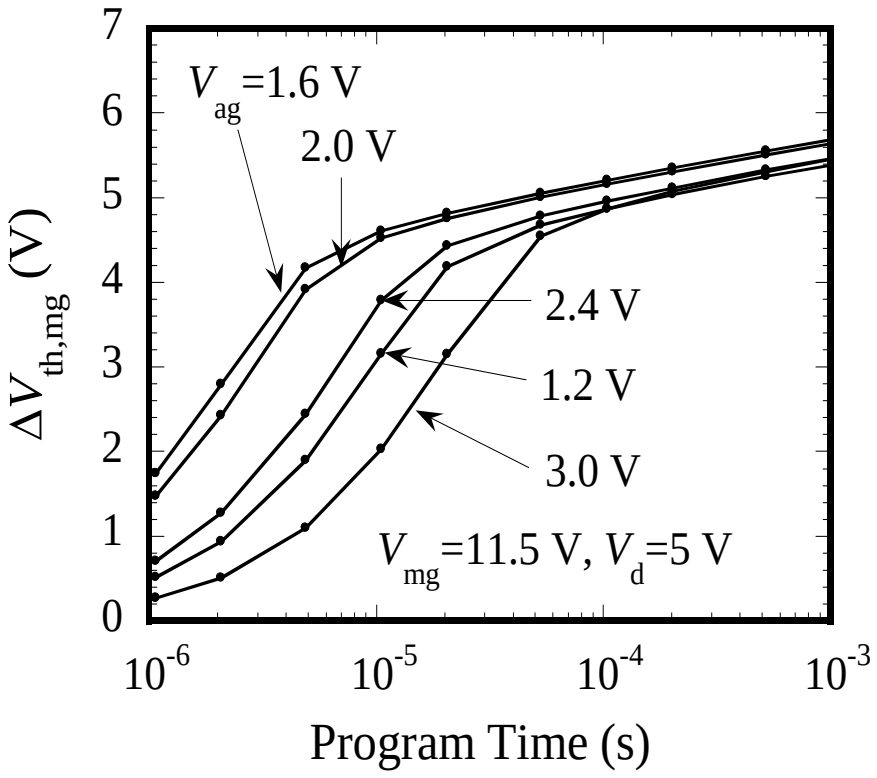


Fig. 3.8. $\Delta V_{th,mg}$ as a function of program time at V_{mg} = 11.5 V and V_d = 5 V with V_{ag} as the parameter [1.37].

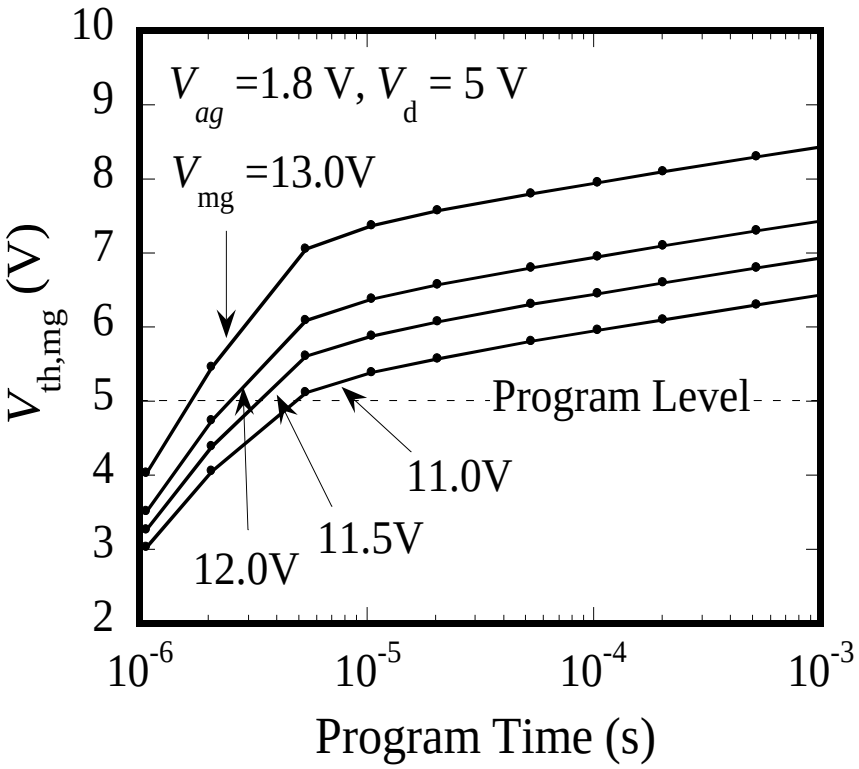


Fig. 3.9. $V_{th,mg}$ as a function of program time at V_{ag} = 1.8 V and V_d = 5 V with V_{mg} as the parameter [1.37].

Figure 3.9 shows $V_{th,mg}$ as a function of program time at V_{ag} = 1.8 V and V_d = 5 V with V_{mg} as the parameter. Fast programming less than 5 μs is achieved with $V_{mg} \geq 11$ V. $V_{th,mg}$ increases exponentially with an increase in program time, and then begins to saturate at a program time above 10 μs, because the FG channel enters the saturation regime at $V_{fg} - V_{th,fg} < V_d$. Here, $V_{th,fg}$ is the threshold voltage of the FG transistor. Program speed is enhanced with an increase in V_{mg} up to 13 V. These results suggest that the enhanced lateral field can be maintained over a wide range of V_{mg}, which is important from the viewpoint of using this concept in a memory cell.

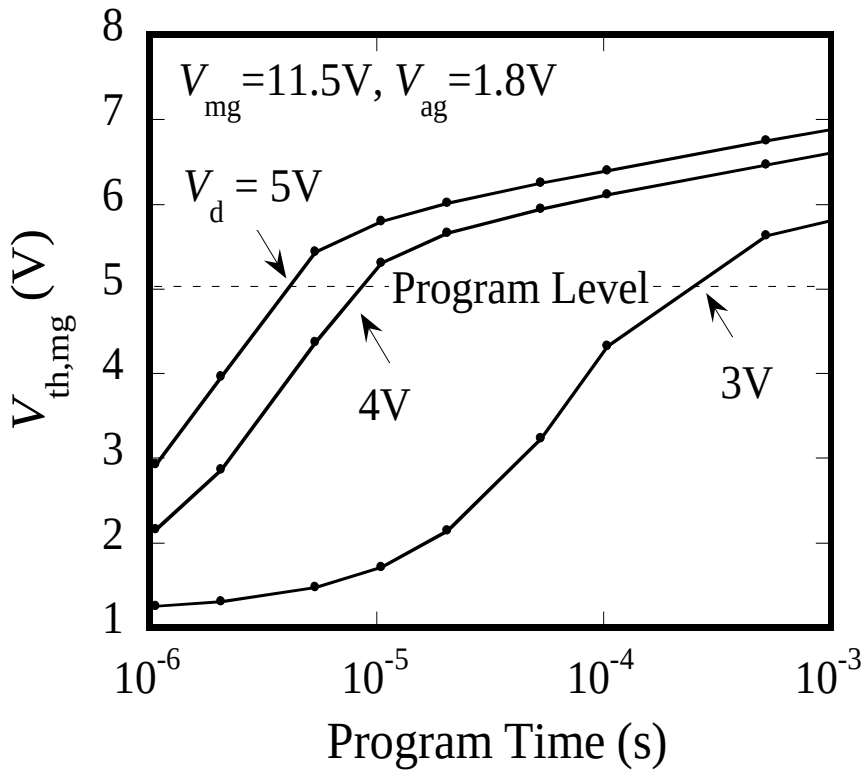


Fig. 3.10. $V_{th,mg}$ as a function of program time at V_{mg} = 11.5 V and V_{ag} = 1.8 V with V_d as the parameter [1.37].

Figure 3.10 shows $V_{th,mg}$ as a function of program time at V_{mg} = 11.5 V and V_{ag} = 1.8 V with V_d as the parameter. The S4-NOR cell can be programmed in less than 10 μs with V_d ≥ 4 V. Its V_d can be reduced by about 1 V as compared with that of the conventional CHE programming used for STD-NOR.

2) Erase

Figures 3.11 and 3.12 shows the effects of R_{mg} and V_d on erase performance. Erase speed is enhanced with an increase in R_{mg} as shown in Fig. 3.10(a). In this chapter, R_{mg}= 0.7 is used from the viewpoint of program disturb immunity, as will be described in Section. V.C. Figure 3.10(b) shows $V_{th,mg}$ as a function of erase time with V_d as the parameter, where the target $V_{th,mg}$ of the erase level is 2 V. The erase operation is completed in 100 ms with V_d = 10.5 V and the one-order-of-magnitude reduction in erase time can be achieved by increasing V_d by about 1 V.

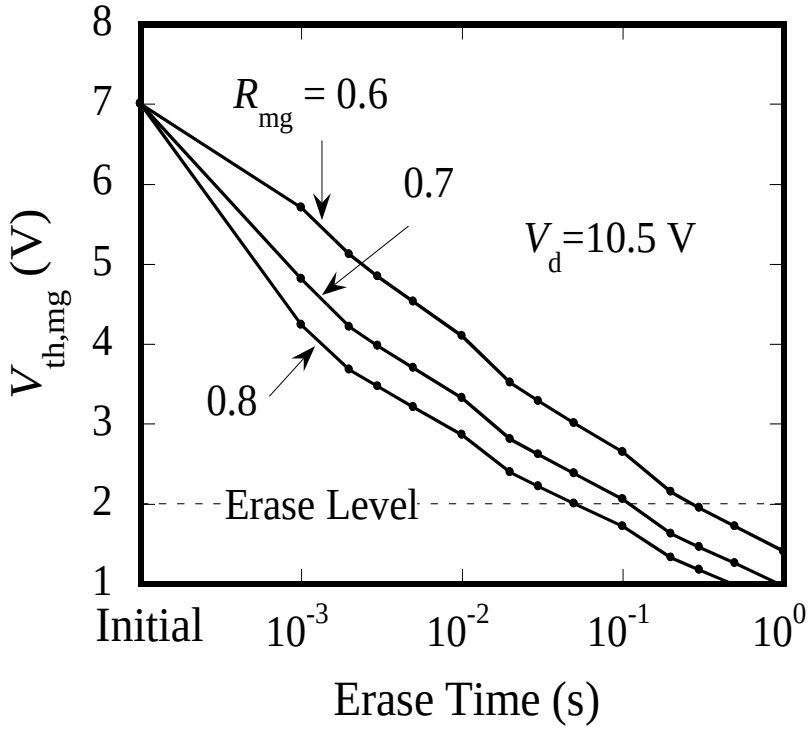


Fig. 3.11. $V_{th,mg}$ as a function of erase time at V_d = 10.5 V with R_{mg} as the parameter [1.37].

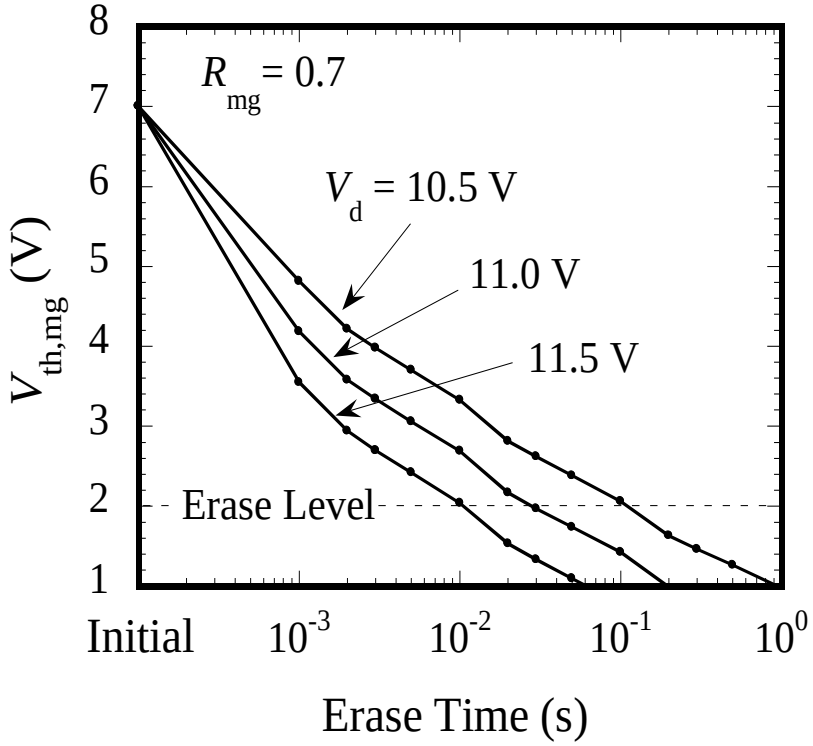


Fig. 3.12. $V_{th,mg}$ as a function of erase time with V_d as the parameter for the memory cells with R_{mg} = 0.7 [1.37].

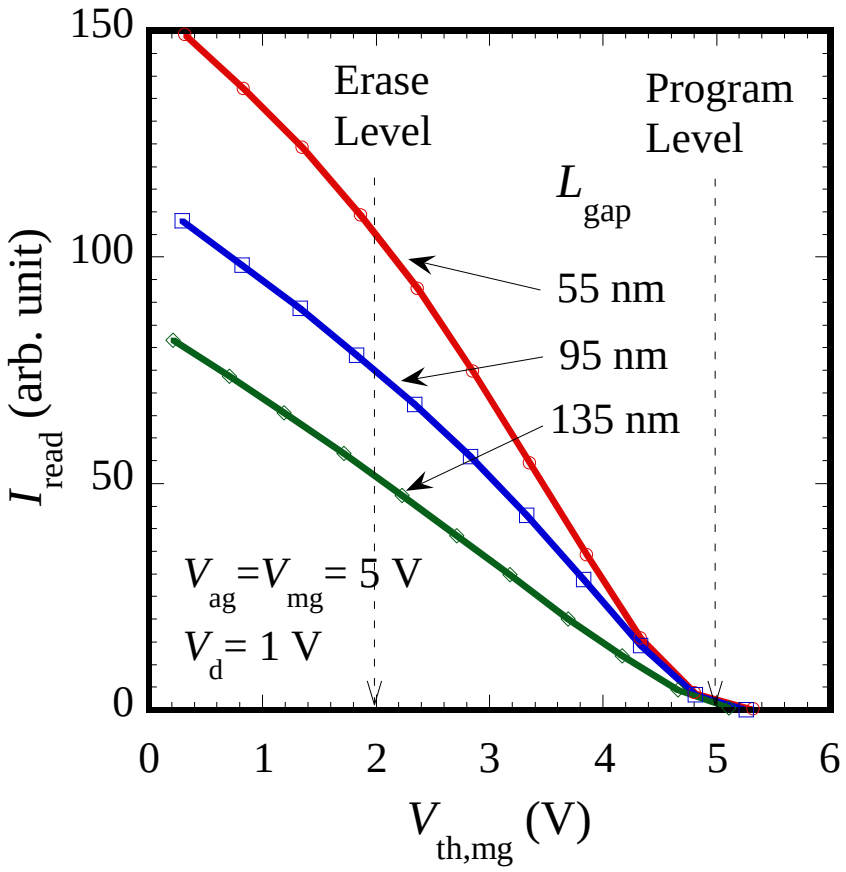


Fig. 3.13. I_{read} as a function of $V_{th,mg}$ at $V_{mg} = V_{ag}$ =5 V and V_d = 1 V with L_{gap} as the parameter [1.37].

3) Read

In the split-gate transistor case, the channel resistance in the gap region (R_{gap}) is controlled by the fringing field from the AG and FG. Figure 3.13 shows the read current (I_{read}) at $V_{mg} = V_{ag}$ = 5 V and V_d = 1 V as a function of $V_{th,mg}$, which was measured on the memory cells with different L_{gap} values. A higher I_{read} is obtained with a smaller L_{gap} for the typical erased memory cell ($V_{th,mg}$ = 2 V), while I_{read} is slightly dependent on L_{gap} for the typical programmed memory cell ($V_{th,mg}$ = 5 V). These results indicate that the transfer conductance of the split-gate transistor is considered to be strongly dependent on R_{gap}.

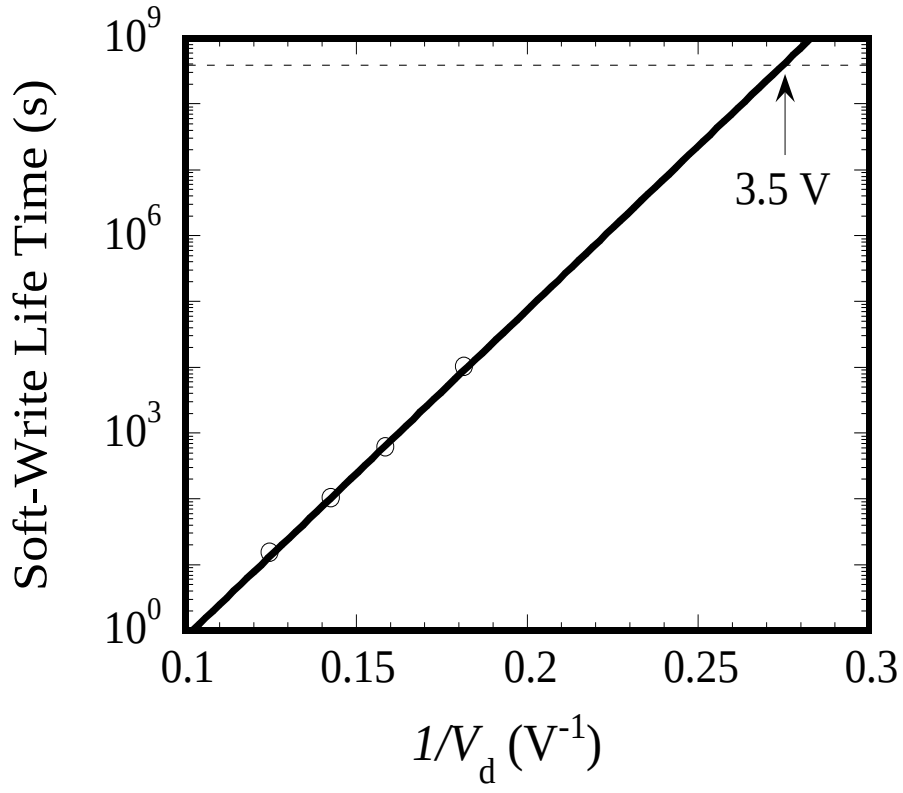


Fig. 3.14. Soft write lifetime as a function of $1/V_d$ [1.37].

Figure 3.14 shows soft write lifetime as a function of $1/V_d$ for a typical single cell. Soft write lifetime is defined as the DC stress time (t_{st}) that induces a $V_{th,mg}$ shift of 0.1 V. A 10-year lifetime, which is predicted by the extrapolation of the slope of soft write lifetime vs $1/V_d$, can be guaranteed at a V_d less than 3.5 V. It is clear that the S4-NOR cell can satisfy a higher soft write immunity than the STD-NOR cell [3.21] – [3.23], [3.33]. This is because of the large reduction in the lateral field in the gap region during read-out operation as discussed in Sections III.C and IV.A

B. Endurance and data retention

The key reliability issues for a flash memory are its endurance and data retention characteristics. Endurance evaluation was conducted on the typical single cells in a memory array and data retention evaluation was conducted on 4 Kbit test chips.

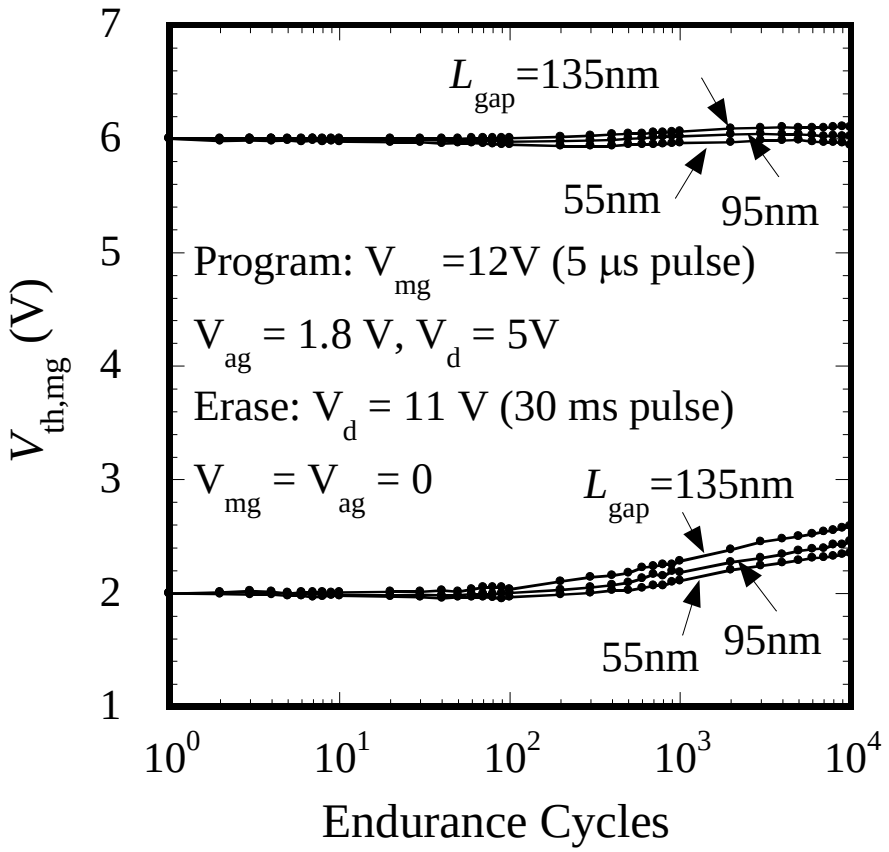


Fig. 3.15. Single-memory-cell endurance characteristics for different L_{gap} values [1.37].

1) Endurance

Figure 3.15 shows the single memory cell endurance characteristics for different L_{gap} values. Fixed program and erase pulses are applied without verifying the cell threshold voltage. The memory cell is programmed by applying a single pulse on the MG with an amplitude of 12 V and a width of 5 μs, while keeping V_{ag} = 1.8 V and V_d = 5 V. Moreover, the memory cell is erased with a single pulse on the drain with an amplitude of 11 V and a width of 30 ms, while keeping V_{mg} = V_{ag} = 0. Little dependence of L_{gap} on endurance characteristics is observed. Moreover, the low state $V_{th,mg}$ shows an increase after 1000 cycles, while the high state $V_{th,mg}$ only shows a minor change, similarly to previous studies [3.34]. A smaller increase in the low state $V_{th,mg}$ is observed with a smaller L_{gap}. The major reasons for the increase in the low state $V_{th,mg}$ are estimated to be the decrease in erase speed because of electron trapping in gate oxide in the gate-to-drain overlap region by FN tunneling and also read current reduction because of electron trapping in the gap region by SSHE injection. The program mechanism is considered to be insensitive to

the charge trapped in the gap region. In real devices, intelligent algorithms as well as process optimization are used to prevent the $V_{th,mg}$ window closing.

2) Data retention

Data retention evaluation was conducted at 250 ℃ after 1000 cycles of program and erase operations at room temperature under typical bias conditions shown in Table 3.1. Figure 3.16 shows the distribution of the minimum $V_{th,mg}$ in each test chip with bake time as the parameter, where L_{gap} is 55 nm. Even after 1000 h of bake time, the decrease in mean $V_{th,mg}$ is less than 100 mV. These results illustrate the good quality of the insulating layers inside the S4-NOR cell.

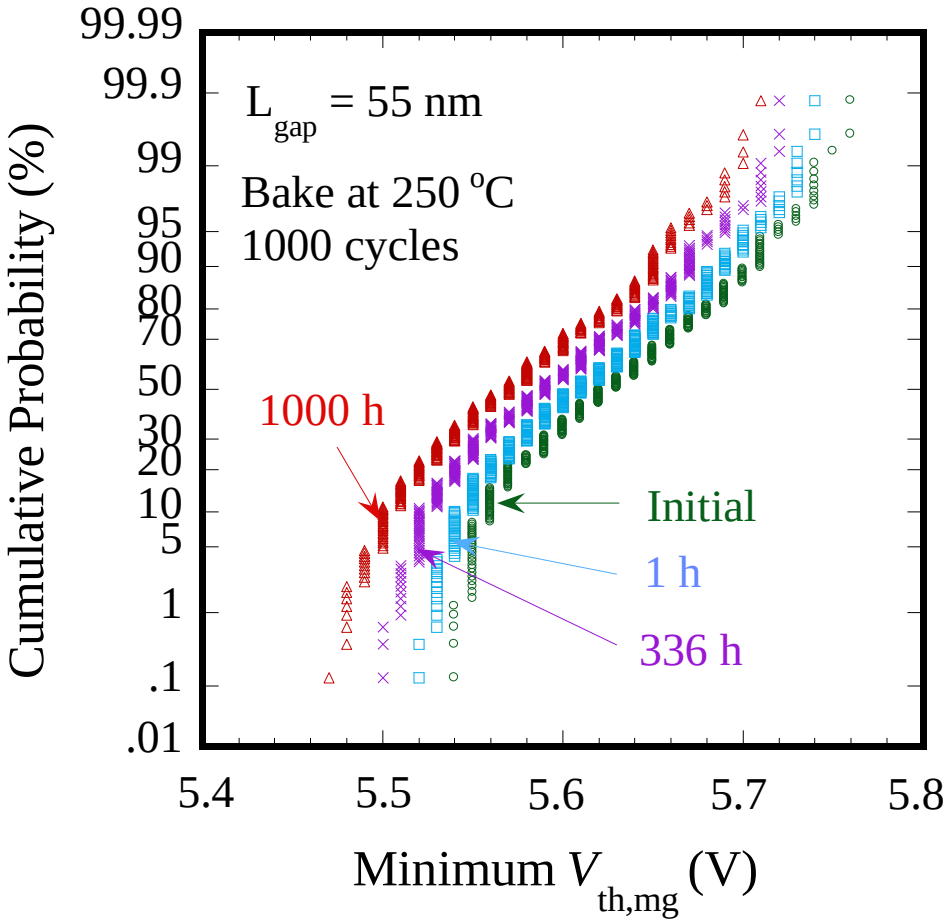


Fig. 3.16. Distribution of the minimum $V_{th,mg}$ in an each test chip for different bake times at 250 °C, which was measured after 1000 cycles of program and erase operations at room temperature [1.37].

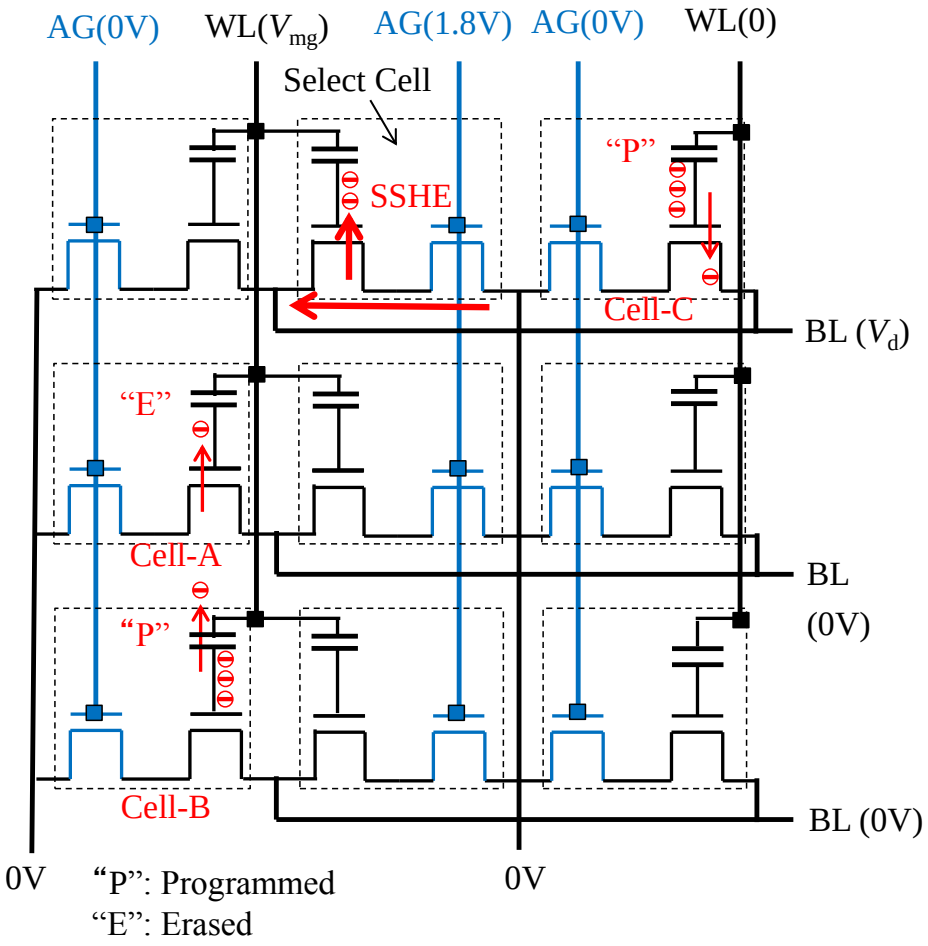


Fig. 3.17. Schematic description of program disturb effects on the different memory cells (Cell-A, Cell-B, and Cell-C) of a memory array [1.37].

C. Program disturb immunity

All the flash memory cells are subjected to program disturb mechanisms in the memory cells sharing a common wordline (WL) or bitline (BL), while one of the memory cells is being programmed [3.35]. The schematic description of the program disturb effects on different cells (Cell-A, Cell-B, and Cell-C) of a memory array is shown in Fig. 3.17. The MGs of plural memory cells are connected to the same WL and the drains of plural memory cells are connected to the same BL. When the number of memory cells connected to a WL or a BL is 1024, the longest disturbance time is about 10 ms. In previous papers, a remarkable degradation of the gate disturb immunity is observed under the initial stage of endurance up to 100 cycles because of hole trapping in the gate oxide during erase [3.35]. The disturb immunity evaluations were conducted on single memory

cells at room temperature after 100 cycles of program and erase operations. The good disturb immunity above 10 s is obtained for both the gate and drain bias stressing during programming as will be described below, which will be acceptable for actual device operation.

1) Gate disturb effect

Gate disturb occurs in the unselected memory cells (Cell-A and Cell-B in Fig. 3.17) connected to the same WL as the selected memory cell being programmed. During programming operation, the common WL is connected to a high V_{mg}.

For the erased Cell-A, the high voltage applied across the gate oxide between the FG and the substrate may cause electron injection into the FG and leads to an increase in $V_{th,mg}$. Figure 3.18 shows $V_{th,mg}$ as a function of the gate bias t_{st} at V_{mg} =12V with R_{mg} as the parameter for Cell-A. $\Delta V_{th,mg}$ increases with an increase in R_{mg} and $V_{th,mg}$ is maintained below the erase-level voltage (2 V) under gate bias stressing for 10 s for $R_{mg} \leq$ 0.75.

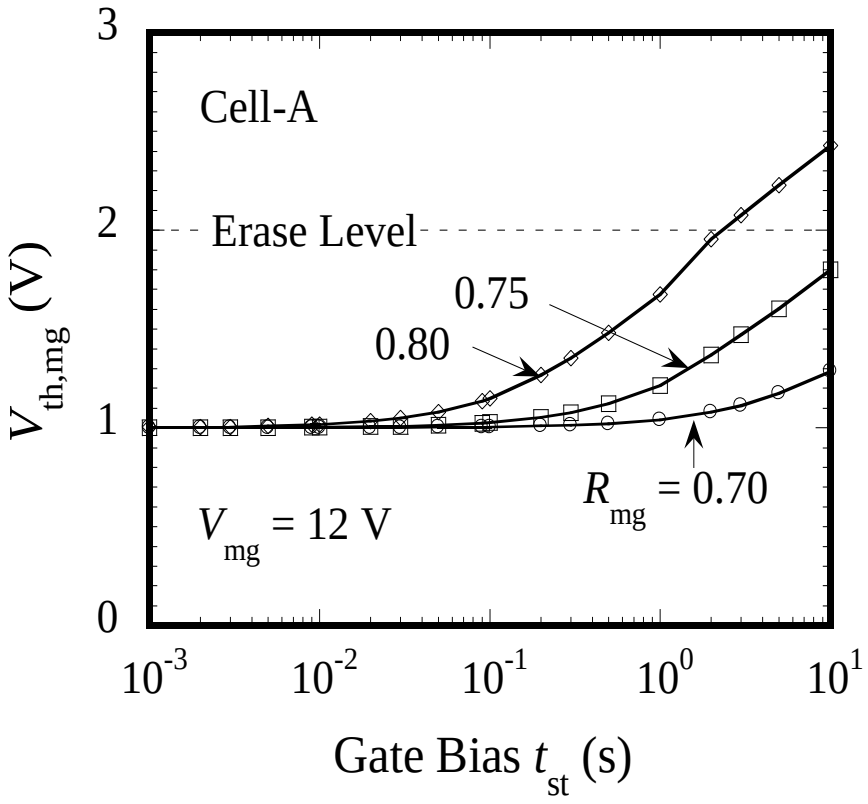


Fig. 3.18. $V_{th,mg}$ as a function of gate bias t_{st} at V_{mg} =12V with R_{mg} as the parameter for Cell-A [1.37].

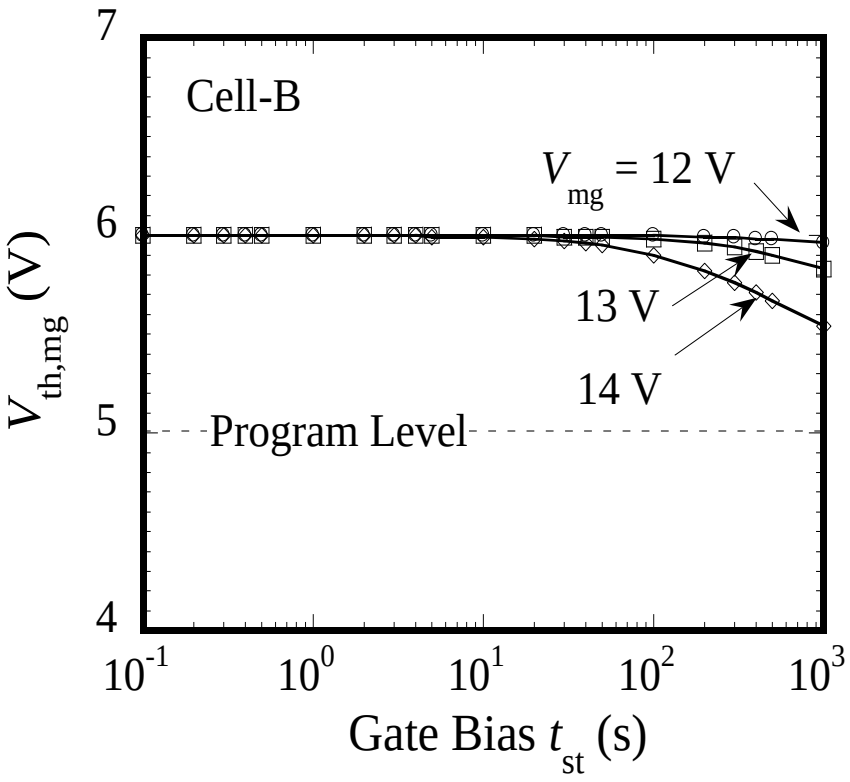


Fig. 3.19. $V_{th,mg}$ as a function of gate bias t_{st} with V_{mg} as the parameter for Cell-B [1.37].

On the other hand, for the programmed Cell-B, the high voltage applied across the gate oxide between the FG and the MG may cause electron extraction from the FG, which leads to a decrease in $V_{th,mg}$. Figure 3.19 shows $V_{th,mg}$ as a function of gate bias t_{st} with V_{mg} as the parameter for Cell-B. No change in $V_{th,mg}$ is observed under gate bias stressing for 10 s at $V_{mg} \leq 14$ V.

The average cell leakage current through the gate oxide between the FG and the substrate (I_{ox}) and the electrical field across its gate oxide (E_{ox}) can be roughly estimated from the change in $V_{th,mg}$ under gate bias stressing on the erased Cell-A with different V_{mg} values, using the following simple expressions:

$$I_{ox} = \frac{C_{mg}\,\Delta V_{thst,mg}}{t_{st}}, \tag{5}$$

$$E_{ox} = R_{mg}\,\frac{V_{mg} - (V_{th0,mg} + \Delta V_{thst,mg} - V_{thuv,mg})}{t_{ox}}, \tag{6}$$

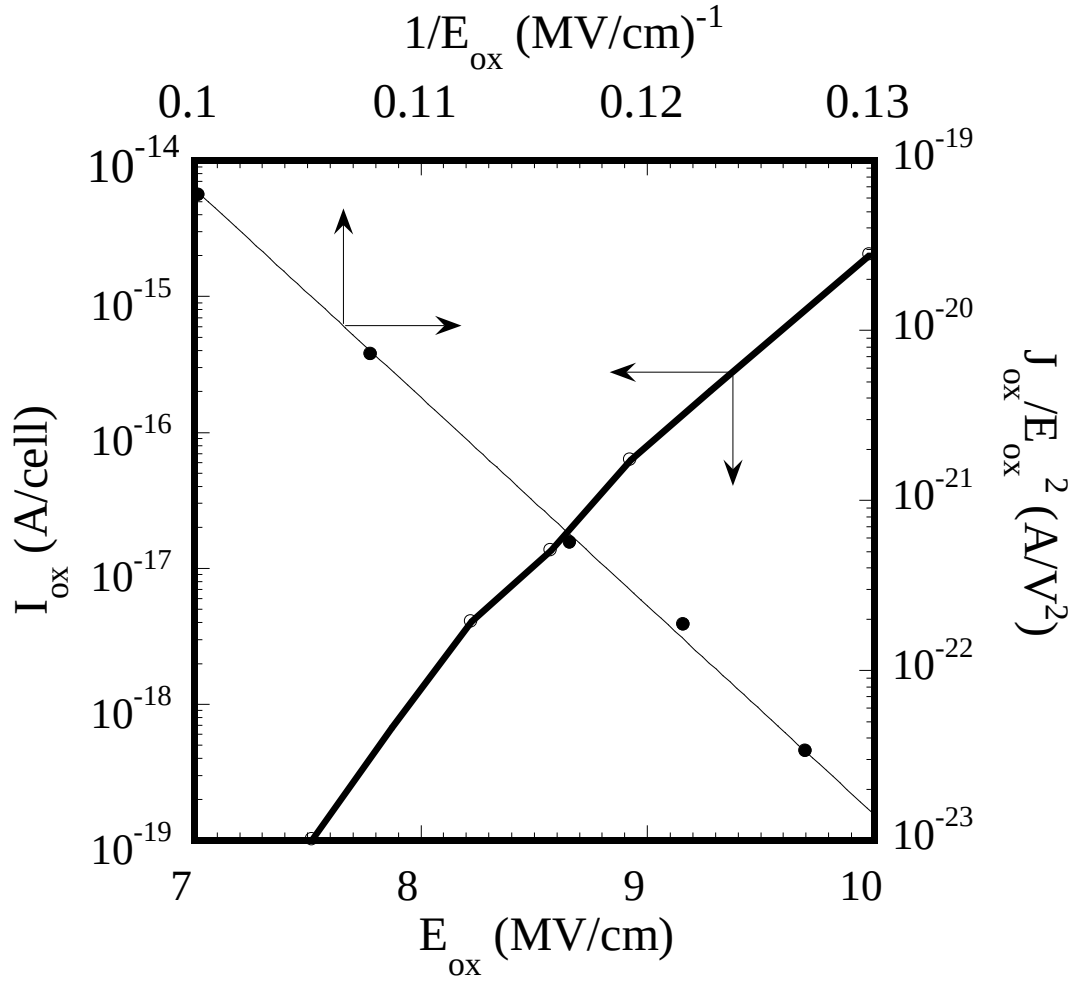


Fig. 3.20. I_{ox} as a function of E_{ox} together with its FN plot of J_{ox}/E_{ox}^2 as a function of $1/E_{ox}$ [1.37].

where $\Delta V_{thst,mg}$ is the change in $V_{th,mg}$ by the gate-bias stressing, t_{st} is the stress time, $V_{th0,mg}$ is the $V_{th,mg}$ before stressing, $V_{thuv,mg}$ is the $V_{th,mg}$ after UV erase, and t_{ox} is the gate oxide thickness (10 nm). Here, R_{mg}= 0.7 and C_{mg} = 0.58 fF.

Figure 3.20 shows I_{ox} as a function of E_{ox}, which is obtained by substituting $\Delta V_{thst,mg}$ = 100 mV and t_{st} for $\Delta V_{thst,mg}$ = 100 mV into Eqs. (3.5) and (3.6). The FN plot of J_{ox}/E_{ox}^2 vs $1/E_{ox}$ is also added in Fig. 3.20, where J_{ox} is defined by I_{ox}/S_{fgsb} and S_{fgsb} is the area of the gate oxide between the FG and the substrate. The FN plot shows a straight line with a slope of 2.8×10^8 V/cm, which is in good agreement with the FN tunneling model [3.36]. This result indicates that the $V_{th,mg}$ shift of Cell-A under gate-bias stressing is mainly caused by FN tunneling current without the effect of the leakage across the gap oxide.

2) Drain disturb effect

Drain disturb occurs in the unselected memory cells in the programmed state (Cell-C in Fig. 3.17) that are on the same bit line as the selected memory cell being programmed. These memory cells will experience a high electric field between the FG and the drain. This may cause tunneling electrons from the FG to the drain and leads to a reduction in $V_{th,mg}$. Figure 3.21 shows $V_{th,mg}$ as a function of the drain bias t_{st} with V_d as the parameter. Under drain-bias stressing for 10 s, the Cell-C shows no change in $V_{th,mg}$ at V_d = 5 V and can maintain $V_{th,mg}$ above the program level (5 V) even at V_d = 6 V.

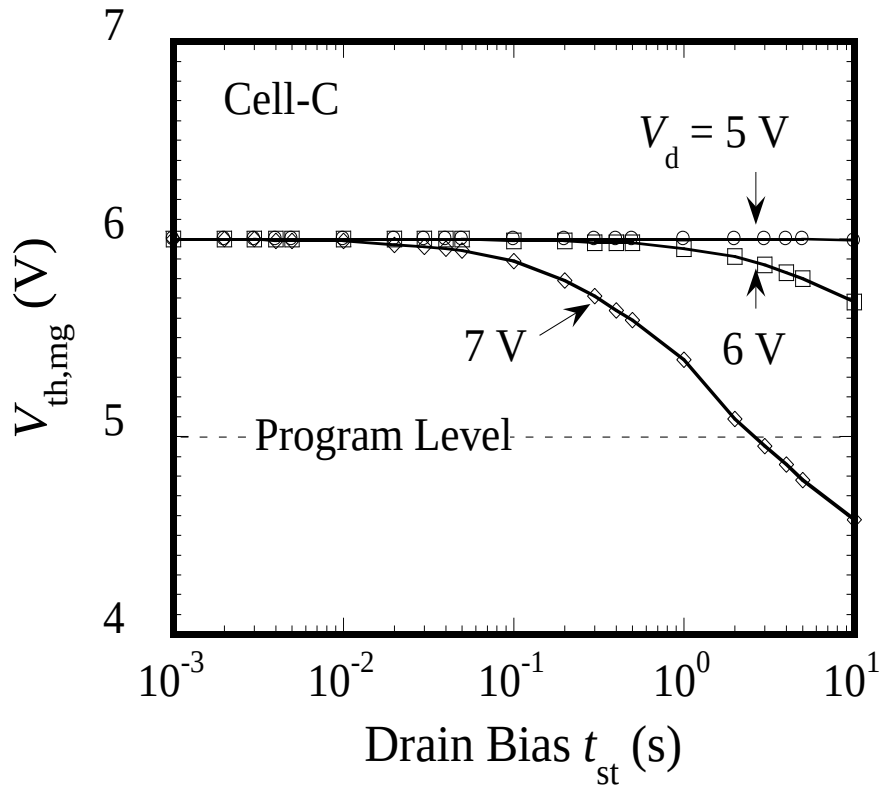


Fig. 3.21. Program disturb characteristics under drain-bias stressing. $V_{th,mg}$ as a function of drain bias t_{st} with V_d as the parameter for Cell-C [1.37].

VI. CONCLUSIONS

A logic-process-compatible S4-NOR flash memory technology is developed. For the S4-NOR cell, the AG and FG are patterned simultaneously with L_{gap} = 55 nm on the channel between the source and the drain, and the FG is capacitively coupled to the n-well MG through the gate oxide. This cell technology can improve the compatibility with a single-polysilicon CMOS logic process. A fast programming time of 5 μs is achieved with a low drain current of 10 μA. Its injection efficiency is higher by four orders of magnitude than the injection efficiency of the STD-NOR cell and is insensitive to the process parameters except L_{gap}, which is suitable for embedded memories from the viewpoint of process control. It also offers low power consumption. High read current is obtained without suffering from soft write effects. In addition, good endurance, data retention, and good program disturb immunity are obtained without the effect of gap oxide leakage. Furthermore, due to the simple structure of the S4-NOR cell, it does not require any specialized process steps.

The proposed S4-NOR flash memory is a promising candidate for next-generation embedded applications that require low power consumption and high performance without an increase in the cost of the logic chip.

REFERENCES

[3.1] A. Niebel, "Embedded Nonvolatile Memory Markets on the Rise," *IEEE International Memory Work Shop*, 2011.

[3.2] R. Strenz, "Embedded Flash Technologies and their Applications: Status and Outlook," *International Electron Devices Meeting Technical Digest*, 2011, pp. 211 – 214.

[3.3] K. Barker, "Embedded Nonvolatile Memories," *IEEE International Memory Work Shop*, 2012.

[3.4] M. Kamiya, Y. Kojima, Y. Kato, K. Tanaka, and Y. Hayashi, "EPROM Cell with Gate Injection Efficiency," *International Electron Devices Meeting Technical Digest*, 1982, pp. 741 – 744.

[3.5] A. T. Wu, T. Y. Chan, P. K. KO, and C. Hu, "A Novel High-Speed, 5-V Programming EPROM Structure with Source-Side Injection," *International Electron Devices Meeting Technical Digest*, 1986, pp. 584 – 587.

[3.6] Y. Yamauchi, M. Yoshimi, K. Omori, S. Nitta, N. Mizukoshi, and K. Suzuki, "A Novel Source-side Injection Split Gate Flash Cell Technology for High Density Low Voltage Applications," *International Conference Solid State Devices and Materials*, 1994, pp. 910 − 912.

[3.7] J. V. Houdt, L. Haspeslagh, D. Wellekens, L. Deferm, G. Groeseneken, and H. E. Maes, *IEEE Transaction on Electron Devices*, vol. 40, no. 12, pp. 2255 – 2263, Dec. 1993.

[3.8] J. V. Houdt, D. Wellekens, L. Haspeslagh, L. Deferm, G. Groeseneken, and H. E. Maes, "A 5-V/3.3-V Compatible Flash EEPROM Cell with a 400-ns/70-µs Programming Time for Embedded Memory Applications," *Biennial Nonvolatile Memory Technology Review*, 1993, p. 54.

[3.9] H. Pein and J. D. Plummer, "A 3-D Sidewall Flash EPROM Cell and Memory Array, *IEEE Electron Device Letters*, vol. 14, no. 8, pp. 415 – 417, Aug. 1993.

[3.10] Y. Ma, C. S. Pang, J. Pathak, S. C. Tsao, C. F. Chang, M. Y. Yamauchi, and M. Yoshimi, "A Novel High Density Contactless Flash Memory Array using Split-Gate Source-Side-Injection Cell for 5V-Only Applications," *Symposium on VLSI Technology Digest of Technical Papers*, 1994, pp. 49 – 50.

[3.11] Y. Ma, C. S. Pang, K. T. Chang, S. C. Tsao, J. F. Frayer, T. Kim, K. Jo, J. Kim, I Choi, and H. Park, "A Dual-bit Split-Gate EEPROM (DSG) Cell in Contactless Array for Single-Vcc High Density Flash memories," *International Electron Devices Meeting Technical Digest*, 1994, pp. 57 – 60.

[3.12] S. Ogura, A. Hori, J. Kato, M. Yamanaka, S. Odanaka, H. Fujimoto, K. Akamatsu, T. Ogura, M. Kojima, and H. Kotani, "Low Voltage, Low Current, High Speed Program Step Split Gate Cell with Ballistic Direct Injection for EEPROM/Flash," *International Electron Devices Meeting Technical Digest*, 1998, pp. 987 – 990.

[3.13] T. Kobayashi, Y. Sasago, H. Kurata, S. Saeki, Y. Goto, T. Arigane, Y. Okuyama, H. Kume, and K. Kimura, "A Giga-Scale Assist-Gate (AG)-AND-Type Flash Memory Cell with 20-MB/s Programming Throughput for Content-Downloading Applications," *International Electron Devices Meeting Technical Digest*, 2001, pp. 29 – 32.

[3.14] Y. Tkachev, X. Liu, and A. Kotov, "Floating-Gate Corner-Enhanced Poly-to-Poly Tunneling in Split-Gate Flash Memory Cells," *IEEE Transaction on Electron Devices*, vol. 59, no.1, pp. 5 – 11, Jan. 2012.

[3.15] S. Bhattacharya, K. Lai, K. Fox, P. Chan, E. Worley, and U. Sharrna, "Improved Performance and Reliability of Split Gate Source-Side Injected Flash Memory Cells," in *International Electron Devices Meeting Technical Digest*,1996, pp. 339 – 342.

[3.16] D. Shum, J. R. Power, R. Ullmann, E. Suryaputra, K. Ho, J. Hsiao, C. H. Tan, W. Langheinrich, C. Bukethal, V. Pissors, G. Tempel, M. Röhrich, A. Gratz, A. Iserhagen, E. O. Andersen, S. Paprotta, W. Dickenscheid, R. Strenz, R. Duschl,

T. Kern, C.T. Hsieh, C.M. Huang, C.W. Ho, H.H. Kuo, C.W. Hung, Y.T. Lin, and L.C. Tran, "Highly Reliable Flash Memory with Self-aligned Split-gate Cell Embedded into High Performance 65nm CMOS for Automotive & Smartcard Applications," *IEEE International Memory Work Shop*, 2012.

[3.17] M. Slotboom, P. Goarin, N. Akil, M.V. Duuren, M. Demand, J.M.D. Wouters, S. Beckx, P. Leray, C. Baerts, N. Heylen, I. Pollentier, "Gate Isolation Technology for Compact Poly-CMP Embedded Flash Memories," *Europen Solid-State Device Research*, 2003, pp. 159 – 162.

[3.18] J. V. Houdt, P. Heremans, L. Defenn, G. Groeseneken, and H. E. Maes, "Analysis of the Enhanced Hot-Electron Injection in Split-Gate Transistors Useful for EEPROM Applications," *IEEE Transaction on Electron Devices*, vol. 39, no. 5, pp. 1150 – 1156, May 1992.

[3.19] J. V. Houdt, G. Groeseneken, and H. E. Maes, "An Analytical Model for the Optimization of Source-Side Injection Flash EEPROM Devices," *IEEE Transaction on Electron Devices*, vol. 42 no. 7, pp. 1314 – 1320, Jul. 1995.

[3.20] H. S. Wong, "Gate Current Injection Split-Gate in MOSFET's with a Split-Gate (Virtual Drain) Structure," *IEEE Electron Device Lett*ers vol. 14, no. 5. May 1993.

[3.21] H. Kume, H. Yamamoto, T. Adachi, T. Hagiwara, K. Komori, T. Nishimoto, A. Koike. S. Meguro, T. Hayashida, and T. Tsukada, "A Flash-Erase EEPROM Cell with an Asymetric Source and Drain Structure," *International Electron Devices Meeting Technical Digest*, 1987, pp. 560 – 563.

[3.22] N. Ajika, M. Ohi, H. Arima, T. Matsukawa and N. Tsubouchi, "A 5 V Only 16M Bit Flash EEPROM Cell with a Simple Stacked Gate Structure," *International Electron Devices Meeting Technical Digest*, 1990, pp. 115 – 118.

[3.23] N. Kodama, K. Saitoh, H. Shirai, T. Okazawa and Y. Hokari, "A 5 V Only 16 Mb Flash EEPROM Cell Using Highly Reliable Write / Erase Technologies," *Symposium VLSI Technology Digest of Technical Papers*, 1991, pp. 75 – 76.

[3.24] E. Takeda, H. Kume, T. Toyabe, and S. Asai, "Submicrometer MOSFET Structure for Minimizing Hot-Carrier Generation," *IEEE Transaction on Electron Devices*, vol. 29, no. 4, pp. 611 – 618, Apr. 1982.

[3.25] C. Hu, "Hot-Electron Effects in MOSFET's," *International Electron Devices Meeting Technical Digest*, 1983, p. 176 – 181.

[3.26] S. Tam, P. K. Ko, and C. Hu, "Lucky-Electron Model of Channel Hot-Electron Injection in MOSFETs," *IEEE Transaction on Electron Devices*, vol. 31, no. 9, pp. 1116 – 1125, Sep. 1984.

[3.27] C. Hu, S. C. Tam, F. Hsu, P. K. Ko, T. Chan, and K. W. TERRILL, "Hot-Electron-Induced MOSFET Degradation – Model, Monitor, and Improvement," *IEEE Transaction on Electron Devices*, vol. 32, no. 2, pp. 375 – 385, Feb. 1985.

[3.28] T. Y. Chan, P. K. Ko, and C. Hu, "Dependence of Channel Electric Field on Device Scaling," *IEEE Electron Device Letters*, vo. 6, no. 10, pp. 551 – 553, Oct. 1985.

[3.29] K. Hasnat, C.-F. Yeap, S. Jallepalli, W.-K. Shih, S.A. Hareland, V.M. Agostinelli, A.F. Tasch, and C. M. Maziar, "A Pseudo-Lucky Electron Model for Simulation of Electron Gate Current in Submicron NMOSFET's," *IEEE Transaction on Electron Devices*, vo. 43, no. 8, pp. 1264 – 1273, Aug. 1996.

[3.30] T. H. Ning, C. M. Osburn, and H. N. Yu, "Emission Probability of Hot Electrons from Silicon into Silicon Dioxide," *Journal of Applied Physics*, vol. 48, pp. 286 – 293, 1977.

[3.31] L. Selmi, E. Sangiorgi, R. Bez, and B. Riccò, "Measurement of the Hot Hole Injection Probability from Si into SiO_2 in p-$MOSFET_S$," *International Electron Devices Meeting Technical Digest*, 1993, pp. 333 – 336.

[3.32] P. Palestri, N. Akil, W. Stefanutti, M. Slotboom, and L. Selmi, "Effect of the Gap Size on the SSI Efficiency of Split-Gate Memory Cells," *IEEE Transaction on Electron Devices*, vol. 53, no. 3, pp. 488 – 493, March 2006.

[3.33] J. F. Van. Houdt, D. Wellekens, G. Groeseneken, and H. E. Maes, "Investigation of the Soft-Write Mechanism in Source-Side Injection Flash EEPROM Devices," *IEEE Electron Device Letters*, vo. 16, no. 5, pp. 181 – 183, May 1995.

[3.34] D. Wellekens, J. Van. Houdt, L. Faraone, G. Groeseneken, and H. E. Maes, "Write/Erase Degradation in Source Side Injection Flash EEPROM's: Characterization Techniques and Wearout Mechanisms," *IEEE Transaction on Electron Devices*, vol. 42, no. 11, pp. 1992 – 1998, Nov. 1995.

[3.35] G. Verma and N. Mielke, "Reliability Performance of ETOX Based Flash Memories," *International Reliability Physics Symposium*, 1988, pp. 158 – 166.

[3.36] M. Lenzlinger and E. H. Snow, "Fowler-Nordheim Tunneling into Thermally Grown SiO_2," *J. Appl. Phys.*, vol. 40, pp. 278 – 283, 1969.

Chapter 4

Floating-Gate MOSFET Nonvolatile DRAM

I. Introduction 73
II. Cell Structure and Process Technology 74
III. Principle of Operations 79
A. DRAM Mode 80
B. Store Mode 80
C. Recall Mode 82
IV. Store Characteristics 84
V. Reliability of Floating-Gate MOSFET 87
VI. Scalable Memory Cell 89
A. Cell Structure 89
B. Operation Concept 90
VII. Conclusions 91
References 91

I. INTRODUCTION

Random access memory (RAM) is a high performance memory which has a less than 100 ns high speed read/write, and an unlimited read/write cycling. However, RAM needs a battery backup for data storage. Read only memory (ROM) can store data for more than 10 years without battery backup. However, a long write time and insufficient write endurance, less than 10^6 cycles, limit nonvolatile memory applications. The current system using RAM as main working memory and ROM as main storage memory, suffers from slow word-by-word serial data transfer between the DRAM and the storage ROM. Hence, to store data being processed in the event of sudden power failure, it must be continuously transferred from the DRAM chip to the ROM chip, leading to an increase in communication power and latency.

Universal memories such as ferro-electric RAM, phase change memory, resistive

RAM, and magnetic RAM [1.20] – [1.26] have been studied for a long time for next generation applications, since they offer the performance of RAM with the nonvolatility of ROM. However, these new memories have not been implemented into high-density stand-alone applications, because they require more time to improve yield in the mass production of memory devices using the new material.

Recently, many studies on nonvolatile SRAM (NV-SRAM) [1.27] – [1.29], which combines SRAM with the new material nonvolatile memory devices, have been reported for early mass production. NV-SRAM works as SRAM in normal operation, the data in which can be stored to nonvolatile memory even during power down. However, NV-SRAM requires a very large cell size because of the complicated cell structure of SRAM, which is not suitable for high-density applications.

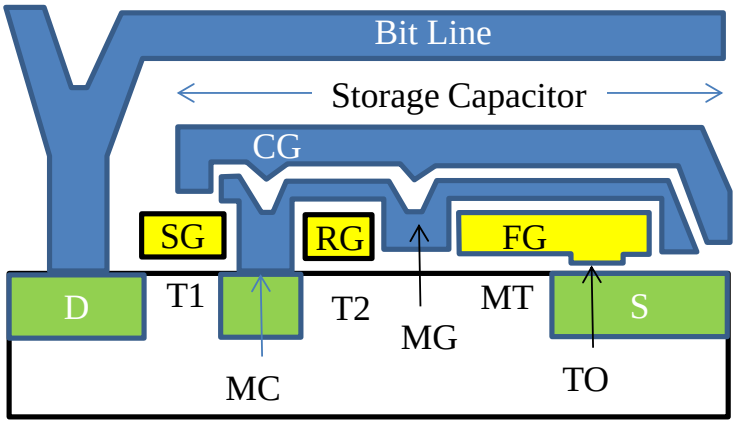
Studies on hybrid main memory [4.1] – [4.4], which combines DRAM with the new material nonvolatile memory, also have been reported for reducing power of low-power and high-performance mobile applications.

This chapter describes a nonvolatile DRAM (NV-DRAM) cell, which combines DRAM with the floating-gate metal-oxide-semiconductor field-effect transistor (FG-MOSFFT) without using any new materials [4.5] – [4.10]. The chip data can be simultaneously backed up to FG-MOSFFT by using the high efficient FN tunneling mechanism without readout from DRAM. These technologies are successfully implemented into a 1 Mbit test chip. In Section II, NV-DRAM cell structure with a merged FG-MOSFET and its key process technologies are illustrated together with a 1 Mbit chip photograph. In Section III, principles of operations are studied. In Section IV, store characteristics are measured for single cells, which are compared with the simulation results. In Section V, endurance and data retention of FG-MOSFFT are evaluated at high temperature in addition to soft-write immunity of FG-MOSFET during DRAM operation at room temperature. In Section VI, scalable NV-DRAM cell is discussed. Finally, conclusions are summarized in Section VII.

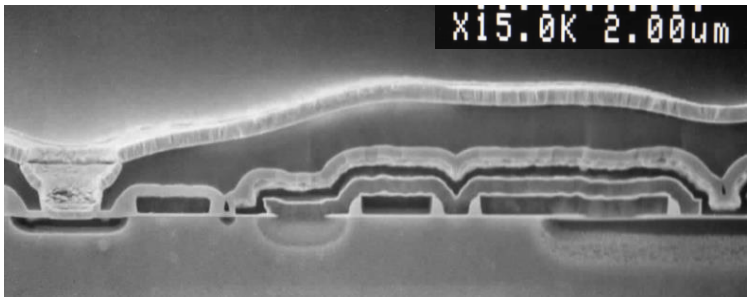
II. CELL STRUCTURE AND PROCESS TECHNOLOGY

The cell structure of FG-MOSFET NV-DRAM is shown in Fig. 4.1. The cell consists of three transistors (T1, T2, MT) and a storage capacitor with a capacitance of C_s. T1 is the select transistor and its drain (D) and gate (SG) are connected to the bit-line and the word-line, respectively. T2 is the recall transistor. MT is the stacked gate memory transistor (MT), which has the FG with tunnel oxide (TO) on the source (S). DRAM data is stored on the gate (MG) of MT, which is connected to the substrate through the

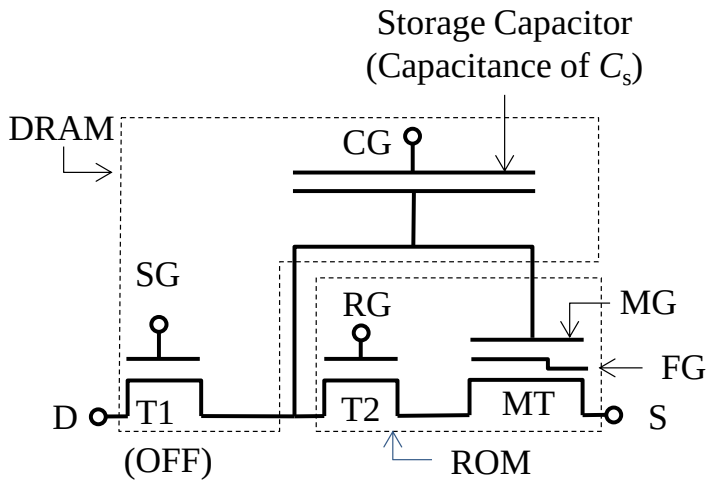
memory contact (MC) between the SG and the recall gate (RG). The DRAM storage capacitor is formed between the MG and the capacitor gate (CG) over the ROM element.



(a) Schematic Cross-Sectional View



(b) SEM View



(c) Equivalent Circuit

Fig. 4.1. Cross-sectional view of the FG-MOSFET NV-DRAM cell [1.38].

Key process steps for the FG-MOSFET NV-DRAM cell are shown in Fig. 4.2. First, the source regions (S) are formed and then the TO is formed on it after the first gate oxidation. The first level polycrystalline silicon (polysilicon) (poly-1) with poly-oxide/ LPCVD-nitride/LPCVD-oxide stacked film on its surface is deposited and patterned to form the cell gates (SG, RG, FG), followed by the lightly doped drain implantation (Fig. 4.2(a)). The peripheral high voltage transistor gates are also formed by the poly-1 with the first gate oxide. Next, an LPCVD-nitride film is deposited and anisotropically etched back stopping on the underlying LPCVD-oxide to form nitride spacers on the side-wall of the cell gates (Fig. 4.2(b)). Then the top LPCVD-oxide on the poly-1 is removed. After reoxidation and removal of oxide on the memory contact (MC), the second level polysilicon (poly-2) is deposited and patterned to form the memory gate (MG) (Fig. 4.2(c)). Next the oxide/nitride/oxide (ONO) stacked film is formed on the poly-2 by poly-oxidation, LPCVD-nitride deposition, and the subsequent second gate oxidation. The tungsten silicide on polysilicon (polycide) was then deposited and patterned to form the capacitor gate (CG) (Fig. 4.2(d)). The peripheral high-speed transistor gates are also formed by the tungsten polycide with the second gate oxide.

A twin-well triple polysilicon (double polysilicon and single polycide) and double metal CMOS process based on 0.8 μm design rules has been developed for 1 Mbit NV-DRAM. A total of 24 masks is used to build the 1 Mbit NV-DRAM. The key process features are shown in Table 4.1. The first gate oxide thickness of 35 nm is used for memory cell and high voltage transistors and the second gate oxide thickness of 17 nm is used for peripheral high-speed transistors. The TO of approximately 8 nm is grown in a vertical and load-locked type furnace, followed by post-oxidation annealing at high temperature. The capacitive equivalent oxide thickness of the interpoly oxide/nitride (ON) film between poly-1 and poly-2 and the interpoly oxide/nitride/oxide (ONO) film between poly-2 and poly-3 are 20 nm and 14 nm, respectively.

By utilizing the above cell structure and process technologies, a 30.94 μm^2 cell with C_s = 50 fF has been successfully implemented into the 1 Mbit NV-DRAM with a chip size of 85.91 mm^2. A photograph of the 1 Mbit NV-DRAM chip is shown in Fig. 4.3

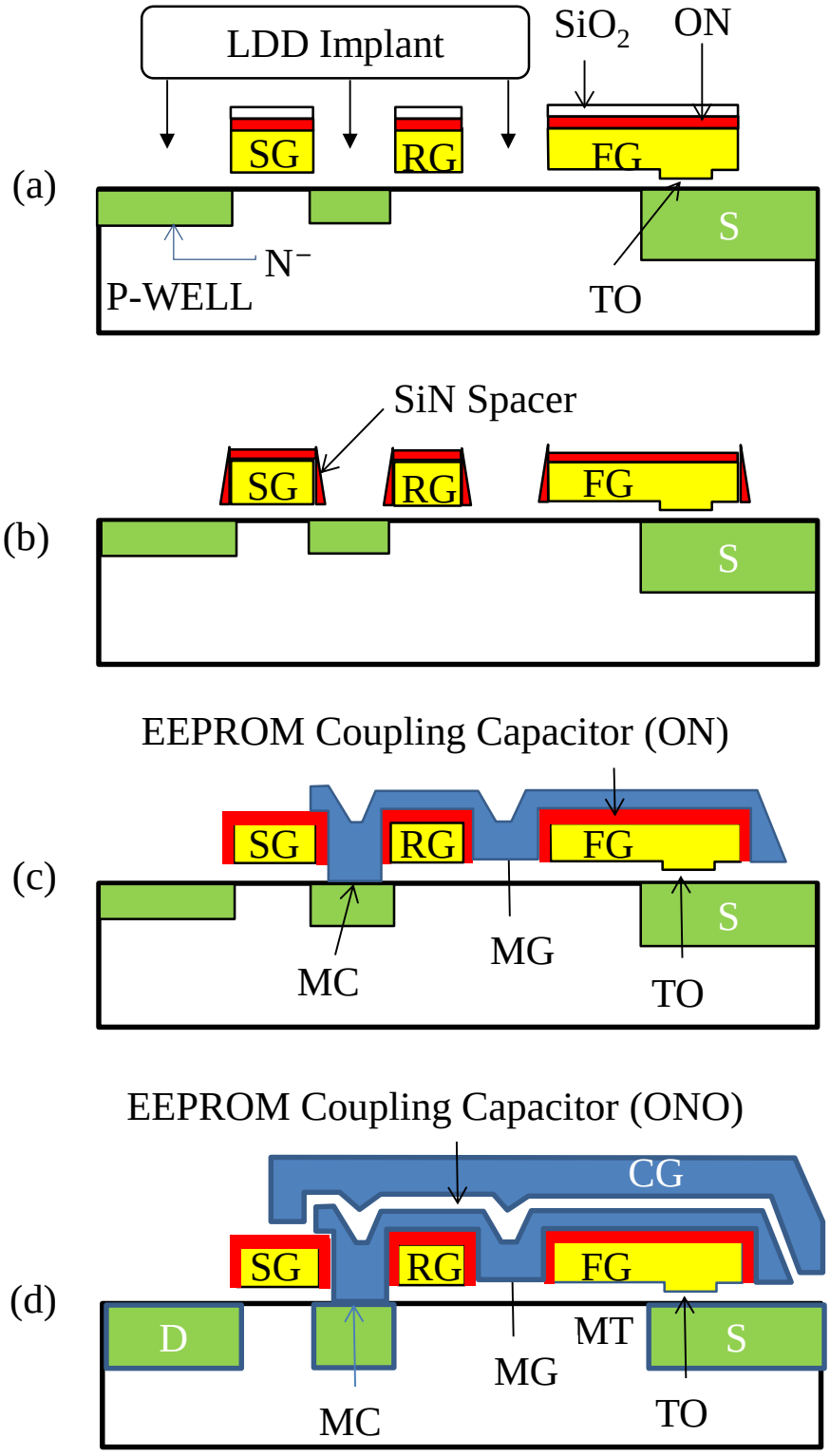


Fig. 4.2. Key process steps for FG-MOSFET NV-DRAM cell [1.38].

Table 4.1

Key Process Technology for 1 Mbit NV-DRAM [1.38].

Process	0.8μm Twin Well CMOS Double Layers of Polysilicon Single Layer of Polycide Double Layers of Metal		
Field Oxide	600	nm	
Minimum Gate Length	Nch	0.8	μm
	Pch	0.9	μm
1st Gate Oxide	35	nm	
2nd Gate Oxide	17	nm	
Tunnel Oxide	8	nm	
Interpoly Dielectric	Poly1-Poly2	20	nm
(Effective Oxide Thickness)	Poly2-Poly3	14	nm
Cell Size	9.1×3.4	= 30.94	μm^2
Chip Size	13.55×6.34	= 85.91	mm^2

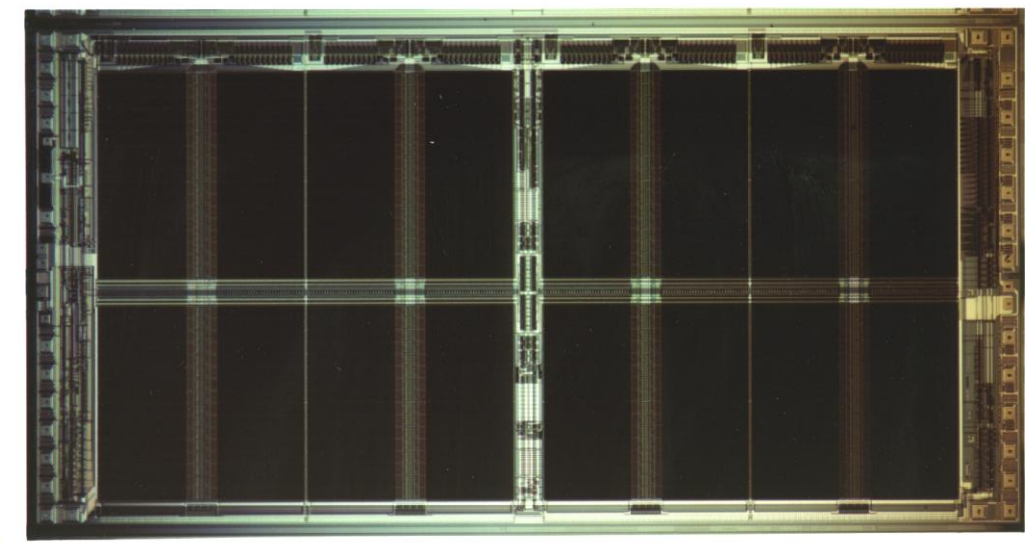


Fig. 4.3. Chip photograph of 1 Mbit NV-DRAM [1.38].

III. PRINCIPLES OF OPERATION

A timing diagram for DRAM-mode, STORE-mode (data transfer from DRAM to FG-MOSFET), and RECALL-mode (data transfer from FG-MOSFET to DRAM)) operations are shown in Fig. 4.4.

In order to gain insight into a store operation, a capacitive equivalent circuit, shown in Fig. 4.5, is used. C_{12} is the poly-1/poly-2 coupling capacitance, C_{g1} is the poly-1 gate capacitance, C_{f1} is the poly-1/field capacitance, C_{s1} is the poly1/source overlap capacitance, C_{23} is the poly-2/poly-3 coupling capacitance, C_{g2} is the poly-2 gate capacitance, C_{f2} is the poly-2/field capacitance, C_{s2} is the poly-2/source overlap capacitance, C_{sg2} is the poly-2/select-gate line capacitance and C_{rg2} is the poly-2/recall-gate line capacitance.

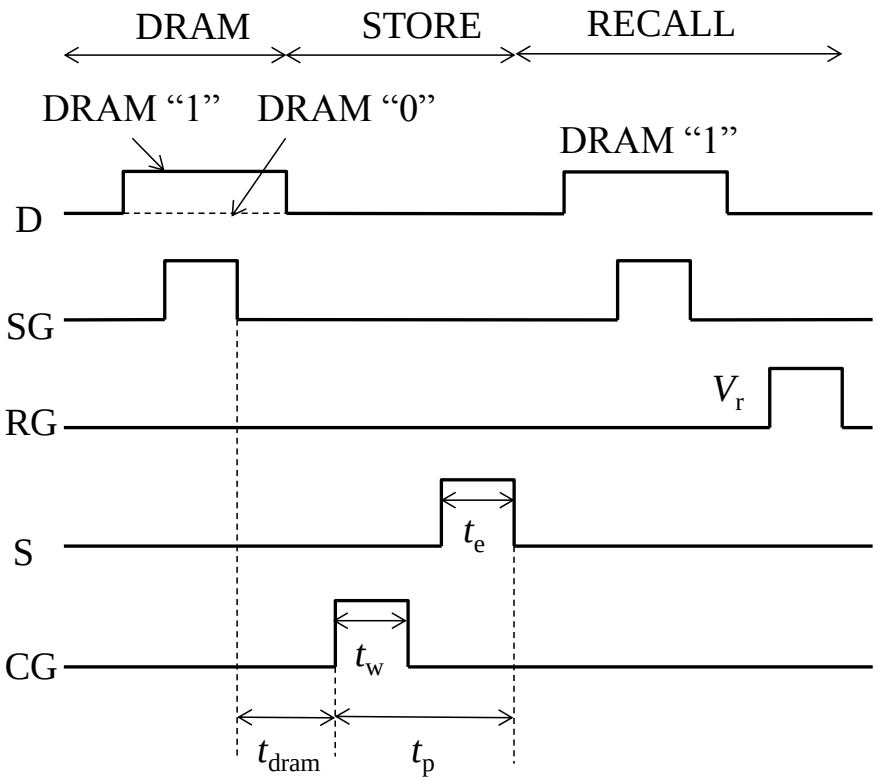


Fig. 4.4. Timing diagram of DRAM-mode, STORE-mode, and RECALL-mode operations for FG-MOSFET NV-DRAM [1.38].

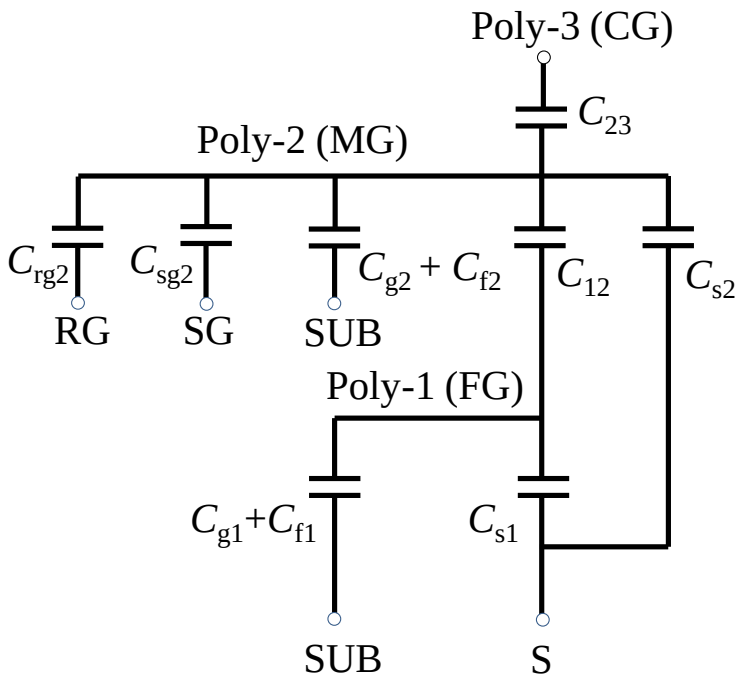


Fig. 4.5. Capacitive equivalent circuit of the FG-MOSFET NV-DRAM cell [1.38].

A. DRAM Mode

The stored charge (Q_{dram}) on the MG is given by

$$Q_{\text{dram}} = \left(C_{2t} - \frac{C_{12}{}^2}{C_{1t}}\right) V_{\text{dram}} - \frac{C_{12}}{C_{1t}} Q_{\text{fg}}, \tag{4.1}$$

where C_{1t} is the total capacitance on the poly-1 FG, and C_{2t} is the total capacitance on the poly-2 MG. V_{dram} is the drain potential in DRAM write mode, which is V_{cc} for DRAM data state of "1" and 0 for DRAM data state of "0", respectively. C_{1t} and C_{2t} are given by

$C_{1t} = C_{12} + C_{g1} + C_{s1} + C_{f1,}$

$C_{2t} = C_{12} + C_{23} + C_{g2} + C_{s2} + C_{f2} + C_{rg2} + C_{sg2.}$

B. Store Mode

The FG-MOSFET is written or erased by tunneling current through the TO. The tunneling current density (J_{tun}) is approximated by the well-known Fowler-Nordheim (FN) equation.

$$J_{\text{tun}} = A E_{\text{ox}}{}^2 \exp\left(-\frac{B}{E_{\text{ox}}}\right), \tag{4.2}$$

where E_{ox} is the electric field in the TO, and A and B are constants. The tunneling parameters A and B are calculated from the tunnel I-V characteristics.

$$A = 1.51 \times 10^{-6}\ \mathrm{A/V^2} \text{ and } B = 2.38 \times 10^{8}\ \mathrm{V/cm}, \tag{4.3}$$

for the write operation, and

$$A = 1.08 \times 10^{-4}\ \mathrm{A/V^2} \text{ and } B = 3.31 \times 10^{8}\ \mathrm{V/cm}, \tag{4.4}$$

for the erase operation. The TO field E_{ox} is given by

$$E_{ox} = \frac{|V_{ox}|}{T_{ox}}, \tag{4.5}$$

where V_{ox} is the voltage drop across the TO and T_{ox} is its thickness.

In a write operation (Fig. 4.6(a)), a programming pulse V_p is applied to the CG, grounding the S, SG and RG. The V_{ox} is enhanced by the positive charge on the MG and reduced by the initial negative charge on the FG according to

$$|V_{ox}|_{write} = K_w V_p + R_{cc}\left(V_{dram} + \frac{Q_{fg}}{C_{12}}\right). \tag{4.6}$$

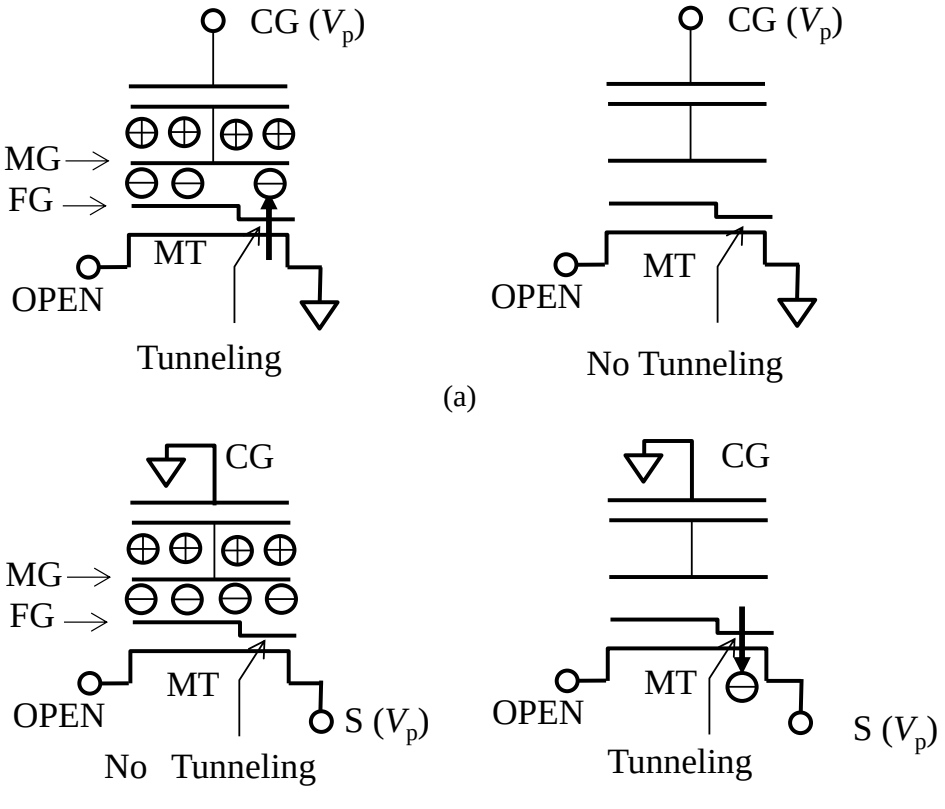


Fig. 4.6. Store operation concept: (a) Write cycle for erased cell, (b) Erase cycle for written cell [1.38].

In an erase operation (Fig. 4.6(b)), a programming pulse V_p is applied to the S, grounding the CG, SG and RG. The V_{ox} is enhanced by the initial negative charge on the FG and reduced by the positive charge on the MG according to

$$|V_{ox}|_{erase} = K_w V_p - R_{cc}\left(V_{dram} + \frac{Q_{fg}}{C_{12}}\right), \quad (4.7)$$

where

$$K_w = \frac{C_{12}\, C_{23}}{C_{1t}\, C_{2t} - C_{12}^{\,2}}$$

$$K_e = 1 - \frac{C_{2t}\, C_{s1} + C_{12}\, C_{s2}}{C_{1t}\, C_{2t} - C_{12}^{\,2}}$$

$$R_{cc} = \frac{C_{12}}{C_{1t}},$$

K_w and K_e denote the fraction of the programming voltage V_p that appears across the TO. R_{cc} is the poly-1/poly-2 capacitive coupling ratio.

The rate of charge injection into the FG due to the tunneling current is given by

$$\frac{\Delta Q_{fg}}{\Delta t} = A_{tun}\, |J_{tun}|, \quad (4.8)$$

where t is the time, A_{tun} is the TO area. Stored charge on the FG shifts the FG-MOSFET V_{th} according to the relation.

$$V_{th} - V_{thi} = -\frac{Q_{fg}}{C_{12}}, \quad (4.9)$$

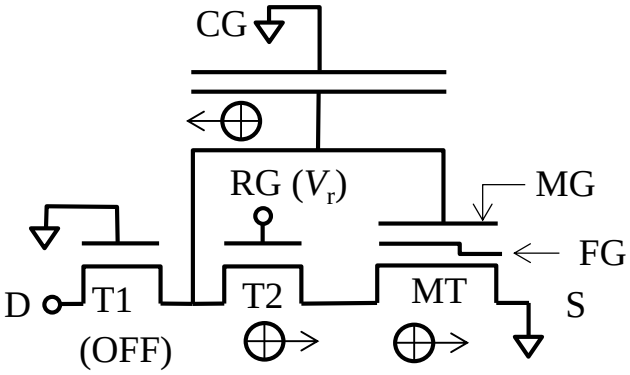
where V_{thi} (1.5V) is the FG-MOSFET V_{th} for an electrically neutral FG. From the above-mentioned model, the FG-MOSFET is written or erased depending on whether the DRAM data state is “1” or “0”, respectively, after a store operation. When the data states in the DRAM and FG-MOSFET are equal prior to a store operation, then the FG-MOSFET is inhibited from being written or erased. In addition, the DRAM data is not disturbed even after a store operation, since the two transistors T1 and T2 are turned off during the store operation.

The FG-MOSFET NV-DRAM cell enables on-chip generation of high voltage V_p from an external single 5 V power supply because of the very small current dissipation during a store operation.

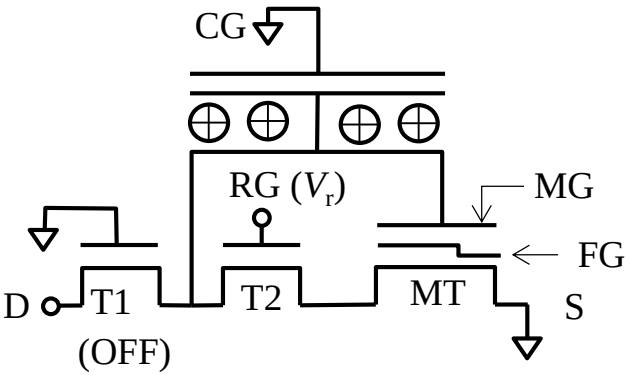
C) Recall Mode

The new recall operation concept is shown in Fig. 4.7. The recall operation is performed by applying a recall pulse V_r to the RG, grounding the SG, CG and S after DRAM write of data state "1". After a recall operation, the MG potential corresponds to the FG-MOSFET V_{th}. As a result, depending on whether the FG-MOSFET V_{th} state is

"high" or "low", the DRAM data state becomes "1" or "0", respectively. The dc current path from the drain is blocked by the T1.



(a) Low V_{th} State



(b) High V_{th} State

Fig. 4.7. Recall operation concept: (a) low V_{th} state, (b) high V_{th} state [1.38].

IV. STORE CHARACTERISTICS

The write and erase characteristics in a store operation are evaluated. In the write operation, as shown in Fig. 4.8, the FG-MOSFET is written to the high V_{th} state in the DRAM data state of "1", even though the FG-MOSFET is inhibited from being programmed in the DRAM data state of "0". On the other hand, in the erase operation, as shown in Fig. 4.9, the FG-MOSFET is erased to the low V_{th} state in the DRAM data state of "0", even though the FG-MOSFET is inhibited from being erased in the DRAM data state of "1". The measured data shows good agreement with the calculated data using the above-mentioned store model.

As a result, an acceptable V_{th} window of approximately 4 V is obtained with a V_p of 13.5 V as shown in Fig. 4.10. The state of the DRAM is stored in the FG-MOSFET in less than 10 ms with a V_p of 13.5 V as shown in Fig. 4.11. The V_{th} window after a store operation is determined by the total amount of stored charge on the MG during the store operation. Figure 4.12 shows the V_{th} window after a store operation as a function of DRAM data storage time (t_{dram}) with storage temperature as a parameter. No V_{th} shift is observed after t_{dram} of 100 ms even at 85 °C, which is sufficient for data transfer from DRAM to FG-MOSFET in the event of power failure.

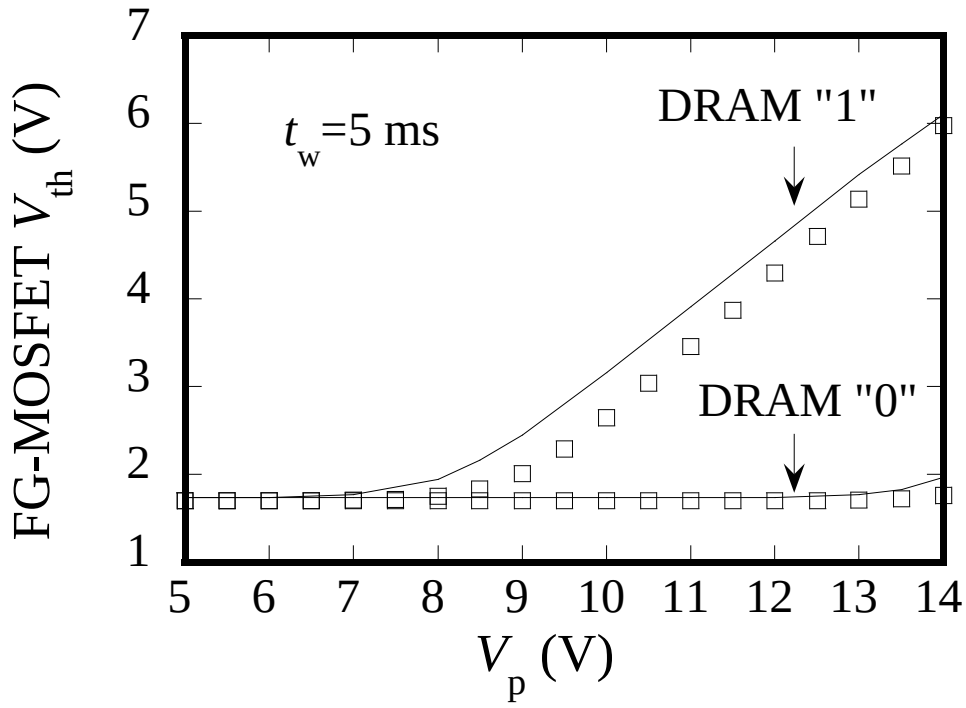


Fig. 4.8. FG-MOSFET V_{th} after a write operation as a function of program voltage V_p. The symbols represent the measured data and lines represent the simulation result [1.38].

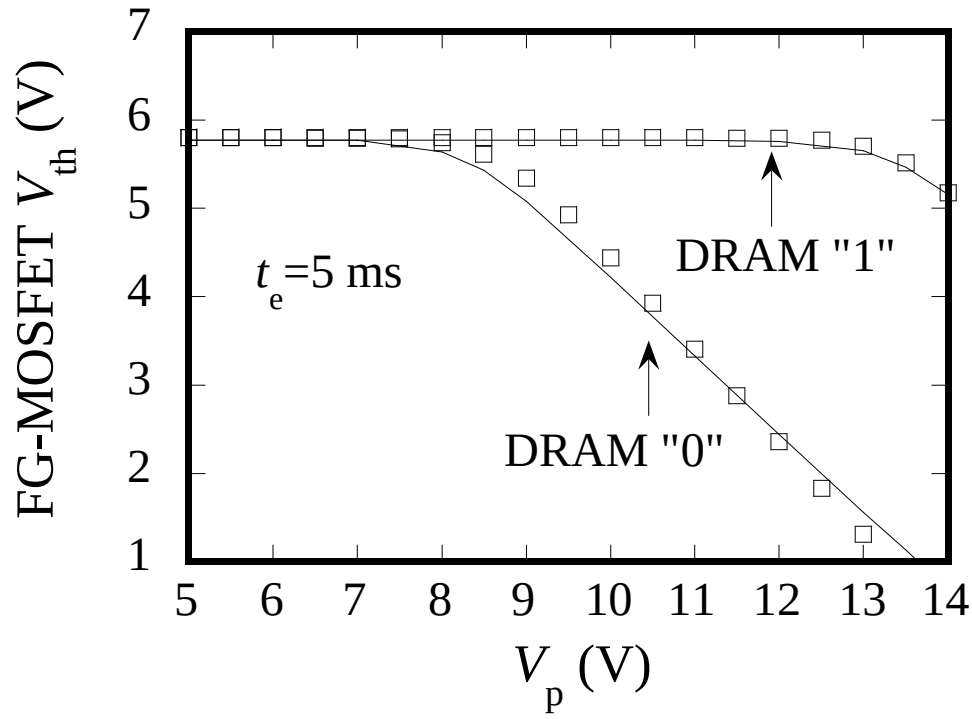


Fig. 4.9. FG-MOSFET V_{th} after an erase operation as a function of program voltage V_p. The symbols represent the measured data and lines represent the simulation result [1.38].

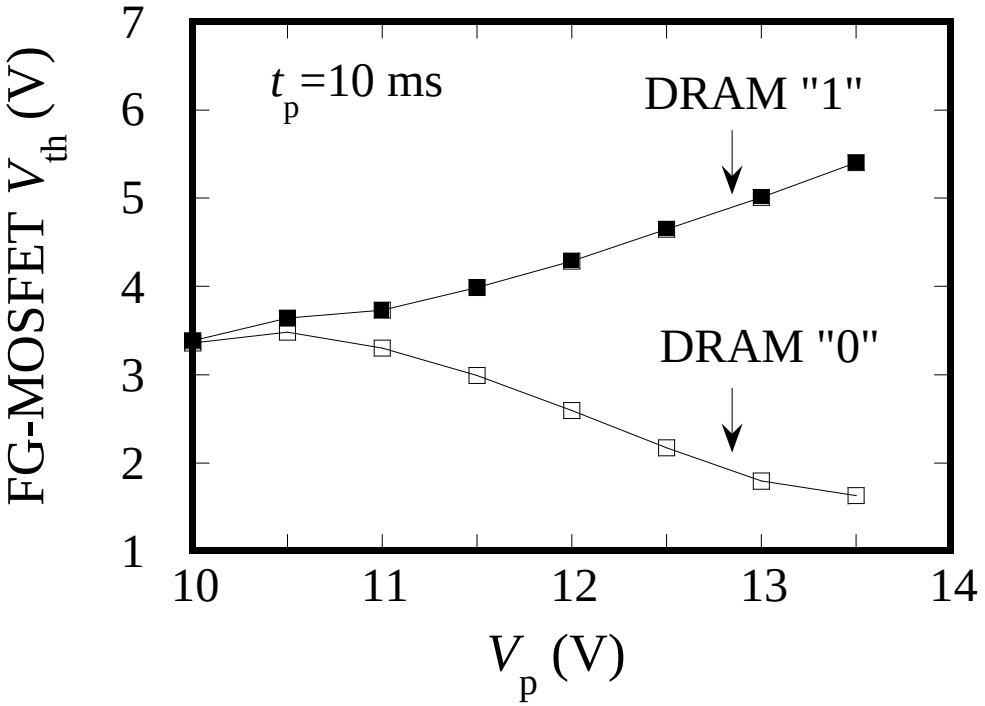


Fig. 4.10. FG-MOSFET V_{th} after a store operation as a function of program voltage (V_p). The symbols represent the measured data and lines represent the simulation result [1.38].

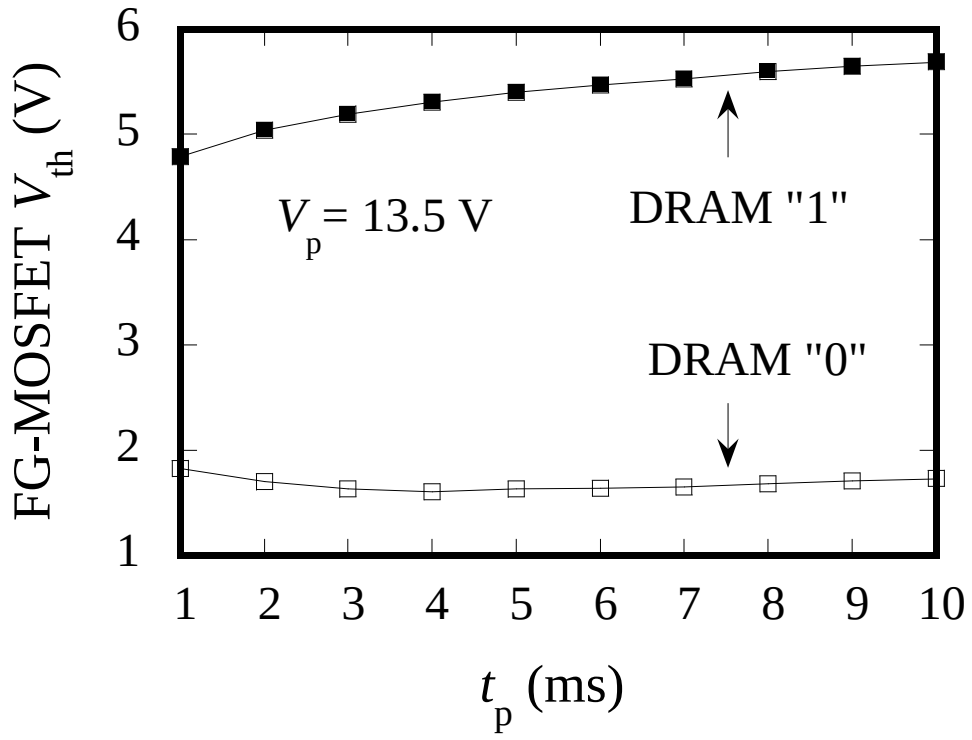


Fig. 4.11. FG-MOSFET Vth after a store operation as a function of program time (t_p). The symbols represent the measured data and lines represent the simulation result.

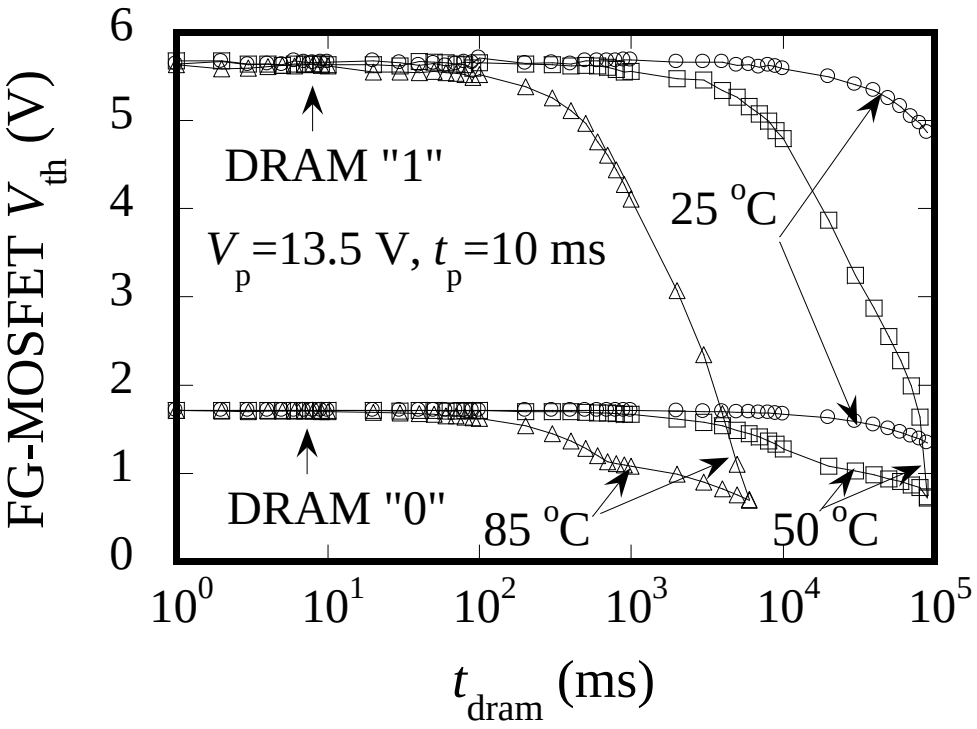


Fig. 4.12. FG-MOSFET V_{th} window after a store operation as a function of DRAM storage time (t_{dram}) [1.38].

V. RELIABILITY OF FG-MOSFET

Evaluation of store endurance, data retention and soft-write immunity are evaluated on single cell.

The intrinsic store endurance is demonstrated to be greater than 10^7 cycles at 85°C for a single cell, as shown in Fig. 4.13. Even after 10^6 store cycles at 85 °C, the closure of the V_{th} window of FG-MOSFET is less than 0.5 V.

High-temperature bake is performed after 10^5 store cycles for written cells with negative charge on the FG. Figure 4.14 shows the mean value of data retention time (t_{ret}) for an FG-MOSFET V_{th} decrease of 0.5 V as a function of storage temperature ($T_{storage}$) for written cells. The t_{ret}, which has an activation energy of 1.47 eV, is predicted to be greater than 10 years at 150 °C.

In DRAM write operation the positive drain voltage (V_{dram}) applied to the MG creates an electric field across the TO, which causes the V_{th} increase of an erased cell. Figure 4.15 shows the soft-write lifetime for an FG-MOSFET V_{th} increase of 50 mV as a function of the drain stress voltage. The soft-write lifetime is estimated to exceed 100 years under the DRAM write operations with the maximum V_{dram} = 5.5 V at 85 °C.

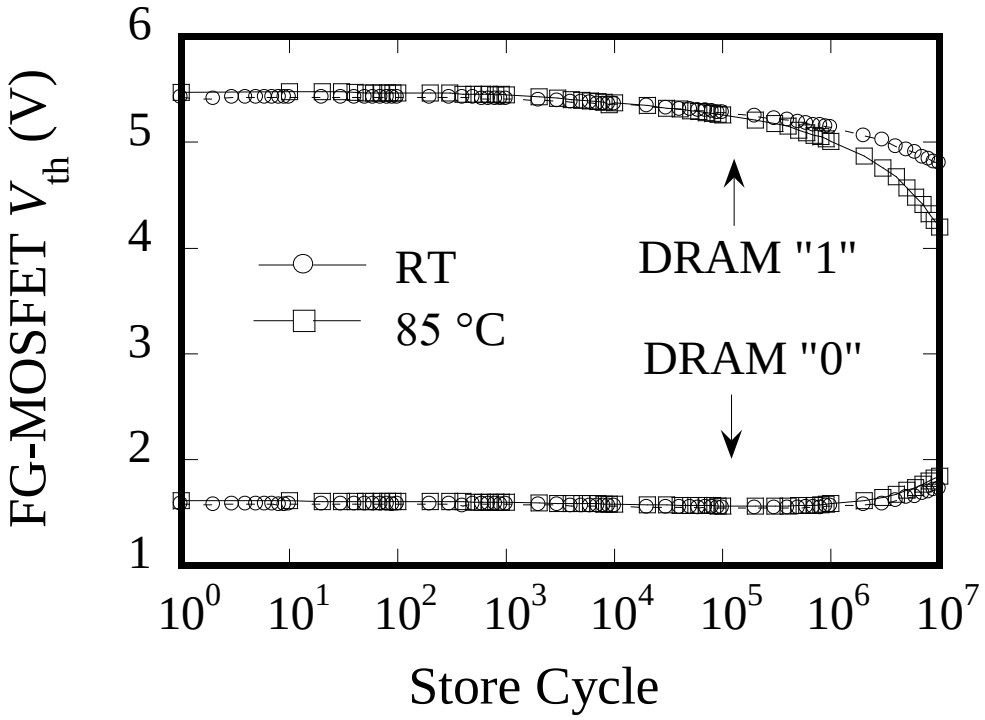


Fig. 4.13. Store endurance for a single cell [1.38].

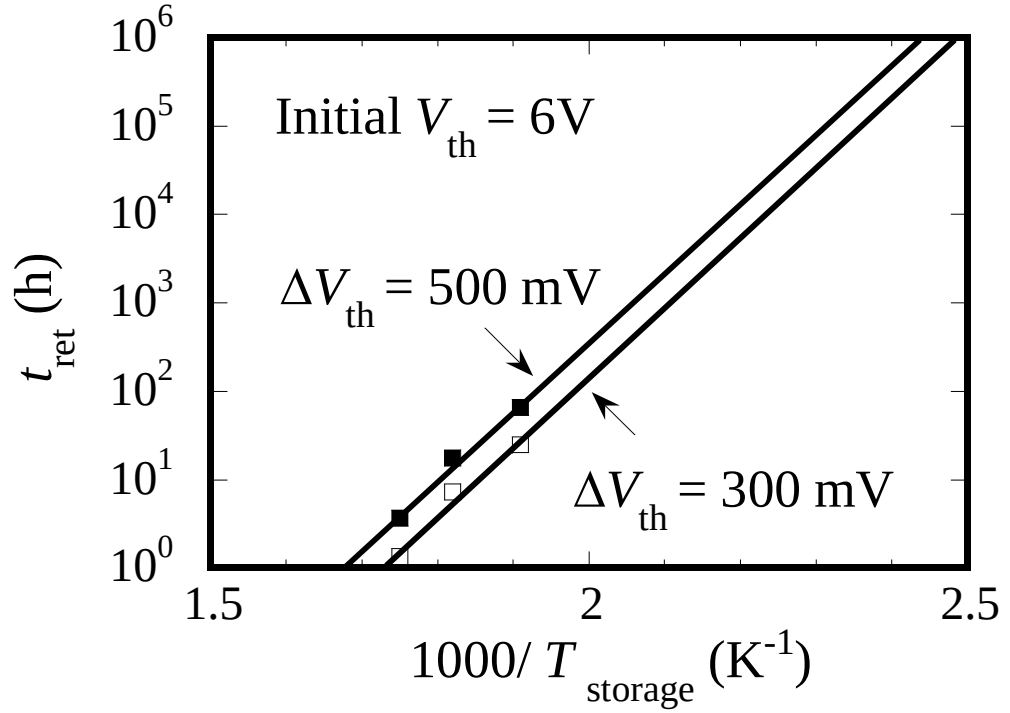


Fig. 4.14. FG-MOSFET data retention time (t_{ret}) for a V_{th} decrease of 300 and 500 mV as a function of storage temperature ($T_{storage}$) for written cells [1.38].

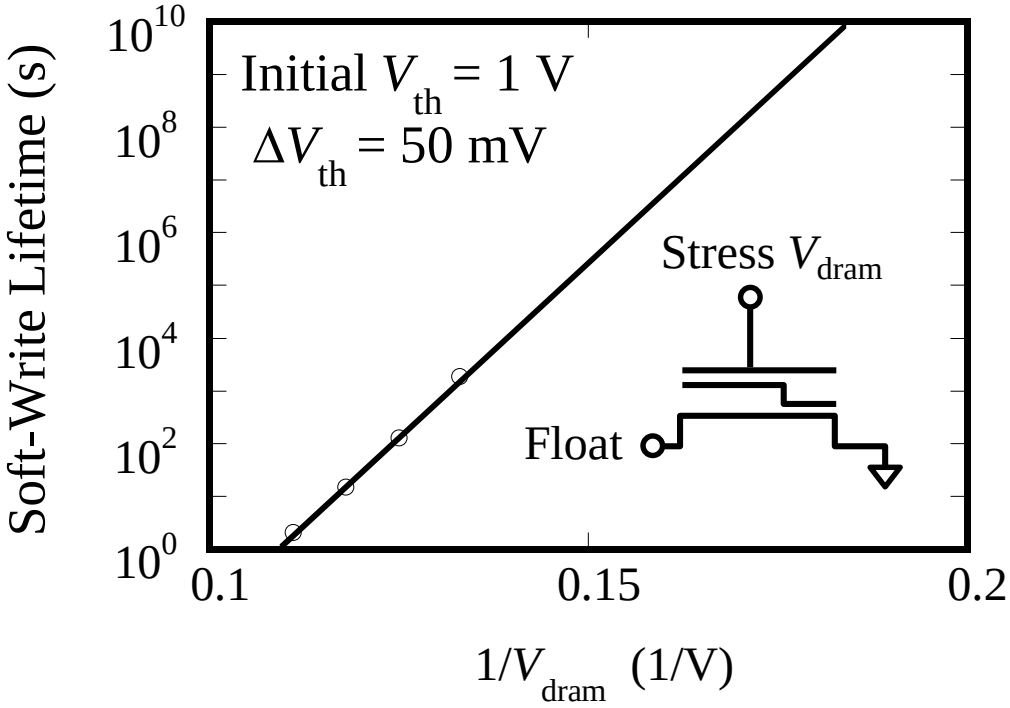


Fig. 4.15. Soft-write lifetime for a V_{th} increase of 50 mV as a function of the drain stress voltage (V_{dram}) for an erased cell [1.38].

VI. SCALABLE MEMORY CELL

A scalable NV-DRAM cell is proposed for further reduction of cell size as shown in Fig. 4.16.

A. Cell Structure

Figure 4.16(a) shows a cross-sectional view of the scalable NV-DRAM cell. The cell consists of two transistors (T1 and MT) and a storage capacitor. The MT is merged in serial between the switching transistor (T1) and the DRAM storage node (N1). T1 is the select transistor and its drain and gate are connected to the bit-line and the word-line, respectively. The gate of T1 is formed by first and second level polysilicon gate electrodes (SG1 and SG2) that are connected through a contact every 32 cells. MT is the stacked gate memory transistor, which has the FG with TO on its channel. This scalable cell technology can eliminate the T2 and MC from the 3 transistor NV-DRAM (3T NV-DRAM) cell shown in Fig. 4.1, and does not need high voltage transistor. Furthermore, the stacked gates (FG and MG) are formed in a self-aligned manner. Therefore, the scalable two transistor NV-DRAM (2T NV-DRAM) cell is expected to be reduced to less than half compared with 3T NV-DRAM cell.

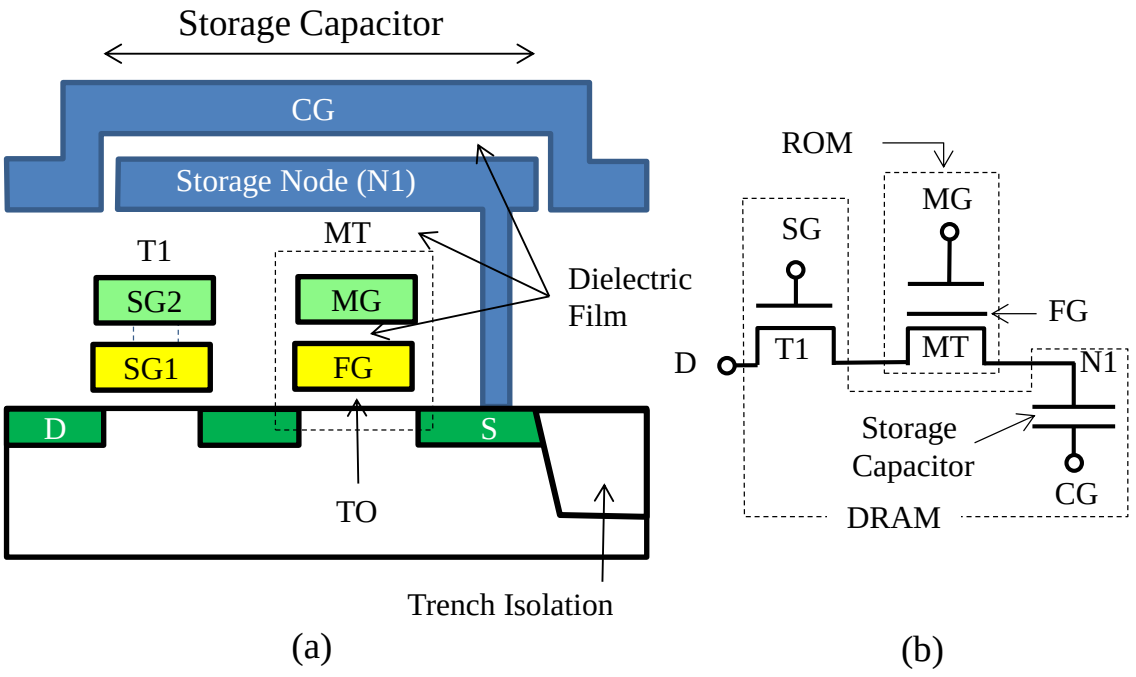


Fig. 4.16. Scalable NV-DRAM cell structure: (a) schematic cross-sectional view, and (b) equivalent circuit.

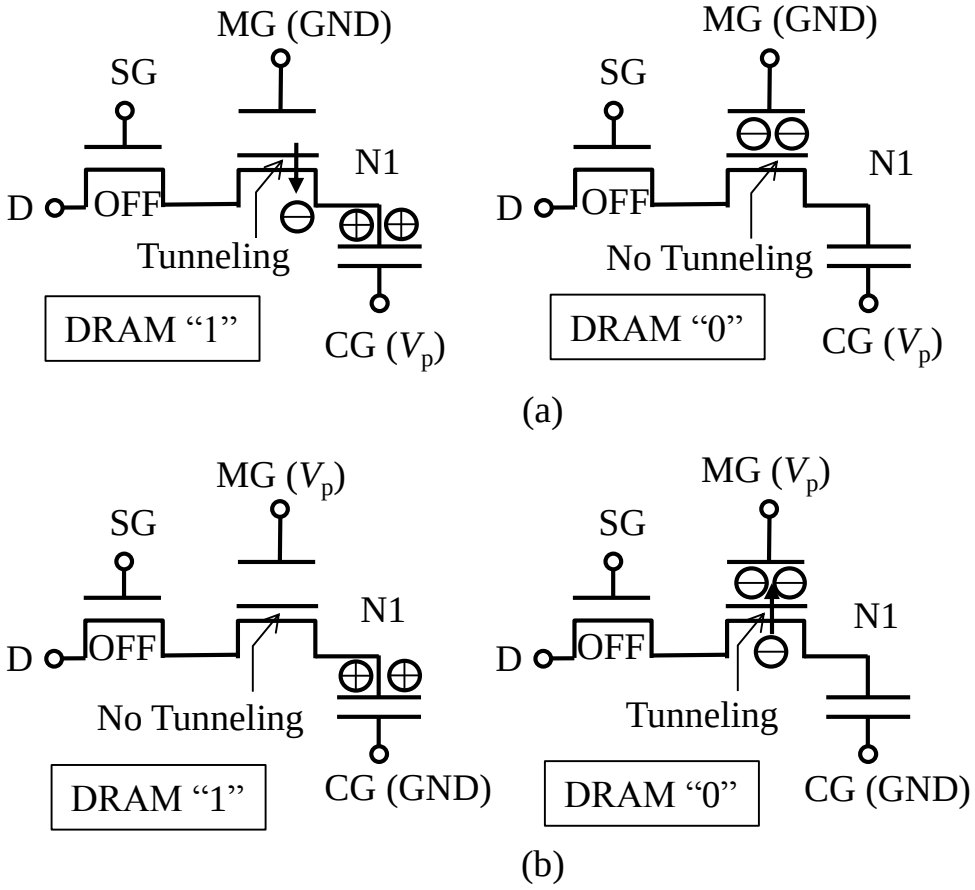


Fig. 4.17. Store operation concept: (a) write cycle for erased cell (negative charge in the FG), and (b) erase cycle for written cell.

B. Operation Concept

1) DRAM Mode

The equivalent circuit of scalable NV-DRAM cell corresponds to the standard DRAM cell by applying a high voltage to MG. The DRAM storage node (N1) is charged or discharged, depending on the DRAM data state "1" or "0".

2) Store Mode

The store of the scalable NV-DRAM cell is achieved by FN tunneling and is similar to that of the 3T NV-DRAM cell as shown in Fig. 4.17. In a program operation for an erased cell, a programming pulse V_p is applied to the CG, grounding the SG and MG. Because the V_{ox} is enhanced by the positive charge on the N1, the FG-MOSFET is programmed to the low-V_{th} by ejection of electrons from FG to N1 in the DRAM state of "1", while it is inhibited from being programmed in the DRAM state of "0". On the

other hand, in an erase operation, a programming pulse V_p is applied to the MG, grounding the SG and CG. The V_{ox} is reduced by the positive charge on the N1. The FG-MOSFET V_{th} can be erased to high-V_{th} by the injection of electrons from the N1 to the FG in the DRAM state of "0", while it is inhibited from being erased in the DRAM state of "1". As a result, the state of FG-MOSFET becomes "low-V_{th}" or "high-V_{th}" depending on the DRAM state of "1" or "0".

VII. CONCLUSIONS

A 3T NV-DRAM cell combining a DRAM cell with an FG-MOSFET was developed, which stores all the data in its FG in the event of sudden power failure and recall the data to DRAM quickly at the time of power-on. The store operation is completed in less than 10 ms with small power consumption by FN tunneling of electrons across the gate oxide between the DRAM storage-node and the FG. The single cell shows excellent reliability such as store endurance greater than 10^7 cycles at 85 °C and data retention in excess of 10 years under high storage temperature of 150 °C. The FG-MOSFET NV-DRAM cell is successfully demonstrated on a 1 Mbit test chip. These technologies allow the CPU to store the data being processed in the event of a sudden power failure, thus overcoming the main drawbacks of silicon MOSFET-based volatile memories

Furthermore, a scalable 2T NV-DRAM cell, for which an FG-MOSFET is merged in serial between a switching transistor and a storage capacitor, is discussed for high-density working memory applications.

REFERENCES

[4.1] H. Park, S. Yoo, and S. Lee, "Power management of Hybrid DRAM/PRAM-based Main Memory," *Design Automation Conference*, 2011, pp. 59 – 64.

[4.2] K. Abe, H. Noguchi, E. Kitagawa, N. Shimomura, J. Ito, and S. Fujita, "Novel Hybrid DRAM/MRAM Design for Reducing Power of High Performance Mobile CPU," *International Electron Devices Meeting Technical Digest*, 2012, pp. 10.5.1 – 10.5.4.

[4.3] W. Wei, K. Namba, and F. Lombard, "Extending Non-Volatile Operation to DRAM Cells," IEEE Access, vol. 1, pp. 758 – 769, 2013

[4.4] W. Tian, Y. Zhao, L. Shi, Q. Li, J. Li, C.J. Xue, M. Li, E. Chen, "Task Allocation on Nonvolatile-Memory-Based Hybrid Main Memory," *IEEE*

Transaction on Very large Scale Integration Systems, Vol. 21, no. 7, pp. 1271 – 1284, Jul. 2013.

[4.5] Y. Terada, K. Kobayashi, T. Nakayama, H. Arima, and T. Yoshihara, "A New Architecture for the NVRAM – An EEPROM Backed-Up Dynamic RAM," *IEEE Journal of Solid-state Circuits*, vol. 23, no. 1, Feb. 1988.

[4.6] Y. Yamauchi, K. Tanaka, K. Sakiyama and R. Miyake, "A Novel NVRAM Cell Technology for High Density Applications," *International Electron Devices Meeting Technical Digest*, 1988, pp. 416 − 419.

[4.7] Y. Yamauchi, H. Ishihara, K. Tanaka, K. Sakiyama, and R. Miyake, "A Versatile Stacked Storage Capacitor on FLOTOX Cell for Megabit NVRAMs," *International Electron Devices Meeting Technical Digest*, 1989, pp. 595 - 598.

[4.8] Y. Yamauchi, H. Tateoka, K. Tanaka, K. Sakiyama, and R. Miyake, "A 4M Bit NVRAM Technology using a Novel Stacked Capacitor on Selectivity Self-Aligned FLOTOX Cell Structure," *International Electron Devices Meeting Technical Digest*, 1990, pp. 931 – 933.

[4.9] J. H. Ahn, S. H. Hong1, S. J. Kim, J. B. Ko, S. W. Shin1, S. D. Lee, Y. W. Kim,, K-S. Lee1, S-K. Lee1, S-E. Jang1, J-H. Choi, S-Y. Kim, G-H. Bae1, S-W. Park, and Y-J. Park, "An Experimental 256 Mbit Non-Volatile DRAM with Cell Plate Boosted Programming Technique," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2004, 42 – 43.

[4.10] K. H. Park, C. M. Park, and S. H. Kong, "Novel Double-Gate 1T-DRAM Cell Using Nonvolatile Memory Functionality for High-Performance and Highly Scalable Embedded DRAMs," *IEEE Transaction on Electron Devices*, vol. 57, no. 3, pp. 614 – 619, Mar. 2010.

Chapter 5

Dual-Mode Liquid-Crystal-Display with Voltage-Loss-Compensation Memory in Pixel

---

I. Introduction 93
II. Pixel Structure 95
III. Refreshing Scheme 96
A. Single-level Sense 96
B. Multi-level Sense 98
IV. Polarity Inversion Scheme 100
V. Dual-Mode LCD Panel 101
A. Architecture 101
B. Performance and Power Consumption 102
VI. Gray-Scale Levels 106
VII. Conclusions 107
References 108

I. INTRODUCTION

In mobile phones, information such as time, battery level and antenna sensitivity is often continuously displayed in black/white monochrome on a part of the screen area. Recently, with the remarkable progress in wireless broadband network infrastructure, strong demand for low power consumption and high image quality display in the whole screen area is growing even for still images. This is especially true in stand-by mode. Ultra-low-power and high-quality still image display in addition to the normal display on the same screen will be required for future mobile information devices such as

smartphones and e-books. However, each pixel, which consists of a switching thin film transistor (TFT) in series with a storage capacitor and stores its image information [1.39] as storage charge on the capacitor, requires refresh every 60 Hz even for still image display similarly to moving image display because of the TFT off-state leakage. Several groups have described the integration of a digital static memory and a liquid crystal AC driver into a standard pixel using a switching transistor and a storage capacitor [5.1] – [5.4]. This technology is a good approach for achieving low power consumption since the SRAM in the pixel can maintain the pixel voltage (V_{pix}) without refreshing. However, due to its complex structure, such SRAM-based pixel technology is not suitable for high-resolution displays with a high transmissive aperture ratio.

A self-refreshing pixel memory with inverter circuit technology has been reported to solve these issues in SRAM-based pixel technology [5.5] – [5.10], and can refresh all pixels simultaneously by inverting the polarity of V_{pix}. The driving frequency of the source line (SL) is reduced to $1/N_{gate}$ [N_{gate}: the number of gate lines (GLs)], which results in a remarkable reduction in the panel power consumption as shown in Fig. 5.1. In addition, the aperture ratio of panels using this technology is increased due to its simple pixel structure compared with a digital static pixel memory technology.

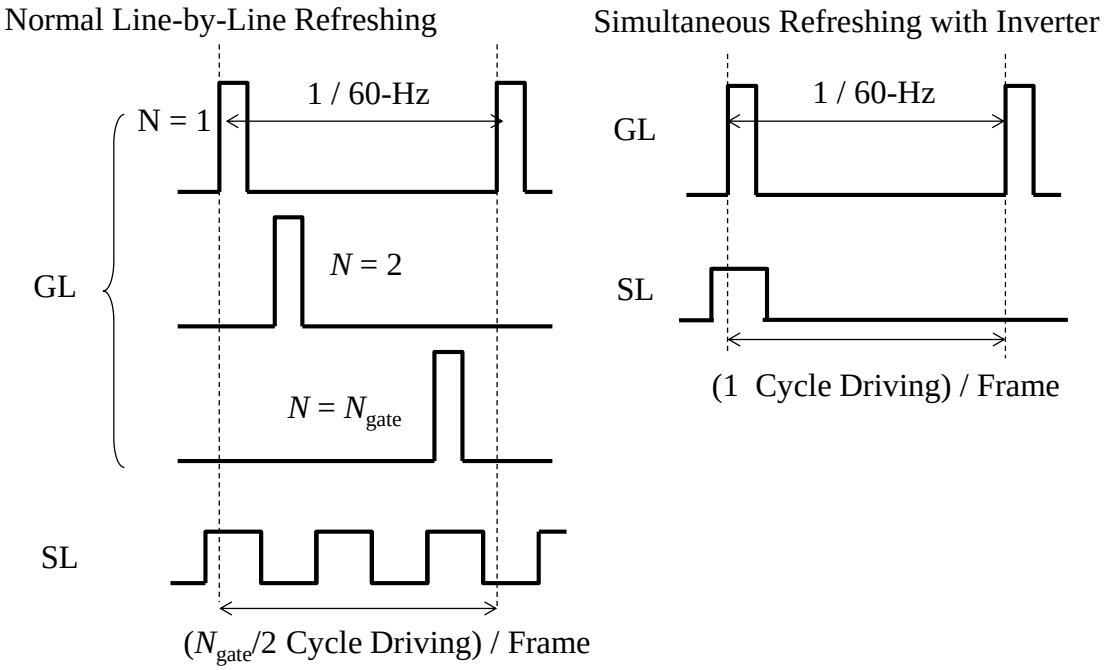


Fig. 5.1. Self-refreshing scheme with inverter circuit [1.39].

However, the above approaches require the use of area ratio gray-scale techniques to increase the color depth for high image quality. These techniques require the multiple sub-pixels for each RGB in a pixel, which results in a decrease of the aperture ratio [5.1] – [5.4].

Recently a multi-level pixel memory that uses an analog voltage gray-scale technique in order to increase the number of the gray-scale levels with high transmissive aperture ratio has been reported [5.5] – [5.10].

This chapter describes a novel pixel memory with a voltage loss compensation (VLC) circuit for an ultra-low power thin film transistor liquid crystal display (TFT-LCD), which enables an increase in color depth with a high transmissive aperture ratio by using the analog voltage gray-scale technique. The technology was successfully implemented into a 3.17 inch high transmissive aperture ratio (> 39%) HVGA transflective panel using a 1.5 μm process technology. The proposed panel has two display modes; a full-color normal mode and a low-power-consumption multi-color mode (memory-mode), both integrated into one display. The driving power consumption for the proposed transflective panel and the quality of the still image display are evaluated in both memory and normal modes.

II. PIXEL STRUCTURE

The proposed pixel structure with integrated VLC memory is shown in Fig. 5.2. The pixel has a memory select transistor (T2) and a VLC memory merged into a standard pixel. A standard pixel consists of a switching transistor (T1) and a storage capacitor with a capacitance of C_s = 200 fF. The gate and the source of T2 are connected to the boost line (BST line) and the source line (SL), respectively. The VLC memory consists of a memory transistor (MT), sensing transistor (T3) and one boost capacitor with a small capacitance of C_{bst}. The gate of MT is connected to the pixel electrode through T3 and is capacitively coupled to the BST line. The gate of T3 is connected to the reference voltage line (REF line). In memory mode, the memory select transistor is active and the SL is connected to the pixel electrode through the MT. In normal mode, the memory select transistor turns off and the source line is disconnected from the VLC memory. The multi-level VLC memory has a refresh diode (RD) in addition to the single-level VLC memory.

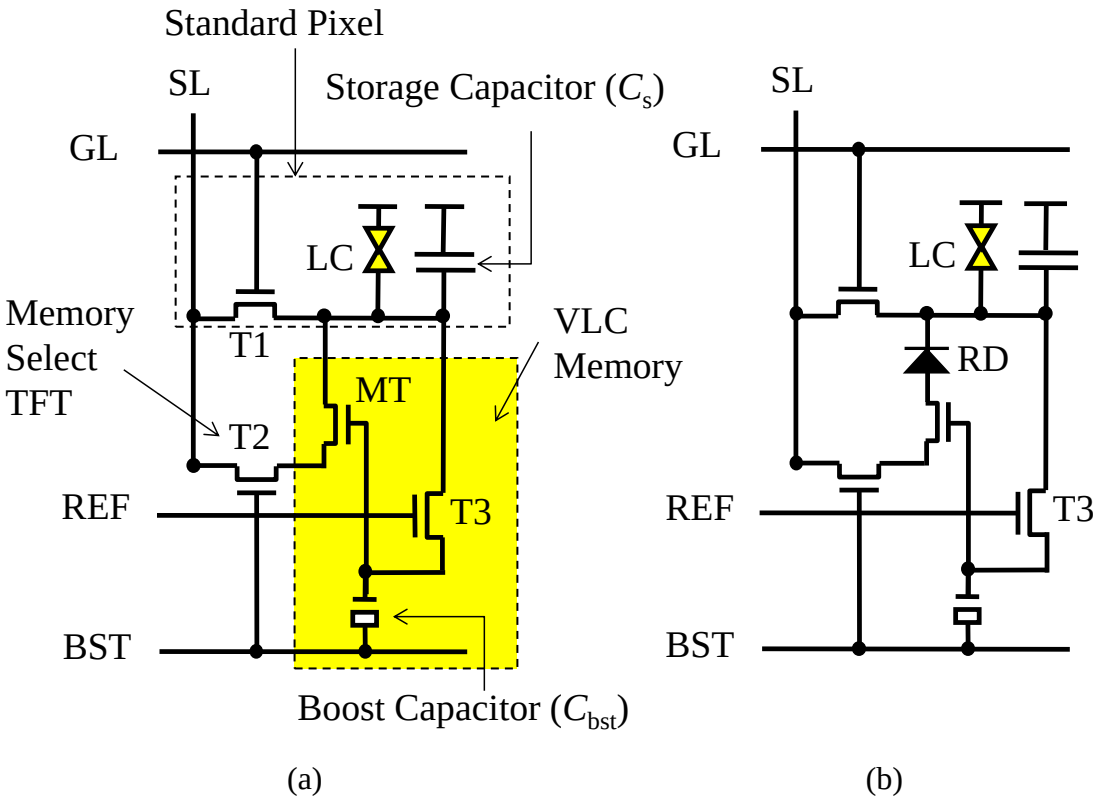
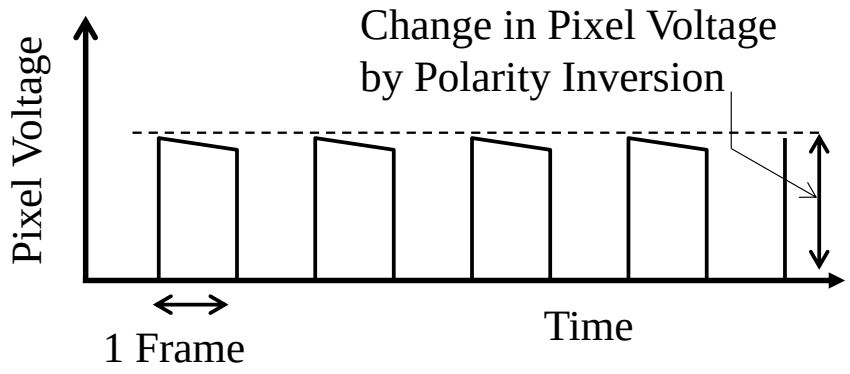


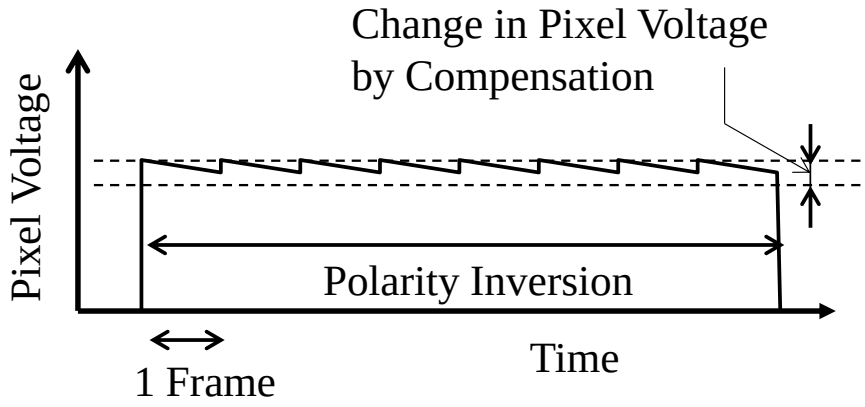
Fig. 5.2. LCD pixel structure with integrated VLC memory [1.35]. (a) Single-level pixel memory. (b) Multi-level pixel memory.

III. REFRESHING SCHEME

As shown in Fig. 5.3, the proposed VLC refreshing is achieved by sensing and compensating all the pixels simultaneously without the polarity inversion driving. The refresh rate is kept at 60Hz for flicker free operation, which also improves immunity against environmental disturbances such as temperature and illumination. The power consumed by charging and discharging the storage capacitor is significantly reduced by compensating the V_{pix} with a very small voltage loss (less than 50 mV). The proposed VLC refreshing scheme is suitable for low-power-consumption and high resolution TFT-LCD.



(a)



(b)

Fig. 5.3. Comparison between (a) the conventional refreshing scheme (both refresh and polarity inversion are performed simultaneously) and (b) the proposed VLC refreshing scheme [1.39].

A. Single-Level Sense

Figure 5.4 shows the sense scheme and timing chart for the single-level VLC memory. If $V_{pix} > V_{ref,eff} = V_{ref} - V_{th3}$, T3 turns off and the gate-node of MT (N2) is boosted by applying a high voltage to the BST line (V_{bst}). Here, V_{ref} is the voltage applied to the gate of the sensing transistor and V_{th3} is the threshold of the sensing transistor. As a result, the pixel electrode (N1) is connected to the SL and $V_{pix} = V_{sl,high}$ (= 5 V). On the other hand, if $V_{pix} < V_{ref,eff}$, T3 turns on and the N2 is not boosted by the high voltage on the BST line because of the very small capacitance coupling ratio of C_{bst} / C_s (< 0.05). As a result, the SL is electrically isolated from the N1 and V_{pix} approaches gradually to $V_{sl,low}$ (= 0) due to the leakage of T1 and T2 in long refresh intervals.

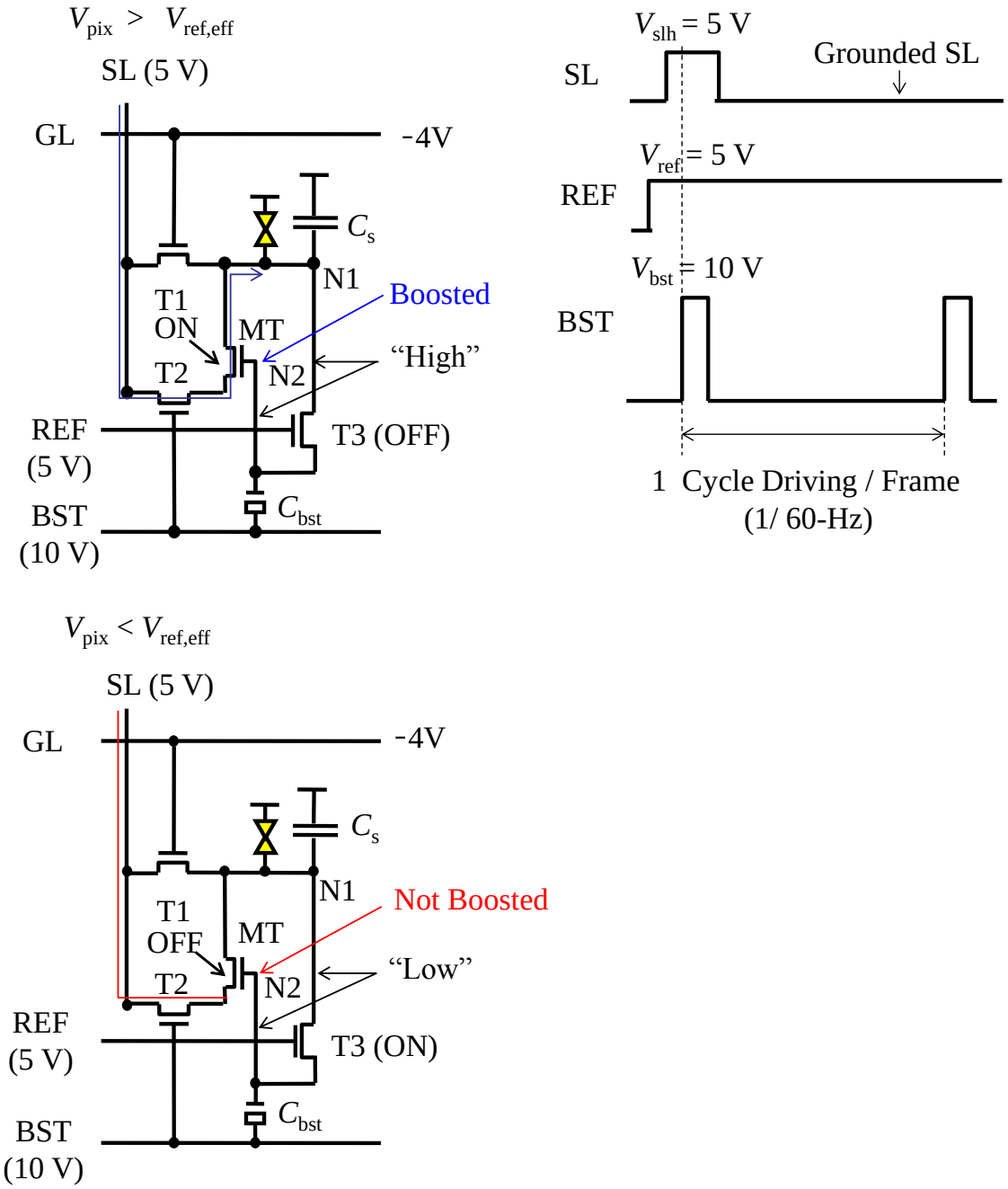


Fig. 5.4. Sense scheme and timing chart for the single-level VLC memory [1.39].

B. Multi-Level Sense

Figure 5.5 shows the sense scheme for the multi-level VLC memory. If $V_{pix} < V_{ref,eff}$ (lower-level), the N2 is not boosted and V_{pix} is maintained as described in single-level sense scheme. In addition, if $V_{pix} > V_{sl1f} = V_{s11} - V_f$ (higher-level), the RD is reverse biased and V_{pix} is maintained. Here, V_{sl1} and V_f are the source line voltage and a forward

threshold voltage for the RD, respectively. However, if V_{pix} is between $V_{ref,eff}$ and V_{sl1f} (refresh-level), the N2 is boosted and the RD is forward biased. As a result, the pixel is refreshed only when V_{pix} is refresh-level. Multi-level sensing self-refresh operations are performed sequentially by increasing the reference voltages (V_{ref1}, V_{ref2} and V_{ref3}) and refresh voltages (V_{sl1}, V_{sl2} and V_{sl3}), level by level in a stepping manner.

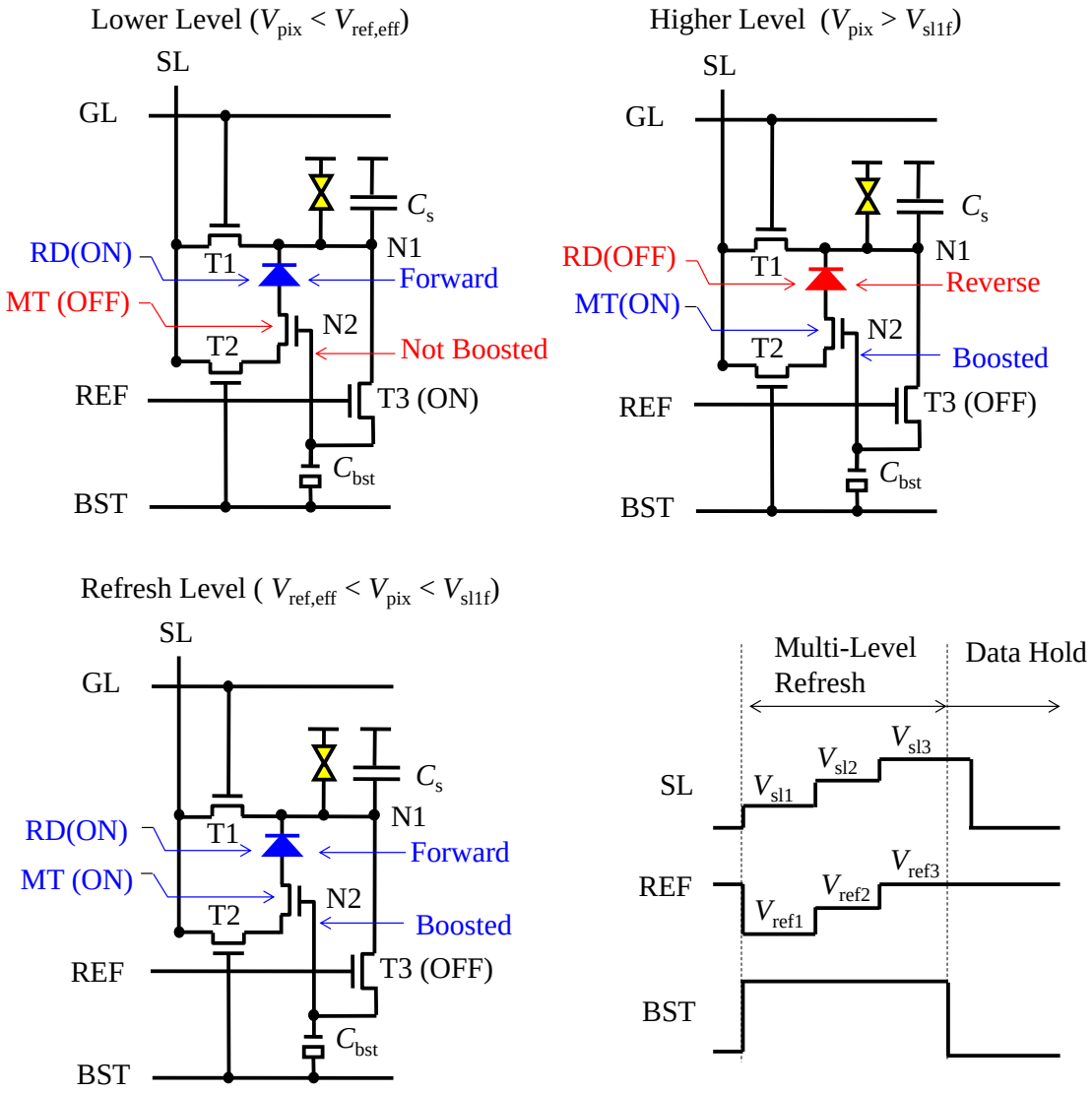


Fig. 5.5. Sense Scheme and timing chart for the multi-level VLC memory [1.39].

IV. POLARITY INVERSION SCHEME

Figure 5.6 shows the timing chart of the internal polarity inversion for the VLC memory. First, in the sampling period, the pixel voltage is sampled by applying V_{bst} = 10 V and then the sampling voltage is stored to the N2 by applying V_{ref} = 0. Next, in the reset period, all pixel electrodes are reset to $V_{sl,high}$ = 5 V by applying the high voltage to the GL. Then, in the inversion period, if the sampling voltage (V_{n2}) is higher than the threshold voltage of MT (V_{thm}), V_{pix} is set to $V_{sl,low}$ = 0. And, if V_{n2} is lower than V_{thm}, V_{pix} is maintained at 5 V.

The single-level sense refreshing is performed followed by the internal polarity inversion using the VLC memory. And, on the other hand, the multi-level sense refreshing is performed followed by the external inversion.

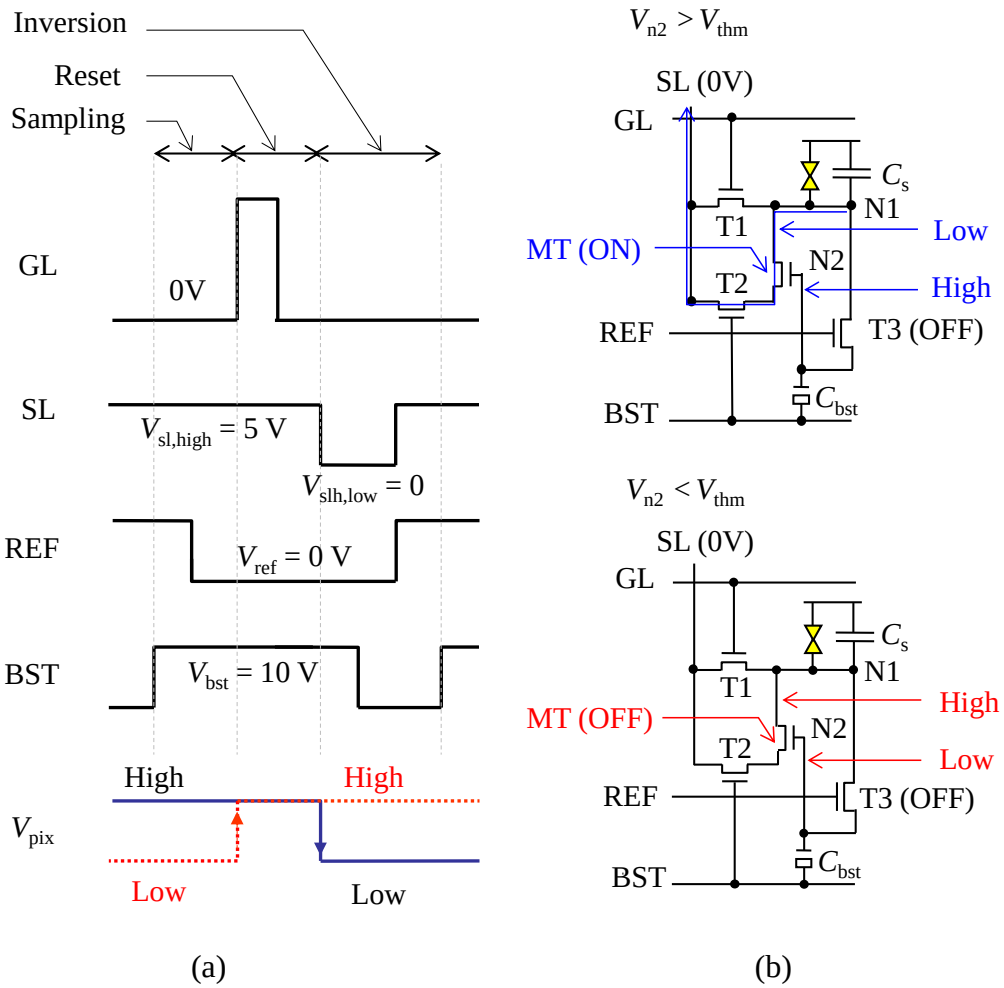


Fig. 5.6. (a) Timing chart of the internal polarity inversion for the single-level VLC memory [1.39]. (b) Internal polarity inversion scheme [1.39].

V. DUAL-MODE LCD PANEL

A. Architecture

The proposed dual-mode display panel has two display modes, “normal mode” and “memory mode” in a single display panel. In normal mode, the panel specification is identical to a standard transmissive display. In memory mode, the panel can display still images with ultra-low-power consumption. The schematic block diagram of the proposed panel is shown in Fig. 5.7. A memory select transistor and a VLC memory are incorporated into each sub-pixel. A memory control circuit is added to drive the two common lines, BST line and REF line in addition to the gate driver and source driver.

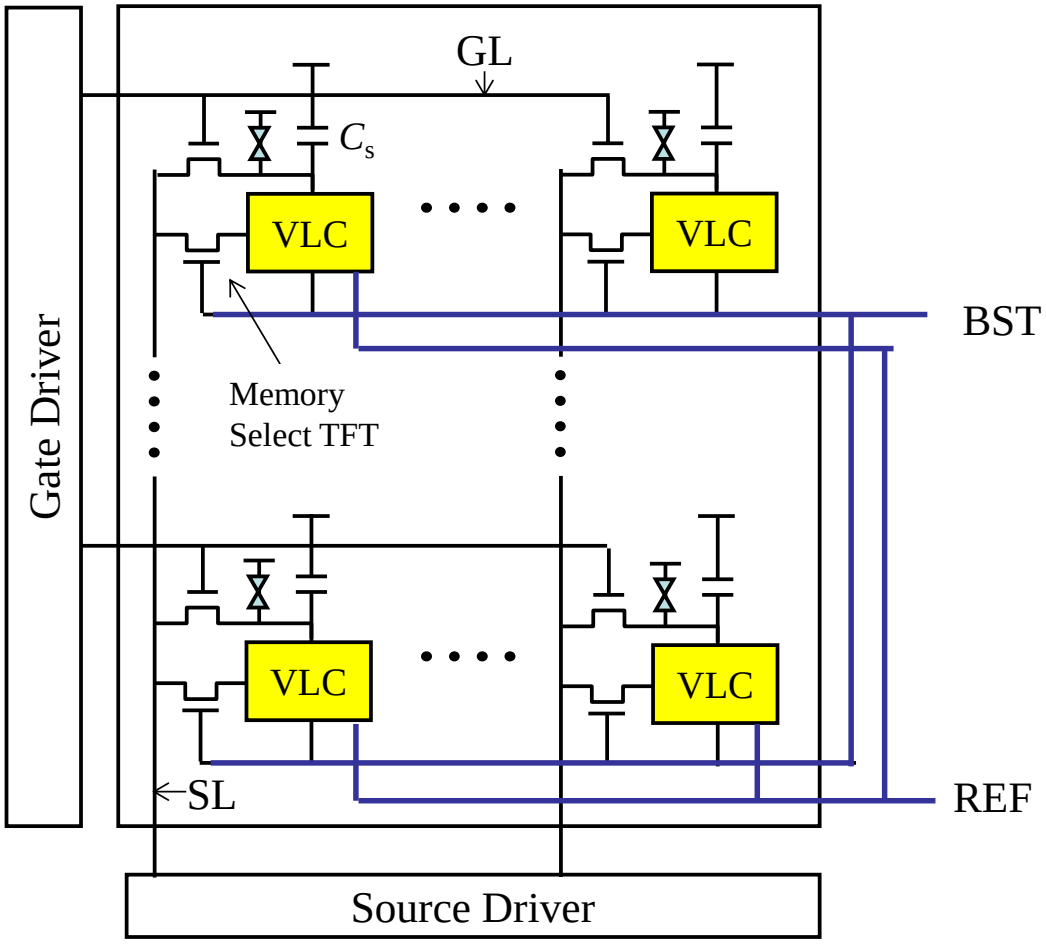


Fig. 5.7. The schematic block diagram of the proposed panel [1.39].

Table 5.1 shows the features of the advanced transflective panels with VLC memory in each pixel. A 3.17 inch HVGA panel (182 ppi pixel pitch) has been developed by using a standard double metal and low temperature polysilicon process based on 1.5 μm design rules. The sub-pixel circuit with a 4 gray-scale memory (64 colors) is formed under the small reflective electrode in the panel, which results in transmissive aperture ratio higher than 39 %.

B. Performance and Power Consumption

The driving power consumption was measured for different polarity inversion rates as shown in Fig. 5.8. The power consumption of VLC refreshing without polarity inversion is as low as 173 μW at a refresh rate of 60 Hz, and is almost independent of the number of gray-scale levels.

Table 5.1
Specifications of Advanced Dual-Mode LCD Panel [1.39].

Item	Specification
Liquid Crystal Mode	Transflective
Display Size	3.17 inch
Number of Sub-Pixels	320 × RGB × 480 (HVGA)
Pixel Pitch	139.5 μm (182 ppi)
Transmissive Aperture Ratio	40% (Single-Level Sense) 39% (Multi-Level Sense)
Process Technology	1.5 μm L/S & 2 metal layers
Number of Memory	1-bit Per Sub-Pixel
Number of Gray-Scale Levels	4 levels Each RGB (64 colors)

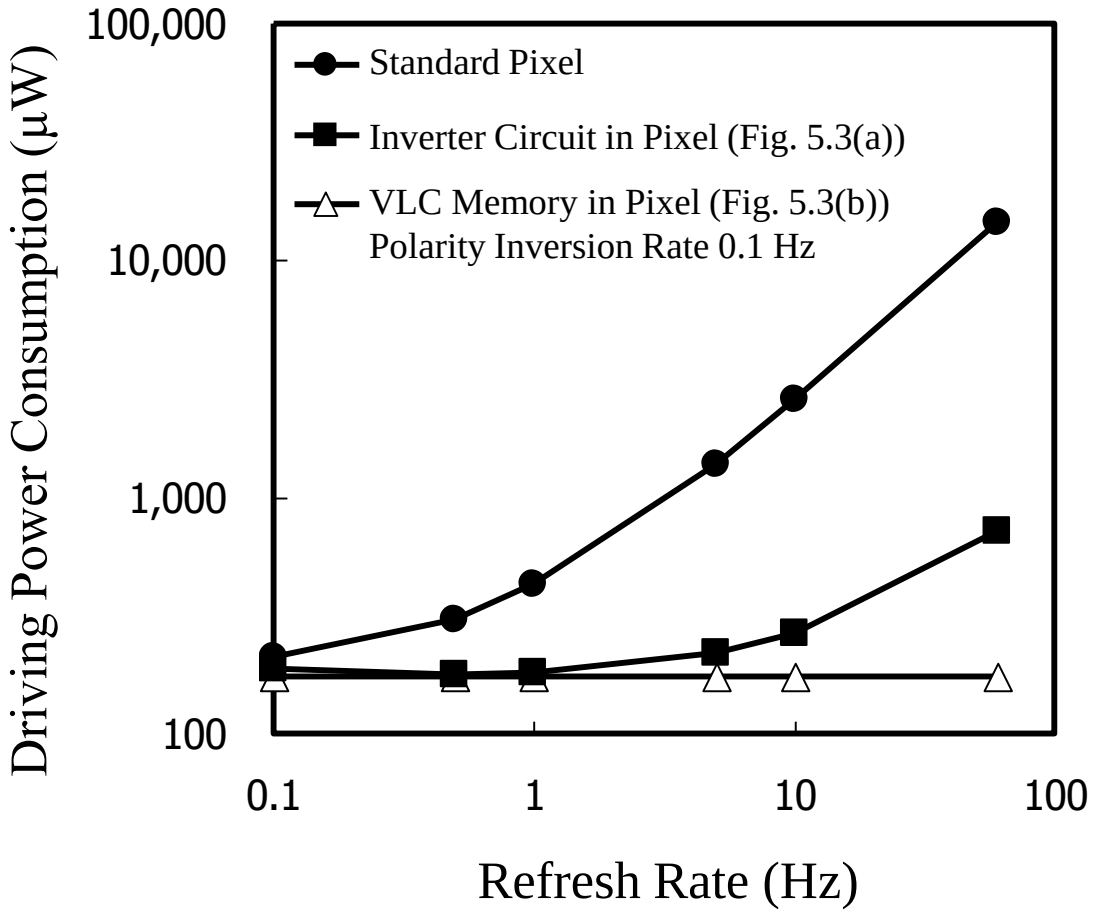


Fig. 5.8. Effect of refresh rate on the driving power consumption for a standard pixel, a pixel with integrated inverter circuit and a pixel with integrated VLC memory [1.39].

The total driving power consumption, including refresh and polarity inversion operations, is reduced in proportional to the polarity inversion rate and is mainly consumed by the refresh operation. With a polarity inversion rate of 0.1Hz, the total driving power consumption results in 188 μW and 210 μW for 2 gray-scales and 4 gray-scales images respectively. The effect of panel resolutions on the driving power consumption was also evaluated as shown in Fig. 5.9. The total driving power consumption for the VLC refreshing method shows little dependence on the pixel resolutions and is significantly reduced compared with the internal inversion method at 60Hz, since the VLC refreshing method can save the power needed for charging and discharging the pixel. The VLC refreshing method is thus suitable for low-power and high-resolution applications.

The multi-color and full-color still image displays are demonstrated for both memory-mode and normal-mode as shown in Fig. 5.10. VLC refreshing at high frequency (60 Hz) successfully suppresses flicker and maintains very good color quality throughout a long data retention period even with brightness of 250 cd/cm^2. The display image quality is improved by increasing the number of gray-scale levels from 2 to 4 and is further improved by using random dithering. This high quality, 64-color image is acceptable for whole-screen continuous still image display. In addition, the quality of the full-color image is as good as the standard transmissive LCD due to its high transmissive aperture ratio.

It is expected that the number of gray-scale levels can be further increased without sacrificing the transmissive aperture ratio for the multi-sense level VLC refreshing pixel memory by tightening the V_{th} distribution of the sensing transistor and optimizing LCD materials.

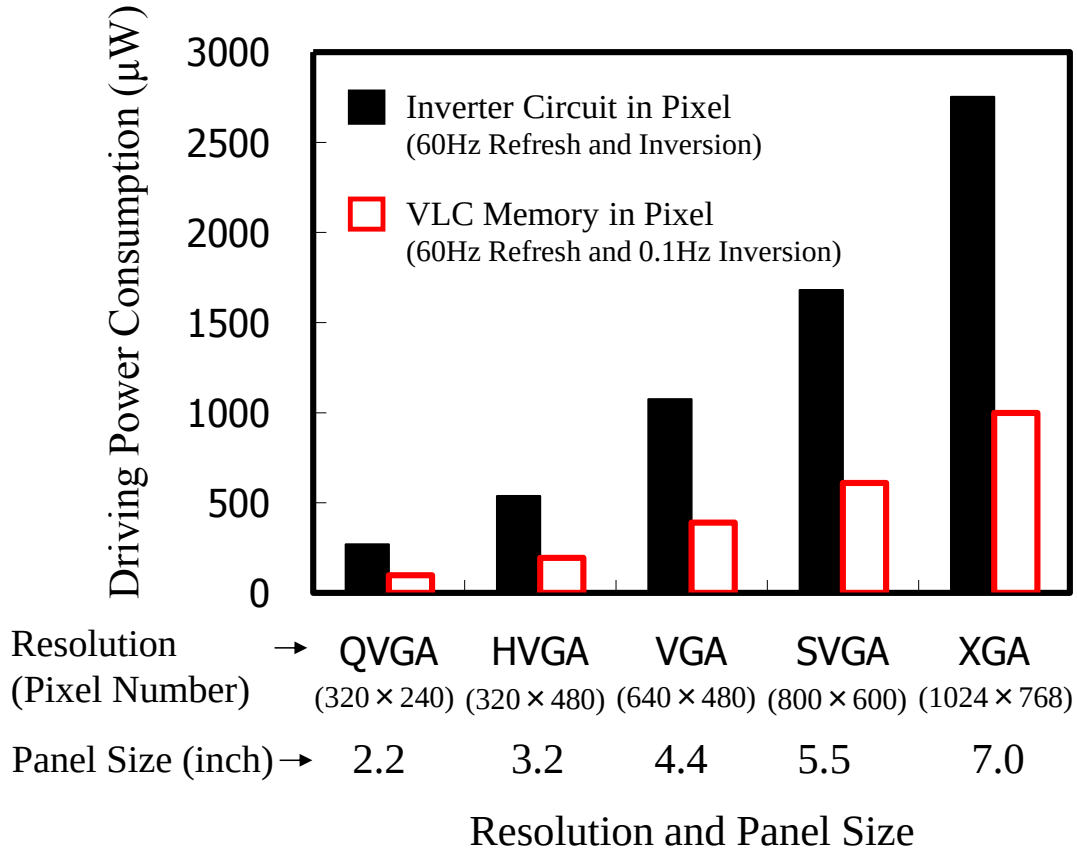
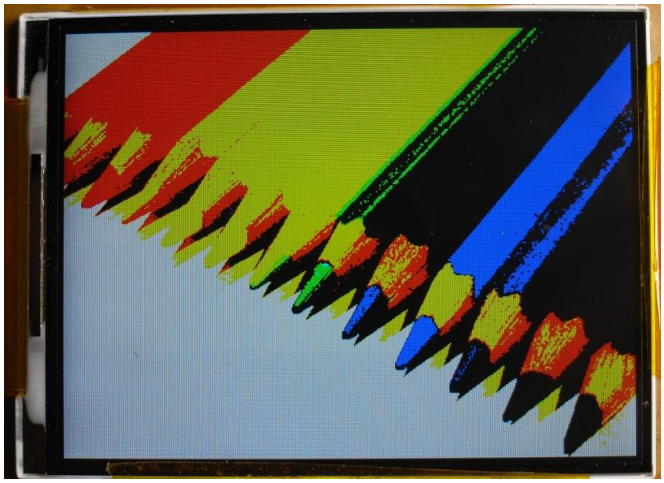


Fig. 5.9. The effect of panel resolution on driving power consumption [1.39].

(a)



(b)

(c)

Fig. 5.10. Still image view displays in memory-mode [(a) 8-color and (b) 64-color] and (c) normal-mode [1.39].

VI. GRAY-SCALE LEVELS

The effect of threshold voltage deviation on the maximum number of gray-scale levels was studied. Figure 5.11 shows the relation between the distribution of V_{pix} before and after refreshing and the distribution of sense-level voltage. The voltage width ($\Delta V_{gray\text{-}scale}$) between the neighboring gray-scale levels is given by

$$\Delta V_{gray\text{-}scale} = V_{pix,n} - V_{pix,n\text{-}1}, \quad (5.1)$$
$$= V_{loss,max} + \Delta V_{th3},$$

where $V_{pix,n}$ is the pixel voltage corresponding to the n-th gray-scale level. $V_{loss,max}$ is the maximum V_{pix} loss by the leakage through T1 and T2, and ΔV_{th3} is the threshold voltage deviation of T3. Here, $V_{loss,max}$ is given by

$$V_{loss,max} = \frac{I_{leak,max}\, t_{ret}}{C_s}, \quad (5.2)$$

where $I_{leak,max}$ is the maximum leakage through TFT, and t_{ret} is retention time.

Figure 5.12 shows the number of gray-scale levels ($N_{gray\text{-}scale}$) as a function of ΔV_{th3}. From these results, 8 levels multi-level-pixel (512 colors) is expected to be achieved with approximately ΔV_{th3} = 120 mV, which is achievable with current low temperature TFT process technologies.

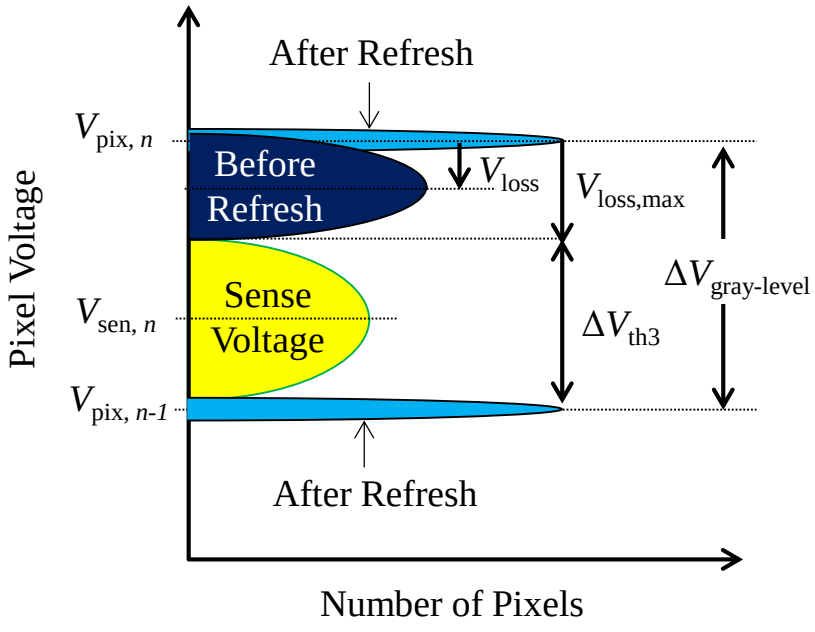


Fig. 5.11. Relation of the distribution of V_{pix} before and after refreshing and the distribution of sense-level voltages ($V_{sen,n}$) corresponding to n-th gray-scale level.

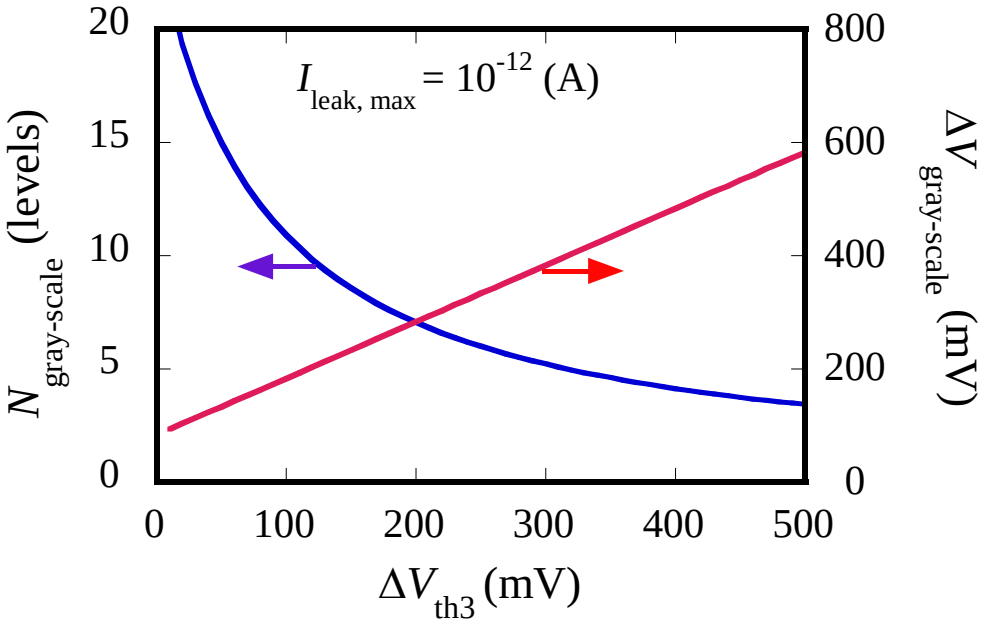


Fig. 5.12. $N_{gray\text{-}scale}$ and $\Delta V_{gray\text{-}scale}$ as a function of ΔV_{th3}.

VII. CONCLUSIONS

A novel VLC memory in pixel has been developed for an advanced dual-mode LCD, which can display a still image with ultra-low power consumption in addition to normal display. The number of gray-scale levels for a single subpixel can be increased using an analog voltage gray-scale technique.

The new pixel with a VLC memory is integrated under a small reflective electrode in a high transmissive aperture ratio (39%) 3.17 inch transflective HVGA TFT-LCD (320 × RGB × 480 subpixels) by using a standard low-temperature-polysilicon process based on 1.5 μm rules. No additional process steps are required. The VLC memory in each pixel enables simultaneous refresh with a very small change in pixel voltage, resulting in a two-orders-of-magnitude reduction in panel power for 64-color image (4 levels per 1 bit memory) display. The measured panel power, as small as 210 μW, is smaller more than one order of magnitude than the system power in standby mode (~ 5 mW).

The advanced transflective TFT-LCD with VLC memory can display high-quality multi-color images anytime and anywhere, due to its low-power-consumption and good outdoor readability. It also offers high transmittance for indoor applications. This technology will increase in importance as a display for future mobile information devices such as electronic books and personal navigation systems.

REFERENCES

[5.1] H. Kimura,T. Maeda, T. Tsunashima, T. Morita, H. Murata, S. Hirata, and H. Sato., "A 2.15 inch QCIF reflective TFT-LCD with Digital Memory on Glass (DMOG)", *SID International Symposium Digest of Technical Papers*, 2001, pp. 268 − 271.

[5.2] M. Senda, Y. Tsutsui, R. Yokoyama, and K. Yoneda, "Ultra-Low-Power Polysillicon AMLCD with Full Integration", *SID International Symposium Digest of Technical Papers*, 2002, pp. 790 − 793.

[5.3] Y. Nakajima, Y. Teranishi, Y. Kida, and Y. Maki, "Ultra-low-power LTPS TFT-LCD Technology Using a Multi-Bit Pixel Memory Circuit", *SID International Symposium Digest of Technical Papers*, 2006, pp. 1185 − 1188.

[5.4] K. Harada , H. Kimura , M. Miyatake , S. Kataoka , T. Tsunashima , T. Motai and T. Kaeamura "A Novel Low Power Consumption All Digital System on Glass Display with Serial Interface", *SID International Symposium Digest of Technical Papers*, 2009, pp.383 − 386.

[5.5] H. Tokioka, M. Agari, M. Inoue, T. Yamamoto, H. Murai, and H. Nagata, "Low Power Consumption TFT-LCD with Dynamic Memory Embedded in Pixels", *SID International Symposium Digest of Technical Papers*, 2001, pp. 280 − 283.

[5.6] K. Yamashita, K. Hashimoto, A. Iwatsu, and M. Yoshiga, "Dynamic Self-Refreshing Memory-in-Pixel Circuit for Low Power Standby Mode in Mobile LTPS TFT-LCD", *SID International Symposium Digest of Technical Papers*, pp. 1096 − 1099.

[5.7] L. W. Chu, P. T. Liu and M. D. Ker, "Design of Analog Pixel Memory for Low Power Application in TFT-LCDs", *Journal of Display Technology*, vol. 7, no. 2, pp. 62 – 69, Feb. 2011.

[5.8] N. Ueda, Y. Ogawa, K. Tanaka, K. Yamamoto, and Y. Yamauchi, "A Novel Multi-level in Pixel Technology for Ultra Low Power LTPS TFT LCD," *SID International Symposium Digest of Technical Papers*, 2010, pp. 615 − 618.

[5.9] Y. Yamauchi and N. Ueda, "Low Power Driving Technologies for High-Resolution AM-LCD", *Active-Matrix Flat-panel Display and Device*, 2010, pp. 85 – 88.

[5.10] H. Ohshima and M. Führen, "High-Performance LTPS Technologies for Advanced Mobile Display Applications", *SID International Symposium Digest of Technical Papers*, 2007, pp. 1482 − 1485.

CHAPTER 6

Floating-Gate MOSTFT Oxide Semiconductor Memory

I. Introduction 109
II. In-Ga-Zn Oxide MOSTFT 111
A. Device Structure 111
B. Basic Operation 112
III. Memory Cell Structure 112
IV. Principle of Cell Operation 114
V. Cell Performance 116
A. Simulation Results 116
B. Experimental Results 119
VI. Reliability 120
A. Prediction of V_{th} shift 120
B. Effect of V_{th} shift on Cell Performance 124
C. Effect of V_{th} shift on Power Dissipation 125
VII. Data Retention Characteristics 126
VIII. Conclusions 128
References 129

I. INTRODUCTION

Oxide semiconductors such as indium-gallium-zinc-oxide (IGZO) are expected to be used as channel layers of metal-oxide-semiconductor thin film transistors (MOSTFTs) in next-generation flat-panel displays owing to their good performance and uniformity [1.32], [6.1] – [6.3]. Several applications, such as active-matrix liquid-crystal displays, active-matrix organic light-emitting-diode displays, and electronic paper displays, have been demonstrated using IGZO MOSTFTs as driving and switching devices [6.4] – [6.7].

Some studies of the off-state leakage current characteristics have been reported for

IGZO MOSTFTs [6.8] – [6.12]. The band gap of IGZO (3.0 – 3.2 eV) is wider than that of Si (1.1 eV) [6.8], [6.9]. The off-state leakage current of IGZO MOSTFTs caused by thermally excited electrons is expected to be reduced by about 30 orders of magnitude in principle as compared with that of Si-based MOSTFTs because the off-state leakage is proportional to exp (-Eg / $k_B T$) under a strong inversion condition [6.9]. Here, E_g is the band gap and k_B is Boltzmann's constant. The ultra-low-leakage of 5×10^{-23} A/μm has been obtained at 85 °C for a MOSTFT with a crystalline IGZO semiconductor channel, which is larger than the expected value [6.10], [6.11]. The Fermi-level pinning effect due to the donor-type valence band deep states is explained in a previous paper as one of the reasons why the leakage value is larger than the expected value [6.12]. An IGZO MOSTFT with an ultra-low off-state leakage current has been successfully implemented in an ultra-low-power-consumption still-image display using a low-frequency driving scheme and a scan driver on glass [6.13], [6.14].

Recently, a study of a three-dimensional (3D) oxide semiconductor memory with an embedded IGZO MOSTFT on a Si chip as a nonvolatile switch device has been reported [6.11]. A memory transistor is fabricated on a bulk-Si substrate to read information. This 3D oxide semiconductor memory shows unlimited write cycles similarly to a DRAM and can store data without refreshing similarly to a nonvolatile memory. These characteristics are very attractive for its use as a next-generation low-power memory. The integration of the oxide semiconductor memory using an IGZO MOSTFT as a nonvolatile switch with pixel and peripheral circuits on a single panel is very attractive for system-on-panel (SoP) applications that can offer additional functionalities including design freedom and power saving. However, IGZO MOSTFT shows a larger threshold voltage shift by bias stressing than bulk-Si transistors due to its low-temperature process. As a result, the effects of threshold variation on the increase in power consumption and endurance degradation become serious issues.

This chapter describes a floating-gate MOSTFT oxide semiconductor (FLOTOS) memory that consists of wide-band-gap IGZO MOSTFTs and a storage capacitor. In Section II, the process technology is described. In Section III, the cell structure of the FLOTOS is described. In Section IV, the principle of operation is explained. In Section V, the cell performance is discussed using simulation results and verified experimentally. In Section VI, long-term reliability is discussed. In Section VII, the performance of the IGZO MOSTFT as a nonvolatile switch is evaluated. Finally, the conclusions of this work are presented in Section VIII.

II. IGZO MOSTFT

A. Device Structure

Figure 6.1 shows a schematic cross-sectional view of a bottom-gate-type IGZO MOSTFT. First, a stacked Ti/Al-Si/Ti film is deposited on a glass substrate and patterned to form the gate electrode. Next, a gate dielectric stack with an effective oxide thickness of 300 nm is deposited by plasma-enhanced chemical vapor deposition. Then, an IGZO layer (50 nm thick) is deposited using a conventional DC sputtering system with an Ar and O_2 gas mixture. The IGZO layers are patterned by a wet etching process. The stacked Ti/Al-Si/Ti film is directly deposited on the IGZO body and patterned to form the source and drain electrodes. The oxide dielectric (150 nm thick) was deposited on the metal interconnection as a passivation film. Finally, the fabricated MOSTFT was annealed at 350 ºC in ambient N_2.

The n-type conducting IGZO films are grown on the glass substrate using the above-mentioned processes. The IGZO film is a wide band-gap semiconductor that has a relatively high concentration of free electrons in its conduction band. These arise either from defects in the material or from extrinsic dopants, the impurity levels of which lie near the conduction band edge. A good ohmic contact can be obtained between the source/drain electrode and the IGZO channel region, because the electron affinity of IGZO is approximately 4.6 eV, which is higher than the work function of the source/drain metal material (~ 3.96 eV). The IGZO MOSTFT does not need source/drain junction required for Si MOSTFT, resulting in low temperature process.

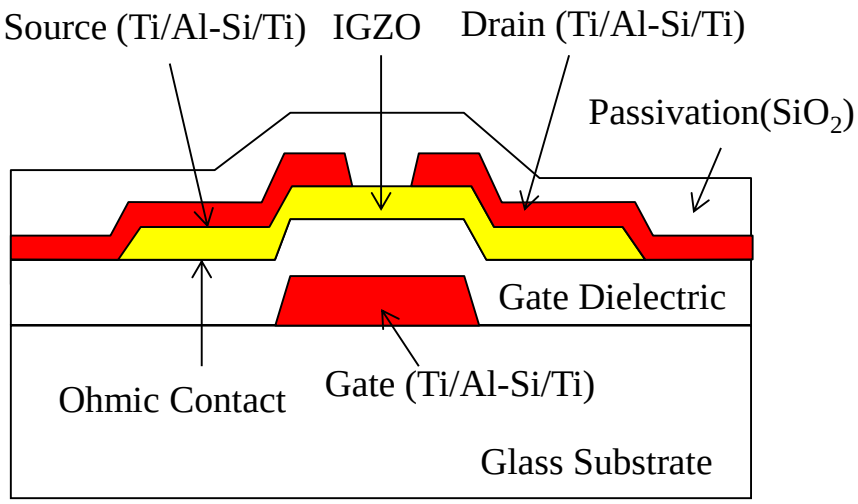


Fig. 6.1. Cross-sectional view of IGZO MOSTFT [1.40].

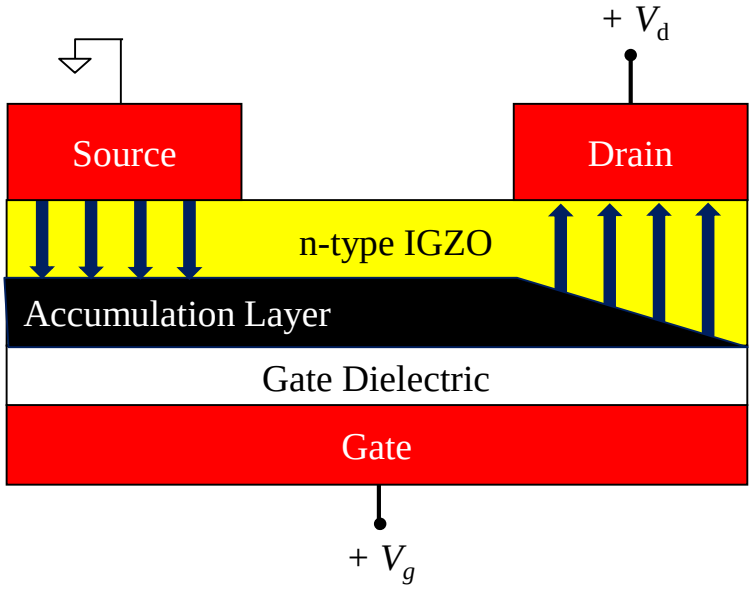


Fig. 6.2. Schematic description of IGZO MOSTFT operating in accumulation.

B. Basic Operation

As shown in Fig. 6.2, the IGZO MOSTFT is turned on by applying a positive voltage to the gate in order to form an accumulation layer on its channel. This operation differs from that of Si MOSTFT, which is turned on by formation of an inversion layer on its channel.

III. Memory Cell Structure

Figure 6.3 shows an equivalent circuit of the proposed FLOTOS memory cell together with a conventional nondestructive readout DRAM cell. The FLOTOS memory cell can be fabricated using the above-mentioned standard IGZO MOSTFT process without any additional process or mask steps.

The proposed FLOTOS memory cell consists of a memory transistor (MT) to read information nondestructively, a write select transistor (T1), a read select transistor (T2), and a storage capacitor with a capacitance of C_s = 150 fF. One side terminal of the storage capacitor is connected to the gate (MG) of MT, while the other side terminal is connected to a ground (V_{ss}) line. The drain (D) of MT is connected to MG through T1 and is also connected to a power supply (V_{dd} =10V) line through T2. Here, the V_{ss} and V_{dd} lines are formed by the source/drain electrodes shown in Fig. 6.1. The source (S) of

MT is connected to a sense-amplifier through a bit line (BL). The T1 and T2 gates are connected to write-select lines (WSELs) and read-select lines (RSELs), respectively. The BLs are arranged perpendicular to the WSELs and RSELs. The read and write operations are carried out using these three control lines.

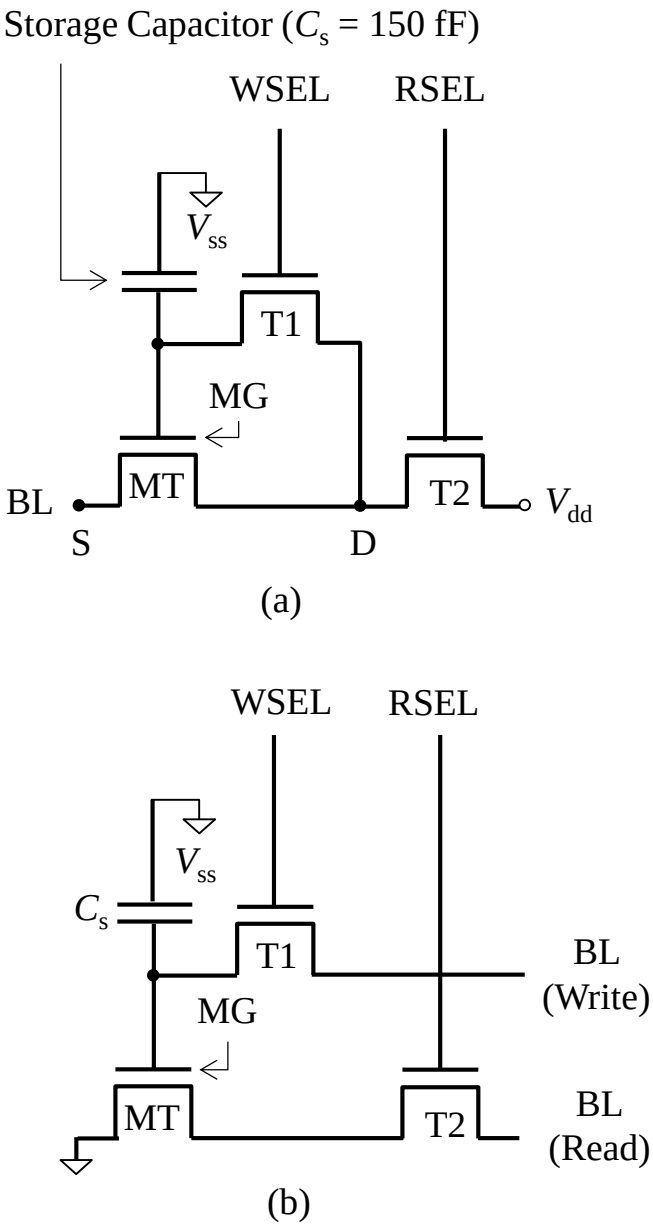


Fig. 6.3. Equivalent circuit of (a) proposed FLOTOS memory cell, and (b) conventional nondestructive readout DRAM cell [1.40].

IV. PRINCIPLE OF CELL OPERATION

Figure 6.4 shows the operation concept of the precharge-assisted threshold compensation (PAC) writing technique for a FLOTOS memory cell. First, a high voltage is applied to the RSEL and WSEL to precharge a storage capacitor with the V_{dd} power supply through T1 and T2, as shown in Fig. 6.4(a). The source of MT is biased to 5 V to reduce the precharge current. The precharge operation is conducted on all cells simultaneously to reduce the power dissipation by the precharge current. Next, the write voltage ($V_{bit,write}$) is applied to the source, and then the drain of the MT is subsequently disconnected from V_{dd} by applying a negative voltage to the RSEL, as shown in Fig. 6.4(b). As a result, the gate potential (V_g) of the MT is reduced closely to the voltage higher than $V_{bit,write}$ by its threshold voltage (V_{th}) by the discharge current from the gate to the source of the MT. The PAC technique is proposed to reduce the voltage swing of the BLs being driven at a high frequency without the influence of the power dissipation caused by the precharge current. Reducing the voltage swing of the BL is an effective way to save power dissipation in write cycles because the driving power increases with the square of the swing of driving voltage.

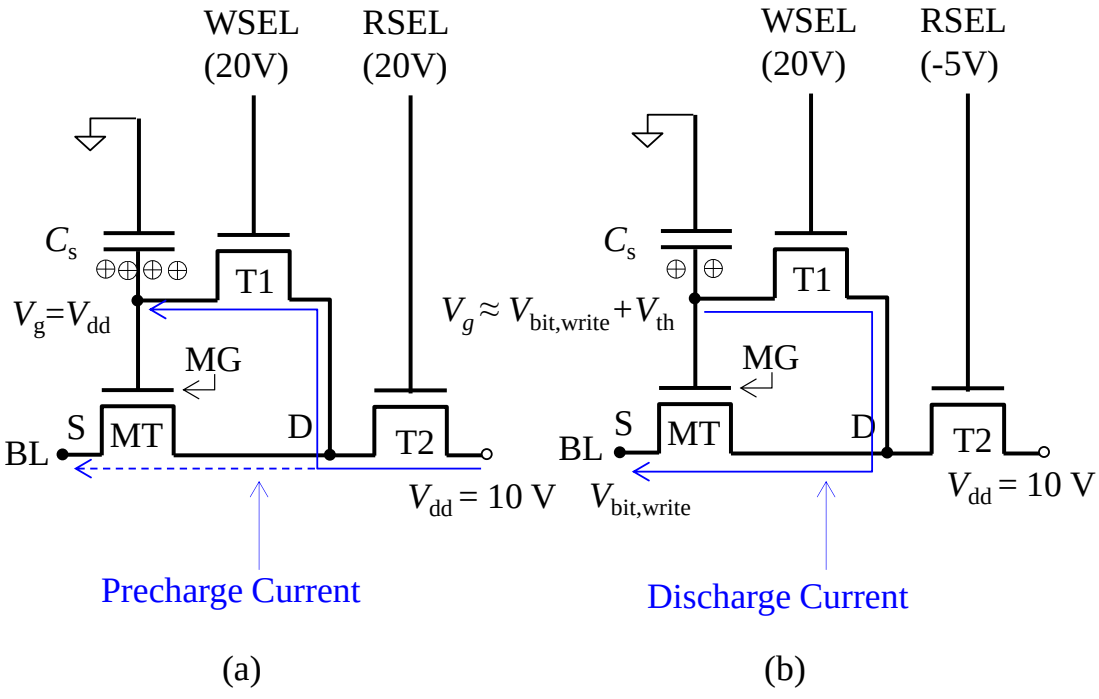


Fig. 6.4. Operation principle of precharge-assisted threshold compensation writing: (a) precharge and (b) discharge in write mode [1.40].

Figure 6.5 shows the equivalent circuit of a 4 bit FLOTOS memory array. The operating conditions of the hold, read, and write modes are shown in Table 6.1 for the memory array shown in Fig. 6.5. In the hold mode to maintain the charges stored on a storage capacitor, a negative voltage of -5 V is applied to all WSELs to turn off T1. In the read mode, the drain of MT is connected to V_{dd} with 20 V on the selected RSEL1 and the BLs are grounded. A negative gate voltage of -5V is also applied to the unselected RSEL2. In the write mode, the precharge operation is carried out and the discharge operation is subsequently achieved by applying 0 V to BL1 for the off-state cell (1,1) and 5 V to BL2 for the on-state cell (1,2).

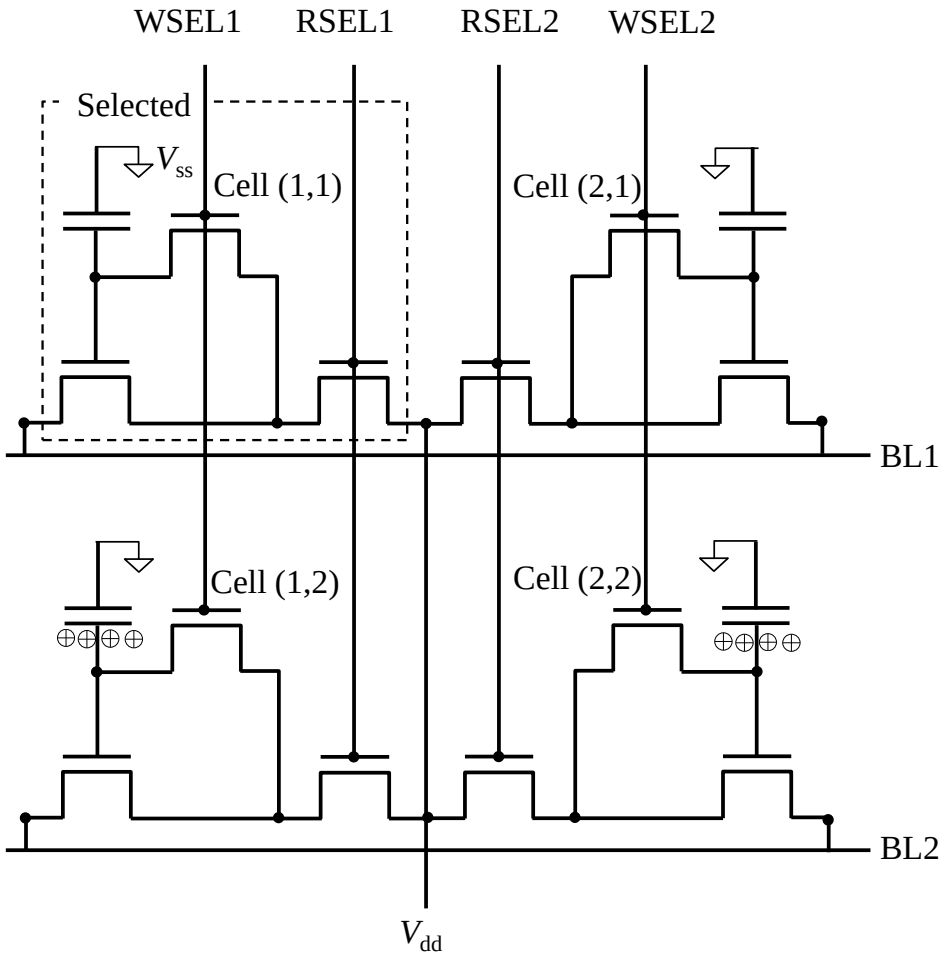


Fig. 6.5. Equivalent circuit of 4-bit memory array [1.40].

Table 6.1

Operating Conditions for the Memory Array Shown in Fig. 6.5 [1.40].

Mode		Applied Voltages (V)					
		WSEL1	WSEL2	RSEL1	RSEL2	BL1	BL2
Hold		-5	-5	-5	-5	0	0
Read		-5	-5	20	-5	0	0
Write	Precharge	20	20	20	20	5	5
		-5	-5	20	20	5	5
	Discharge	20	-5	-5	-5	0	5

V. CELL PERFORMANCE

The write speed, read current, and compensation voltage were simulated with a simple model, and the validity of the results was then verified experimentally.

A. Simulation results

The memory cell is written using a discharge current from the gate to the source of MT through its channel. The discharge current (I_{write1}) to the source of MT in the initial strong accumulation region is calculated using

$$I_{\text{write1}} = K\left(V_{\text{g}} - V_{\text{bit,write}} - V_{\text{th}}\right)^2 + I_0 \;\; for\; V_{\text{g}} - V_{\text{bit,write}} \geq V_{\text{th}}, \tag{6.1}$$

where K is a transfer conductance parameter. V_{th} is defined as the intercept with the gate-to-source voltage axis of the extrapolation of the linear part in the square root of the source current as a function of the gate-to-source voltage. I_0 is a fitting parameter that depends on the leakage current at $V_{\text{g}} - V_{\text{bit,write}} = V_{\text{th}}$. On the other hand, the discharge current (I_{write2}) to the source in the weak accumulation region is calculated using

$$I_{\text{write2}} = I_0 \exp\left(\frac{V_{\text{g}} - V_{\text{bit,write}} - V_{\text{th}}}{S \log e}\right) \; for\; V_{\text{g}} - V_{\text{bit,write}} \leq V_{\text{th}}, \tag{6.2}$$

where S is a fitting parameter that depends on the subthreshold swing. The discharge current (I_{write}) in all regions from weak to strong accumulation is described by

$$I_{\text{write}} = \left(\frac{1}{I_{\text{write1}}} + \frac{1}{I_{\text{write2}}}\right)^{-1}. \tag{6.3}$$

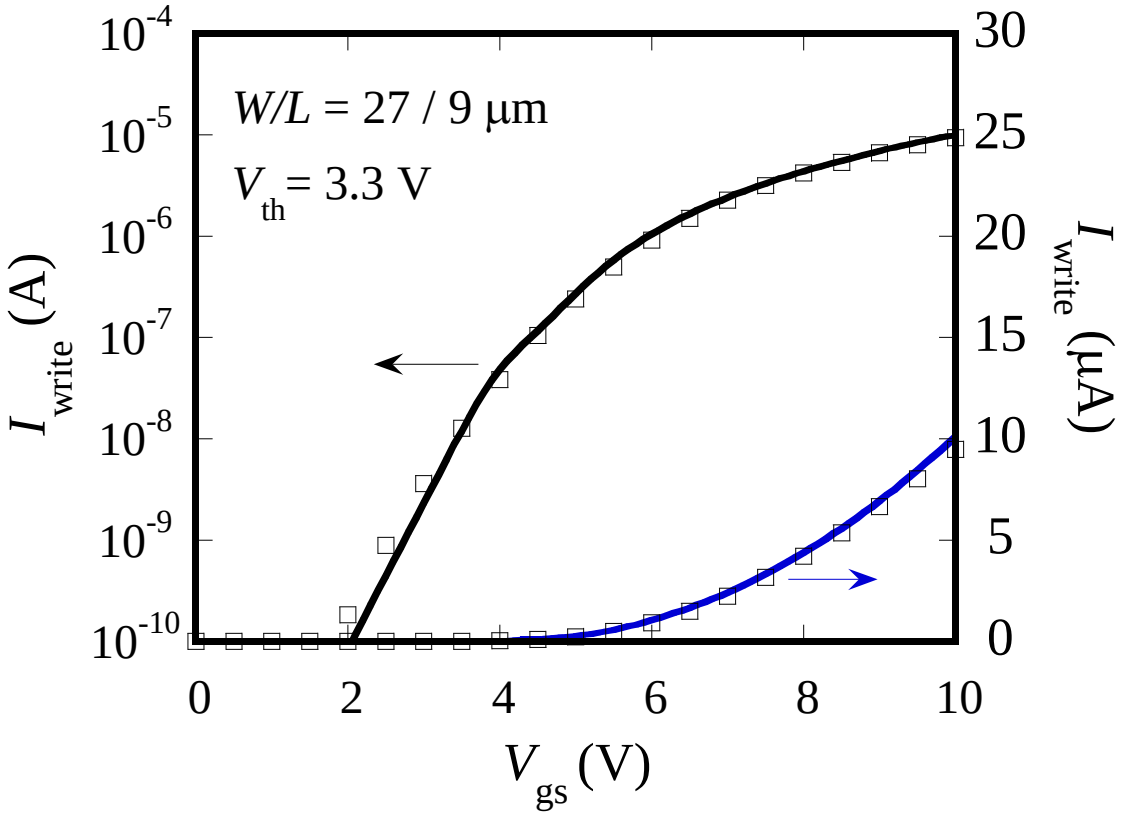


Fig. 6.6. Discharge current (I_{write}) as a function of gate-to-source voltage (V_{gs}) of MT. The symbols represent the measured data and the lines represent the curve fitting to Eq. (6.3) [1.40].

The gate-to-source voltage V_{gs} dependence of I_{write} obtained from Eq. (6.3) shows good fitting to the measured data for a single MOSTFT, as shown in Fig. 6.6. The discharge time (t_w) dependence of V_{gs} (=V_g – $V_{bit,write}$) is given by

$$V_{gs} = V_{dd} - V_{bit,write} - \frac{1}{C_s}\int_0^{t_w} I_{write}(t)dt, \qquad (6.4)$$

Figure 6.7 shows the simulation results of the t_w dependence of the V_{gs} values of MTs in the on and off states, which are obtained using Eq. (6.4). The V_{gs} values of MTs in the on and off states decrease with an increase in t_w and become very close to V_{th} in 1 μs. The read current of MT in the on-state (strong accumulation region) is given by

$$I_{on} = K{V_{bit,write}}^2\left(1 + \frac{\Delta V_{g,write} - V_{th}}{V_{bit,write}}\right)^2, \qquad (6.5)$$

where $\Delta V_{g,write}$ is the compensation voltage defined as V_{gs} when t_w is 1 μs in Eq. (6.4). On the other hand, the read current I_{off} of MT in the off state (weak accumulation region) is

given by

$$I_{\text{off}} = I_0 \exp\left(\frac{\Delta V_{\text{g,write}} - V_{\text{th}}}{S \log e}\right). \tag{6.6}$$

The differences ($\Delta V_{\text{g,write}} - V_{\text{th}}$) between $\Delta V_{\text{g,write}}$ and V_{th} are 80 and 210 mV for the typical MTs in the on and off states, respectively, as shown in Fig. 6.7. The high I_{on} of 7.2 μA was obtained with a low I_{off} of 99 nA by substituting the $\Delta V_{\text{g,write}}$ obtained using Eq. (6.4) into Eqs. (6.5) and (6.6). Since the deviation of $\Delta V_{\text{g,write}} - V_{\text{th}}$ is very small compared with $V_{\text{bit,write}}$ in Eq. (6.5), a tight I_{on} distribution is obtained. On the other hand, a change in $\Delta V_{\text{g,write}} - V_{\text{th}}$ causes an exponential change in I_{off} in Eq. (6.6). The deviation of I_{off} becomes larger than that of I_{on}.

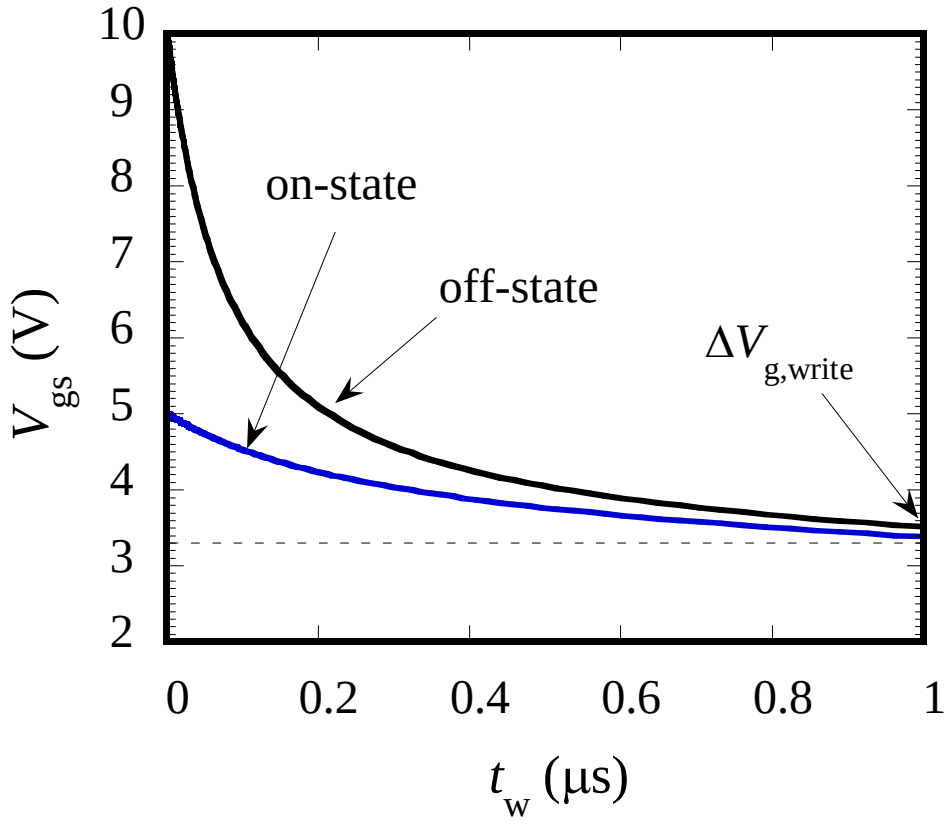


Fig. 6.7. Simulation results of discharge time (t_w) dependence of gate-to-source voltage (V_{gs}) of MT [1.40].

B. Experimental results

The effect of V_{th} compensation on the distribution of I_{on} and I_{off} after a write operation with t_w = 1 μs was evaluated on the FLOTOS memory cells as shown in Fig. 6.8. A mean value of I_{on} as high as 7.1 μA is obtained with a standard deviation as small as 7.0% for the proposed cell, while a mean value of I_{on} as low as 0.74 μA with a standard deviation as large as 76% is obtained for the conventional cell (Fig. 6.3(b)). The proposed cell using the threshold compensation write technique can increase the mean value of I_{on} by about one order of magnitude and reduces the relative standard deviation by about one order of magnitude. From these results the mean threshold voltage is estimated to be as high as 3.5V. On the other hand, I_{off} increases by using the compensation write technique. However, the mean value of I_{off} is as low as 54 nA with a standard deviation of 22%, and is acceptable for readout.

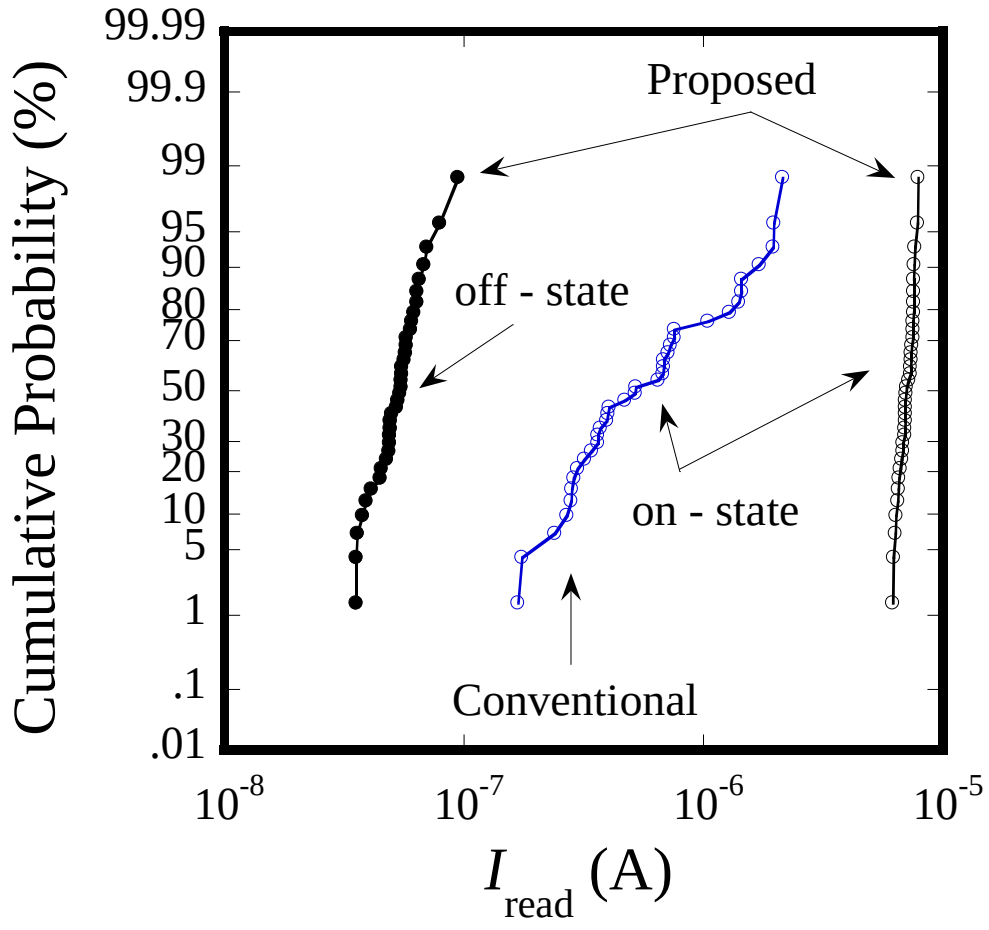


Fig. 6.8. Measured read current (I_{read}) distribution after write operation under bias conditions shown in Table 6.1 for on and off state cells, where discharge time is 1 μs [1.40].

The FLOTOS can satisfy a high I_{on} > 5.7 μA with an overdrive voltage as low as 5 V on BL and a read current window I_{on}/I_{off} = 50, which are sufficient for reading out the data stored in a memory cell. In addition, the effect of the variation in V_{th} on the read current window degradation can be compensated for after every write operation by the PAC technique, which makes it possible to realize an infinite number of write endurance cycles at a low BL voltage of 5 V.

VI. RELIABILITY

A number of low-field (< 3 MV/cm) gate-bias stressing studies have been conducted on IGZO MOSTFTs. These studies indicated that positive gate-bias stressing produces a positive V_{th} shift with little changes in the mobility and subthreshold slope [6.15] – [6.27]. However, negative gate-bias stressing results in a negligible change in V_{th} in the dark [6.15]. The positive bias stress-induced V_{th} shift of MT becomes a serious issue for a continuous and long-term reading operation because the V_{th} shift cannot be compensated for. In this section, the saturation value of the V_{th} shift of MT was predicted under the read stress conditions shown in Table 6.1. In addition, the V_{th} shift effects on cell performance and power consumption were also studied.

A. Prediction of V_{th} shift

The DC stress time (t_{st}) dependence of the stress-induced V_{th} shift ($\Delta V_{th,st}$) was evaluated under different gate-bias stress conditions with a drain voltage of 10 V (type A) and without the drain bias (type B), as shown in Fig. 6.9, where the gate-bias stress voltages ($V_{g,st}$) were $V_{th,init}$ + 5 V and $V_{th,init}$ + 10 V. Here, $V_{th,init}$ is the threshold voltage of MT before bias stressing. The stress conditions of the type A with a drain voltage and type B without a drain voltage correspond to those of cells (1,2) and (2,2) in the memory array shown in Fig. 6.5, respectively. The larger amount of V_{th} shift is observed with a higher $V_{g,st}$ under the type B stress conditions. However, no apparent effect of $V_{g,st}$ on the V_{th} shift was observed under the type A stress conditions. The largest V_{th} shift is observed with $V_{th,init}$ + 10 V under the type B stressing. The Joule heating generated by the bias-current stressing effect on the enhancement of the V_{th} shift is not observed under the type A stressing [6.16]. The V_{th} shift was reduced by applying the drain bias. These results suggest that the amount of V_{th} shift is mainly determined by the effective positive electrical fields across the gate dielectric.

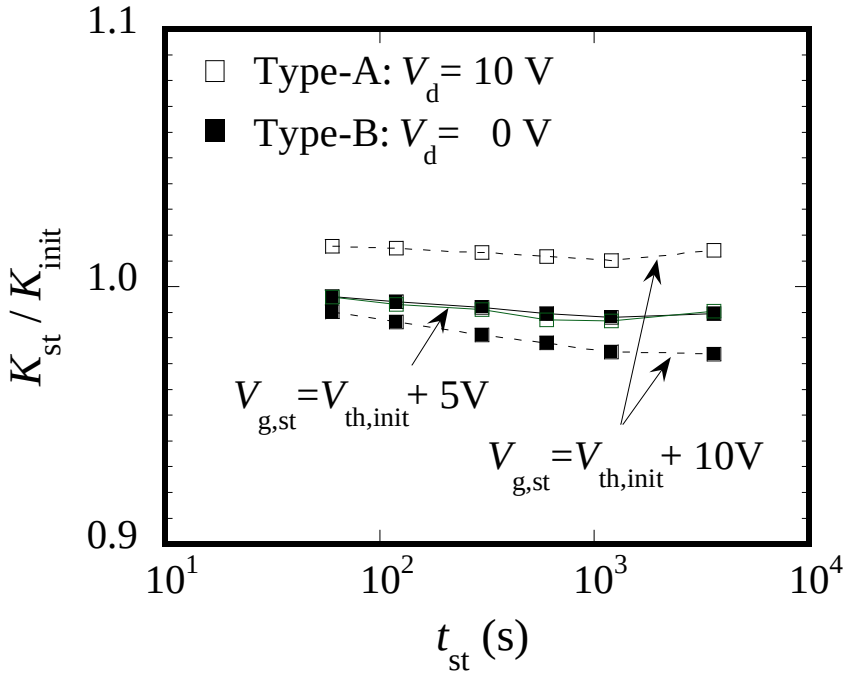


Fig. 6.9. DC stress time (t_{st}) dependence of V_{th} shift ($\Delta V_{th,st}$) under different gate-bias stress conditions with drain voltages ($V_{d,st}$) of 10 V (type A) and 0 V (type B), where gate voltages ($V_{g,st}$) are $V_{th,init}$ +5 V and $V_{th,init}$ + 10 V [1.40].

Figure 6.10 shows the DC stress time dependence of the normalized transfer conductance parameter (K_{st} / K_{init}), where K_{init} and K_{st} are the transfer conductance parameters before and after gate-bias stressing. Little change in K_{st} / K_{init} was observed for all stress bias conditions. These results suggest that no new interfacial traps were created during the gate-bias stressing with and without the drain bias.

Empirically, the stress-induced V_{th} shift, attributed to the trapping of charges in the existing traps without the creation of additional traps, can be predicted by fitting the well-known stretched-exponential function to the measured data [6.17] – [6.19], [6.21], [6.26] – [6.29]. The stretched-exponential model is described by

$$\Delta V_{\mathrm{th,st}} = \Delta V_{\mathrm{thmax,st}} \left\{ 1 - \exp\left[-\left(\frac{t_{\mathrm{st}}}{\tau} \right)^{\beta} \right] \right\}, \tag{6.7}$$

where $\Delta V_{\mathrm{thmax,st}}$ [1.40]is the saturation value of the V_{th} shift, τ is the characteristic trapping time of carriers, and β is the stretched-exponential exponent. In addition, the $\Delta V_{\mathrm{thmax,st}}$ dependence on the effective gate-bias stress voltage ($V_{g,st} - V_{th,init}$) is given by

$$\Delta V_{thmax,st} = \left(V_{g,st} - V_{th,init} \right)^{\alpha}, \tag{6.8}$$

where α is a parameter associated with the interface qualities [6.17], [6.28].

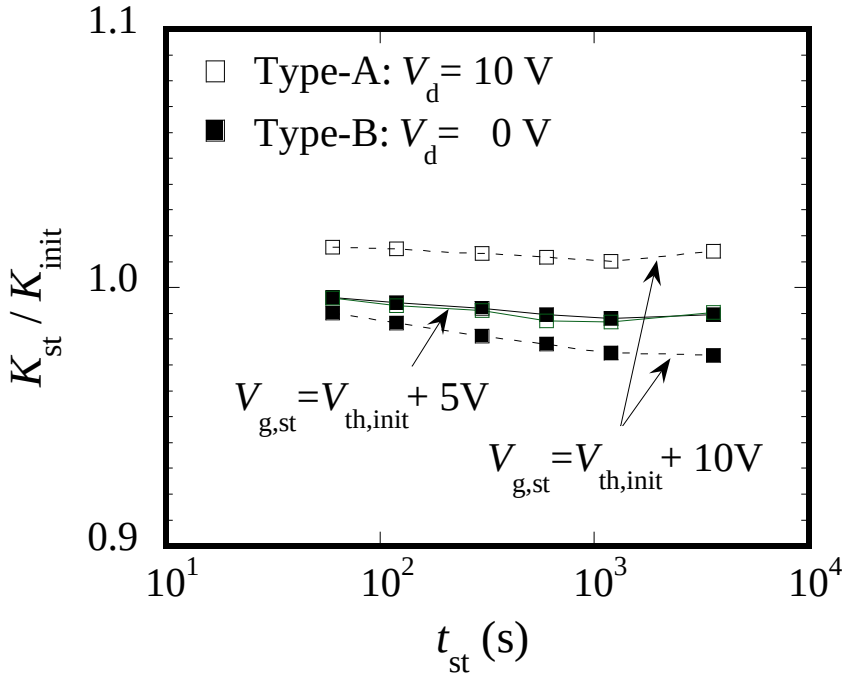


Fig. 6.10. DC stress time (t_{st}) dependence of normalized transfer conductance (K_{st}/K_{init}) under different gate-bias stress conditions with drain voltages (V_{dst}) of 10 V (type A) and 0 V (type B), where gate voltages ($V_{g,st}$) are $V_{th,init}$ +5 V and $V_{th,init}$ +10 V [1.40].

The proposed PAC technique can keep the effective gate-bias stress voltage as low as 5 V for all the on-state cells, which is a good approach to reducing the V_{th} shift as well as for low-power operation.

The stretched-exponential fitting curve showed good agreement with the measured data under type B stressing with different $V_{g,st}$ values as shown in Fig. 6.11. The parameter values are listed in Table 6.2, which are similar to the stretch-exponential fitting parameters investigated in the previous paper [6.12] – [6.14], [6.20], [6.21]. The saturation values of the V_{th} shift are 2.2 and 3.5 V for $V_{g,st} = V_{th,init} + 5$ V and $V_{g,st} = V_{th,init} + 10$ V, respectively, with the same β and τ. The saturation value of the stress-induced V_{th} shift is predicted to be as low as 2.2 V with $V_{g,st} = V_{th,init} + 5$ V under the type B stressing for a FLOTOS, which is the worst stress condition because the effective gate-bias stress voltage is kept as low as 5 V by the proposed PAC technique.

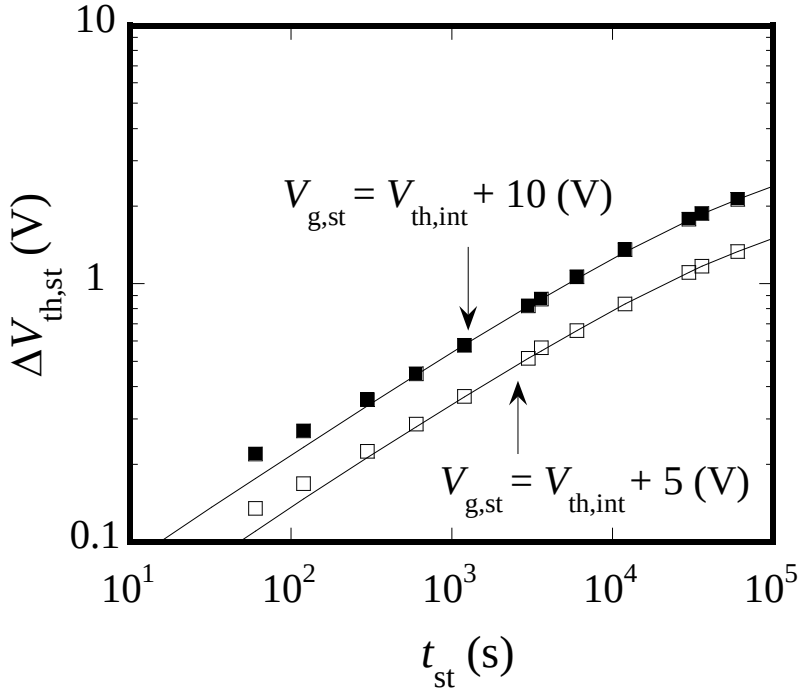


Fig. 6.11. DC stress time (t_{st}) dependence of V_{th} shift ($\Delta V_{th,st}$) for different $V_{g,st}$ values. The symbols represent measured data and the lines represent the curve fitted to Eq. (6.7) with the parameters shown in Table 6.2 [1.36].

Table 6.2

Fitting Parameter Values for Stretched Exponential Model for Different Gate-Bias Stress Voltages $V_{th,init}$ + 5 V and $V_{th,init}$ +10 V [1.40].

Parameter	Fitting Values	
	$V_{g,st} = V_{th,init} + 5$ V	$V_{g,st} = V_{th,init} + 10$ V
β	0.42	0.42
τ (s)	7.0×10^4	7.0×10^4
$\Delta V_{thmax,st}$ (V)	2.2	3.5

B. Effect of V_{th} shift on cell performance

The on-state read currents before and after gate-bias stressing are given by

$$I_{\mathrm{on,init}} = K_{\mathrm{init}}\left(V_{\mathrm{g,st}} - V_{\mathrm{th,init}}\right)^2, \tag{6.9}$$

$$I_{\mathrm{on,st}} = K_{\mathrm{st}}[V_{\mathrm{g,st}} - (V_{\mathrm{th,init}} + \Delta V_{\mathrm{th,st}})]^2, \tag{6.10}$$

where $I_{\mathrm{on,init}}$ and $I_{\mathrm{on,st}}$ are the on-state read currents before and after the gate-bias stressing, respectively. From Eqs. (6.9) and (6.10), the normalized on-state read current ($I_{\mathrm{on,st}}$ / $I_{on,init}$) is given by

$$\frac{I_{\mathrm{on,st}}}{I_{\mathrm{on,init}}} = \frac{K_{\mathrm{st}}}{K_{\mathrm{init}}}\left(1 - \frac{\Delta V_{\mathrm{th,st}}}{V_{\mathrm{g,st}} - V_{\mathrm{th,init}}}\right)^2. \tag{6.11}$$

The increase in $V_{\mathrm{g,st}}$ causes an increase in $\Delta V_{\mathrm{th,st}}$ from Eqs. (6.7) and (6.8), while it reduces the effect of $\Delta V_{\mathrm{th,st}}$ on the degradation of $I_{\mathrm{on,st}}$ / $I_{\mathrm{on,init}}$ in Eq. (6.11). Figure 6.12 shows the DC stress time dependence of the normalized I_{on} under the type B stressing. The largest degradation of the normalized I_{on} was observed with a lower $V_{\mathrm{g,st}}$ = 5 V under the type B stressing. This result suggests that the effect of $V_{\mathrm{g,st}}$ on the degradation of I_{on} is larger than the increase in $\Delta V_{\mathrm{th,st}}$ with an increase in $V_{\mathrm{g,st}}$ in Eq. (6.11). By substituting Eq. (6.7) into Eq. (6.11), the normalized $I_{\mathrm{on,st}}$ can be rewritten as the following equation, assuming K_{st}/K_{init} = 1:

$$\frac{I_{\mathrm{on,st}}}{I_{\mathrm{on,init}}} = \left\{1 - \frac{\Delta V_{\mathrm{thmax,st}}}{V_{\mathrm{g}} - V_{\mathrm{th,init}}}\left\{1 - \exp\left[-\left(\frac{t_{\mathrm{st}}}{\tau}\right)^{\beta}\right]\right\}\right\}^2. \tag{6.12}$$

From the results of the fitting to Eq. (6.12), as shown in Fig. 6.12, the lifetime for the read operation with $V_{\mathrm{g,st}} = V_{\mathrm{th,init}}$ + 5 V under the type B stressing (worst condition) is estimated to be 13.8 h. Here, the lifetime is defined as the DC gate-bias stress time when the normalized $\mathrm{I_{on,st}}$ reaches 0.5. This lifetime is acceptable for battery-powered wireless mobile terminal applications such as smartphones because it is possible to compensate for the V_{th} shift each time the battery is charged.

C. Effect of V_{th} shift on power dissipation

In the proposed PAC write scheme, all the operating voltages except V_{dd} are not affected by the V_{th} variation. In addition, there is only one cycle of BL driving with a high V_{dd} for writing all cells in a memory array. The increase in V_{dd} caused by the stress-induced V_{th} shift is about 2.2 V, indicating little impact of the V_{th} shift on power consumption.

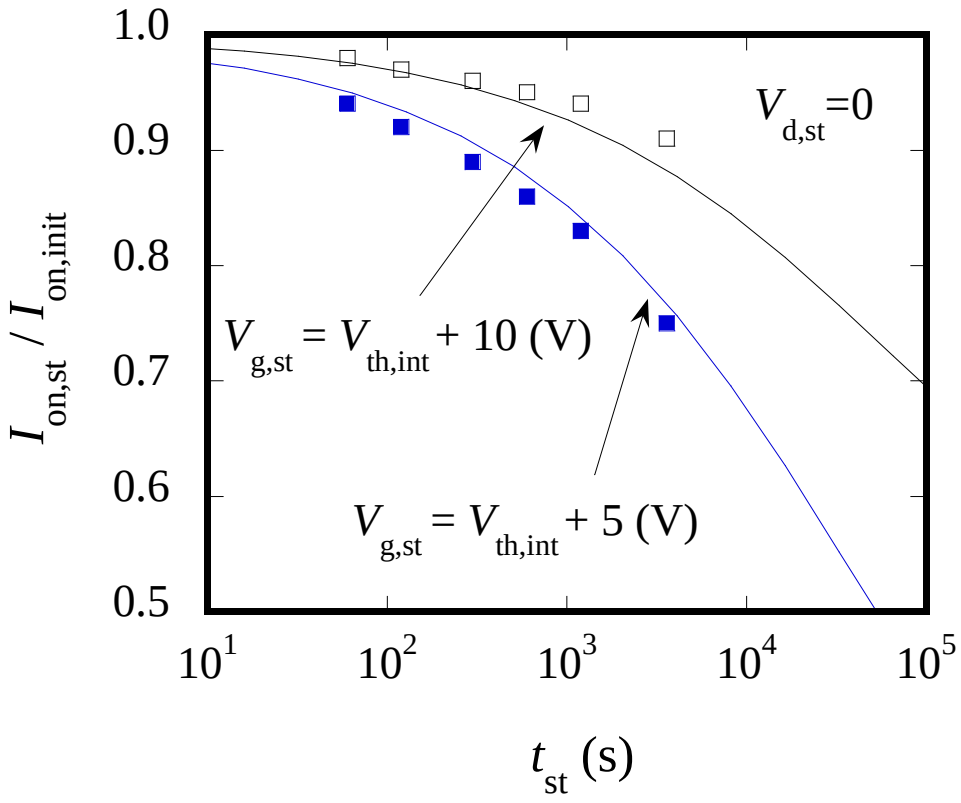


Fig. 6.12. DC stress time (t_w) dependence of normalized read current ($I_{on,st}/I_{on,init}$) for different gate voltages ($V_{g,st}$) under type B stressing. The symbols represent the measured data and the lines represent the calculation results using Eq. (6.12) [1.35].

VII. DATA RETENTION CHARACTERISTICS

The bias stressing during the read operation causes a charge gain for the selected memory cell (1,1) in the off state whose drain is connected to V_{dd} and the charge loss for the unselected memory cell (2,2) in the on-state whose drain is grounded, as shown in Fig. 6.5. The data retention characteristics were evaluated under the worst bias-stress condition of read operation at 60 °C after 10^9 cycles of write operation.

Figures 6.13(a) and 6.13(b) show the equivalent circuit of the test device used to evaluate data retention together with the bias conditions for the charge gain in the off-state cell and the charge loss in the on-state cell, respectively. First, a write operation is carried out with $V_g = V_{th,init}$ and $V_g = V_{th,init} + 5$ V for the off and on state memory cells, respectively. Then, the data on a storage capacitor is stored under the bias conditions shown in Fig. 6.13.

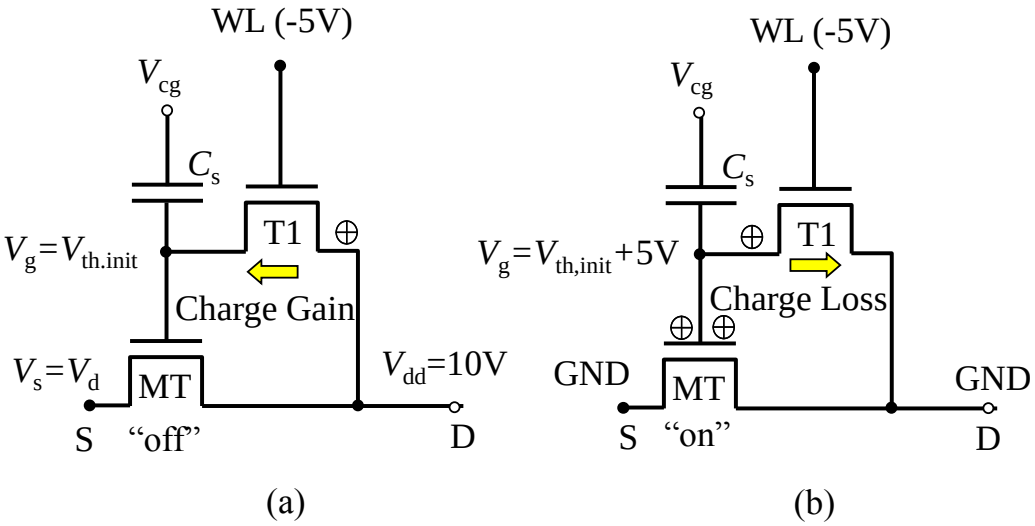


Fig. 6.13. Equivalent circuit for test device with bias conditions for data retention. (a) Charge gain in off-state cell. (b) Charge loss in on-state cell [1.40].

Next, the first measurement of the transfer curve was conducted after 10 h of data retention and the second measurement of the transfer curve was subsequently conducted after refreshing by rewriting the initial data. Little change in the transfer characteristics was obtained between the first and second measurements for the on-state and off-state memory cells, as shown in Fig. 6.14.

The change in threshold voltage ($\Delta V_{\mathrm{th,ret}}$) between the first and second measurements of the transfer curves is less than + 184 mV for the on-state cell and - 70 mV for the off-state cell. The stress-induced V_{th} shift of MT during data retention is subtracted from $\Delta V_{th,ret}$ by subtracting the threshold voltage after refreshing from that after data retention. The average leak current ($I_{\mathrm{leak,ret}}$) can be calculated using

$$I_{\mathrm{leak,ret}} = C_s \frac{\Delta V_{\mathrm{th,ret}}}{t_{\mathrm{ret}}}, \qquad (6.13)$$

where t_{ret} is the retention time (10 h). From Eq. (6.13), $I_{\mathrm{leak,ret}}$ is estimated to be 2.8×10^{-20} A/μm at 60 °C for the on-state cell, while the $I_{\mathrm{leak,ret}}$ of the low-temperature polycrystalline silicon MOSTFT is higher than 0.5 pA/μm [6.30]. The amount of refresh power in the IGZO-based FLOTOS is expected to be reduced by more than 7 orders of magnitude compared with that in the Si-based memory. The FLOTOS does not need a refresh circuit for battery-powered wireless mobile terminal applications such as smartphones because the retention time of 10 h at 60 °C corresponds to the maximum usage time on a single battery charge.

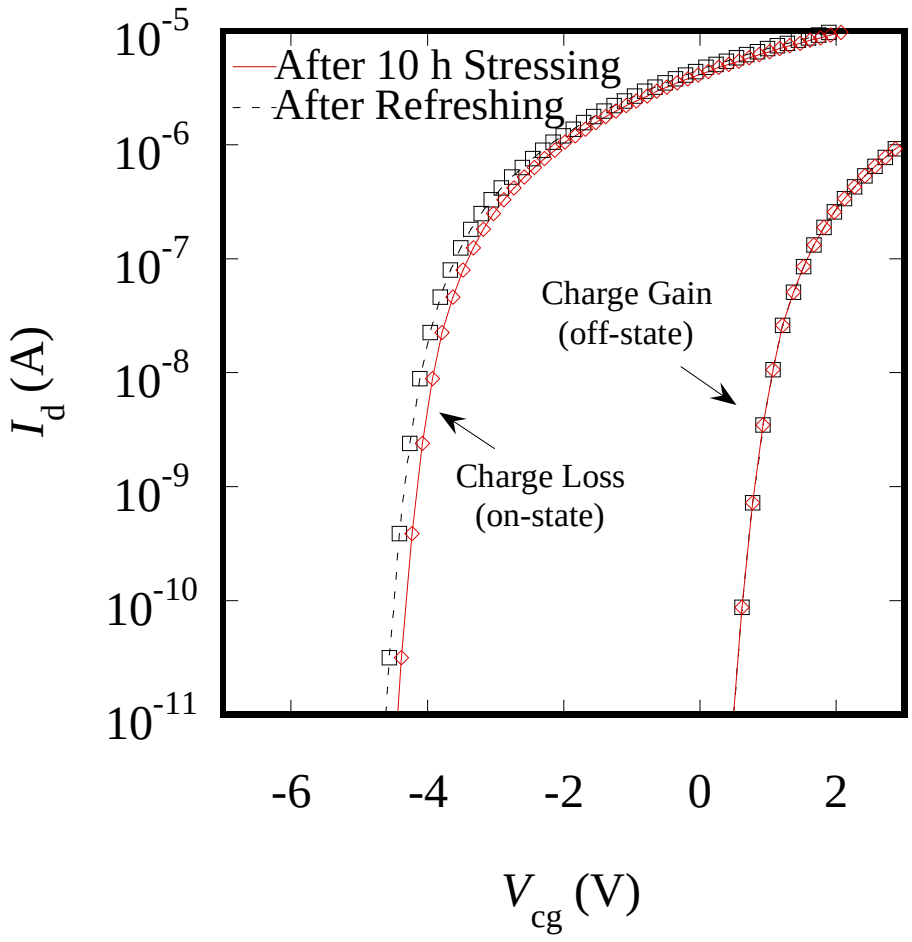


Fig. 6.14. Transfer characteristics after 10 h data retention (first measurement) and after refreshing (second measurement) for charge loss in on-state cell and charge gain in off-state cell [1.40].

VIII. CONCLUSIONS

A FLOTOS memory, that consists of wide-band-gap IGZO MOSTFTs and a storage capacitor, has been developed. The FLOTOS is fabricated using a standard IGZO MOSTFT process without any additional process or mask steps. The proposed precharge-assisted PAC technique makes it possible to realize an infinite number of write cycles and a low-power write operation with a low bit-line voltage of 5 V. Furthermore, excellent data retention longer than 10 h is obtained at 60 °C even under the worst bias-stress conditions. This is due to the ultra-low off-state leakage (2.8×10^{-20} A/μm) of the IGZO MOSTFTs, which is estimated to be smaller by more than 7 orders of

magnitude than that of polycrystalline silicon MOSTFTs. The off-state leakage current of the IGZO MOSTFT is expected to be further reduced to 3×10^{-24} A/μm at 60 °C from a previous study [6.9]. The FLOTOS has possibility to store charge on a storage capacitor (C_s = 150fF) for 10 years without refreshing similarly to a nonvolatile memory. It is expected that FLOTOS will become a vital technology to meet the demands for ever thinner and more lightweight mobile devices.

REFERENCES

[6.1] K. Nomura, A. Takagi, T. Kamiya, H. Ohta, M. Hirano, and H. Hosono, "Amorphous Oxide Semiconductors for High-Performance Flexible Thin Film Transistors," *Japanese Journal of Applied Physics*, vol. 45, pp. 4303 – 4308, May. 2006.

[6.2] R. Hayashi, A. Sato, M. Ofuji, K. Abe, H. Yabuta, M. Sano, H. Kumomi, K. Nomura, T. Kamiya, M. Hirano, and H. Hosono, "Improved Amorphous In-Ga-Zn-O TFTs," *SID International Symposium Digest of Technical Papers*, 2008, pp. 621 – 624.

[6.3] J. F. Wager, "Amorphous Oxide Semiconductor Thin-Film Transistors: Performance & Manufacturability for Display Applications," *SID International Symposium Digest of Technical Papers*, 2009, pp. 181 – 183.

[6.4] H. H. Lu, H. C. Ting, T. H. Shih, C. Y. Chen, C. S. Chuang, and Y. Lin, "32-inch LCD Panel Using Amorphous Indium-Gallium-Zinc-Oxide TFTs," *SID International Symposium Digest of Technical Papers*, 2010, pp. 1136 – 1138.

[6.5] T. Arai, N. Morosawa, K. Tokunaga, Y. Terai, E. Fukumoto, T. Fujimori, T. Nakayama, T. Yamaguchi, and T. Sasaoka, "Highly Reliable Oxide-Semiconductor TFT for AM-OLED Display," *SID International Symposium Digest of Technical Papers*, 2010, pp. 1033 – 1036.

[6.6] Y. G. Mo, M. Kim, C. K. Kang, J. H. Jeong, Y. S. Park, C. G. Choi, H. D. Kim, and S. S. Kim, "Amorphous Oxide TFT Backplane for Large Size AMOLED TVs," *SID International Symposium Digest of Technical Papers*, 2010, pp. 1037 – 1040.

[6.7] H. Yamaguchi, T. Ueda, K. Miura, N. Saito, S. Nakano, T. Sakano, K. Sugi, I. Amemiya, and M. Hiramatsu, "11.7-inch Flexible AMOLED Display Driven by a-IGZO TFTs on Plastic Substrate," *SID International Symposium Digest of Technical Papers*, 2012, pp. 1002 – 1005.

[6.8] H. Godo, A. Miyanaga, K. Kusunoki, K. Toyotaka, T. Nishi, K. Moriya, S. Fukai, Y. Kubota, K. Wakimoto, J. Koyama, S. Yamazaki, Y. Oikawa, K.

Okazaki, and M. Sakakura, "Modeling and Measurement of Ultra-low Leakage Current of IGZO TFTs and New Driving Method of LCDs," *International Display Workshop*, 2010, pp. 235 – 238.

[6.9] M. Murakami, K. Kato, K. Inada, T. Matsuzaki, Y. Takahashi, and S. Yamazaki, "Theoretical Examination on a Significantly Low Off-State Current of a Transistors Using Crystalline In-Ga-Zn-Oxide," *Active-Matrix Flat-panel Display and Device*, 2012, pp. 171 – 174.

[6.10] H. Inoue, T. Matsuzaki, S. Nagatsuka, Y. Okazaki, T. Sasaki, K. Noda, D. Matsubayashi, T. Ishizu, T. Onuki, A. Isobe, Y. Shionoiri, K. Kato, T. Okuda, J. Koyama, and S. Yamazaki, "Nonvolatile Memory with Extremely Low-Leakage Indium-Gallium-Zinc-Oxide Thin-Film Transistor," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 9, pp. 2258 – 2265, Sep. 2012.

[6.11] K. Kato, Y. Shionori, Y. Sekine, K. Furutani, T. Hatano, T. Aoki, M. Sasaki, H. Tomatsu, J. Koyama, and S. Yamazaki, "Evaluation of Off-State Current Characteristics of Transistor Using Oxide Semiconductor Material, Indium–Gallium–Zinc Oxide," *Japanese Journal of Applied Physics*, 51, pp. 021201-1 – 021201-7, Jan. 2012.

[6.12] J. Jeong and Y. Hong, "Debye Length and Active Layer Thickness-Dependent Performance Variations of Amorphous Oxide-Based TFTs," *IEEE Transactions on Electron Devices*, vol. 59, no. 3, pp. 710 – 714, Mar. 2012.

[6.13] S. Amano, H. Harada, K. Akimoto, J. Sakata, T. Nishi, K. Moriya, K. Wakimoto, J. Koyama, and S. Yamazaki, "Low Power LC Display Using In-Ga-Zn-Oxide TFTs Based on Variable Frame Frequency," *SID International Symposium Digest of Technical Papers*, 2010, pp. 626 – 629.

[6.14] C. K. Kang, Y. S. Park, S. Park, Y. G. Mo, B. H. Kim, and S. S. Kim, "Integrated Scan Driver with Oxide TFTs Using Floating Gate Method," *SID International Symposium Digest of Technical Papers*, 2011, pp. 25 – 27.

[6.15] A. Suresh and J. F. Mutha, "Bias Stress Stability of Indium Gallium Zinc Oxide Channel Based transparent thin film transistors," *Applied Physics Letters*, vol. 92, pp. 033502-1 – 033502-3, Jan. 2008.

[6.16] M. Fujii, H. Yano, T. Hatayama, Y. Uraoka, T. Fuyuki, J. S. Jung, and J. Y. Kwon, "Thermal Analysis of Degradation in Ga_2O_3–In_2O_3–ZnO Thin Film Transistors," *Japanese Journal of Applied Physics*, vol. 47, no.8, pp. 6236 – 6240, Aug. 2008.

[6.17] J. M. Lee, I. T. Cho, J. H. Lee, and H. I. Kwon, "Bias-Stress-Induced Stretched-Exponential Time Dependence of Threshold Voltage Shift in

InGaZnO Thin Film Transistors," *Applied Physics Letters*, vol. 93, pp. 093504-1 – 093504-3, Sep. 2008.

[6.18] K. Nomura, T. Kamiya, M. Hirano, and H. Hosono, "Origins of Threshold Voltage Shifts in Room-Temperature Deposited and Annealed In–Ga–Zn–O Thin Film Transistors," *Applied Physics Letters*, vol. 95, pp. 013502-1 – 013502-3, Jul. 2009.

[6.19] S. Lee, K. Jeon, J. H. Park, S. Kim, D. Kong, D. M. Kim, D. H. Kim, S. Kim, S. Kim, J. Hur, J. C. Park, I. Song, C. J. Kim, Y. Park, and U. I. Jung, "Electrical Stress-Induced Instability of Amorphous Indium-Gallium-Zinc Oxide Thin Film Transistors under Bipolar AC Stress," *Applied Physics Letters*, vol. 95, pp. 132101-1 – 132101-3, Sep. 2009.

[6.20] J. Lee, J. S. Park, Y. S. Pyo, D. B. Lee, E. H. Kim, D. Stryakhilev, T. W. Kim, D. U. Jin, and Y. G. Mo, "The Influence of the Gate Dielectrics on Threshold Voltage Instability in Amorphous Indium-Gallium-Zinc Oxide Thin Film Transistors," *Applied Physics Letters*, vol. 95, pp. 123502-1 – 123502-3, Sep. 2009.

[6.21] J. F. Conley, "Instabilities in Amorphous Oxide Semiconductor Thin Film Transistors," *IEEE Transactions on Device and Materials Reliability*, vol. 10, no.4, pp. 460 – 475, Dec. 2010.

[6.22] M. Kimura and S. Imai, "Degradation Evaluation of α-IGZO TFTs for Application to AM-OLEDs," *IEEE Electron Device Letters*, vol. 31, no. 9, pp. 963 – 965, Sep. 2010.

[6.23] E. N. Cho, J. H. Kang, C. E. Kim, P. Moon, and I. Yun, "Analysis of Bias Stress Instability in Amorphous InGaZnO Thin-Film Transistors, "*IEEE Transactions on Device and Materials Reliability*, vol. 11, no. 1, pp. 112 – 117, Mar. 2011.

[6.24] T. C. Chen, T. C. Chang, T. Y. Hsieh, W. S. Lu, F. Y. Jian, C. T. Tsai, S. Y. Huang, and C. S. Lin, "Investigating the Degradation Behavior Caused by Charge Trapping Effect under DC and AC Gate-Bias Stress for InGaZnO Thin Film Transistor," *Applied Physics Letters*, vol. 99, pp. 022104-1 – 022104-4, Jul. 2011.

[6.25] T. C. Chen, T. C. Chang, T. Y. Hsieh, M. Y. Tsai, Y. T Chen, Y. C. Chung, H. C. Ting, and C. Y. Chen, "Self-heating Enhanced Charge Trapping Effect for InGaZnO Thin Film Transistor," *Applied Physics Letters*, vol. 101, pp. 042101-1 – 042101-4, Jul. 2012.

[6.26] S. Park, E. N. Cho, and I. Yun, "Threshold Voltage Shift Prediction for Gate Bias Stress on Amorphous InGaZnO Thin Film Transistors," *Microelectronics Reliability*, vol. 52, pp. 2215 – 2219, Jul. 2012.

[6.27] S. J. Kim, S. Y. Lee, Y. W. Lee, S. H. Kuk, J. Y. Kwon, and M. K. Hana, "Effect of Charge Trapping/Detrapping on Threshold Voltage Shift of IGZO TFTs under AC Bias Stress," *Electrochemical and Solid-State Letters*, vol. 15, no. 4, pp. H108 – H110, Feb. 2012.

[6.28] F. R. Libsch and J. Kanicki, "Bias-Stress-Induced Stretched-Exponential Time Dependence of Charge Injection and Trapping in Amorphous Thin Film Transistors," *Applied Physics Letters*, vol. 62, no. 11, pp. 1286 – 1288, Mar. 1993.

[6.29] S. Zafar, A. Callegari, E. Gusev, and M. V. Fischetti, "Charge Trapping Related Threshold Voltage Instabilities in High Permittivity Gate Dielectric Stacks," *Journal of Applied Phys*ics, vol. 93, no. 11, pp. 9298 – 9303, Jun. 2003.

[6.30] C. H. Kim, K. S. Sohn, and J. Jang, "Temperature Dependent Leakage Currents in Polycrystalline Silicon Thin Film Transistors," *Journal of Applied Phys*ics, vol. 81, no. 12, pp. 8084 – 8090, Jun. 1997.

CHAPTER 7
Conclusions

As described in Chapter 1, with the dynamic growth in wireless broadband networking infrastructure, the processing speeds and memory densities of mobile applications are rapidly increasing and this results in an increase in the power consumption of the system and liquid crystal display (LCD) panel. Furthermore, battery capacity is severely restricted by the demands for ever thinner and more lightweight mobile devices. This thesis has proposed and described various types of semiconductor memory devices and discussed their basic operation and design concepts, in order to address the issues caused by drawbacks in current volatile memories composed of silicon metal-oxide-semiconductor field-effect transistors (MOSFET).

Chapter 2 described a scalable multi-level-cell (MLC) virtual-ground flash memory cell technology that has potential for continued scaling of fast-access NOR flash memory. The virtual-ground cell is suitable for high-density nonvolatile storage applications because of its contactless structure. However, the cell-to-cell interference caused by the virtual-ground architecture makes it difficult to achieve MLC, the mechanisms of which can be explained by both the neighboring cell leakage and capacitive coupling models. An adequate gate coupling ratio and compensation for cell-to-cell interference are maintained to enable further scaling of virtual-ground MLC flash memory. A high gate coupling ratio of 0.6 is obtained using a novel bowl-shaped floating-gate structure cell technology without sacrificing cell size. Increasing the gate coupling ratio is important for reducing cell-to-cell interference and achieving low voltage operation. A novel array segmented virtual-ground architecture with bit-line isolation between neighboring segments and two-step programming utilizing a channel hot electron injection threshold voltage (V_{th}) compensation technique were proposed to reduce the number of neighboring cells that need to be programmed after programming a given cell, as well as to reduce the V_{th} shift of the neighboring cells. Adoption of this programming approach realizes a reduction in the V_{th} shift caused by the floating-gate-to-floating-gate coupling interference to less than 25% in the bit-line direction, compared to that with conventional

programming approaches, and the V_{th} shift is almost completely eliminated in the word-line direction without sacrificing program throughput. The proposed virtual-ground MLC floating-gate cell, which is as small as $3F^2$/bit (F is the minimum feature size) was successfully implemented into a test chip. Good reliability is also obtained up to 10^5 cycles. The proposed scalable MLC virtual-ground technology allows the CPU to execute a program directly without going through volatile memory, that leading to reduction in DRAM refresh power consumption. Thus, the technology is suitable for internet mobile applications that require faster execute-in-place performance and higher density.

Chapter 3 described a source-side injection single-polysilicon split-gate NOR (S4-NOR) flash memory cell compatible with a CMOS logic process. For the proposed cell, two gates are patterned simultaneously with a small gap length on the channel between the source and drain by using conventional photolithography technology beyond the 100 nm generation, and the floating-gate is capacitively coupled to the n-well memory gate through the gate oxide. This cell technology enables improved compatibility with CMOS logic processes. Fast programming, 5 μs, is achieved with a low, 10 μA, drain current. Injection efficiency is higher by four orders of magnitude as compared with that of the standard NOR cell and is insensitive to process parameters except the gap length in the split-gate transistor, the mechanism of which was clarified by the modified lucky-electron model. The proposed cell is suitable for embedded memories from the process control viewpoint and it also offers low power consumption. A high read current is obtained without suffering from soft-write effects. In addition, good endurance, long data retention and good program disturb immunity are also obtained without gap oxide leakage effect. Furthermore, due to the simple structure of the S4-NOR cell, it does not require any specialized process steps. Thus, the proposed S4-NOR flash memory is a promising candidate for next-generation embedded applications that require low power consumption and high performance without an increase in the cost of the logic chip.

Chapter 4 described a nonvolatile DRAM cell combining a DRAM cell with a floating-gate MOSFET, which stores all the data in its floating-gate in the event of sudden power failure and recalls the data to DRAM quickly at power-up. The store operation is completed in less than 10 ms with low power consumption by utilizing Fowler-Nordheim tunneling of electrons across the gate oxide between the DRAM storage-node and the floating-gate. The single cell shows excellent reliability such as

store endurance greater than 10^7 cycles at 85 °C and data retention in excess of 10 years under high, 150 °C, storage temperature. This floating-gate based nonvolatile DRAM cell was successfully demonstrated on a 1 Mbit test chip. A scalable nonvolatile DRAM cell, in which a floating-gate MOSFET is merged in serial between a switching transistor and a storage capacitor, was discussed for further reduction of cell area using the same store concept as the above-mentioned store scheme. These technologies allow the CPU to store the data being processed in the event of a sudden power failure, thus overcoming the main drawbacks of silicon MOSFET-based volatile memories.

Chapter 5 described a novel voltage-loss-compensation (VLC) memory in pixel that was developed for an advanced dual-mode LCDs, and which can display still image using an ultra-low-power-consumption memory mode in addition to the normal display mode. The number of gray-scale levels can be increased for a single subpixel using an analog voltage gray-scale technique. The new pixel with VLC memory is integrated under a small reflective electrode in a high transmissive aperture ratio (39%) 3.17 inch transflective HVGA (320 × RGB × 480 subpixels) panel by using a standard low-temperature-polysilicon process based on 1.5 μm rule. No additional process steps are required. The VLC circuit in each pixel enables simultaneous refresh with a very small change in pixel voltage, resulting in a two-orders-of-magnitude reduction in panel power for a 64-color image (4 levels per 1 bit memory) display. The measured panel power is as small as 210 μW and is smaller by about one order of magnitude than the system power in standby mode (~ 5 mW). The advanced transflective thin film transistor (TFT) LCD using the newly proposed pixel with VLC can display high-quality multi-color images anytime and anywhere, due to its low-power-consumption and good outdoor readability. It also offers high transmittance for indoor applications. This technology will become increasingly important in displays for future mobile information devices such as e-books and personal navigation systems.

Chapter 6 describes a floating gate MOSTFT oxide semiconductor (FLOTOS) memory that consists of wide-band-gap indium-gallium-zinc oxide (IGZO) MOSTFTs and a storage capacitor. The FLOTOS is fabricated using a standard IGZO MOSTFT process without any additional process or mask steps. The proposed precharge-assisted PAC technique makes it possible to realize an infinite number of write cycles and a low-power write operation with a low bit-line voltage of 5 V. Furthermore, excellent data retention longer than 10 h is obtained at 60 °C even under the worst bias-stress conditions.

This is due to the ultra-low off-state leakage (2.8×10^{-20} A/μm) of the IGZO MOSTFTs, which is estimated to be smaller by more than 7 orders of magnitude than that of polycrystalline silicon MOSTFTs. The off-state leakage current of the IGZO MOSTFT is expected to be further reduced to 3×10^{-24} A/μm at 60 °C from a previous study. The oxide semiconductor memory has possibility to store charge on a storage capacitor for 10 years without refreshing similarly to a nonvolatile memory. It is expected that FLOTOS will become a vital technology to meet the demands for ever thinner and more lightweight mobile devices.

Finally, in conclusion, Chapters 2 – 4 described the development of nonvolatile memory technologies that were developed to allow the CPU to execute a large volume instructions directly without going through volatile memory and to back up the data being processed without chip-to-chip communication, which leads to reduced dynamic power during standby and reduced latency. Chapter 5 described a refresh rate reduction technique based on the charge-loss-compensate memory that was developed to enable long-time continuous still-images to be displayed without sacrificing battery life. Chapter 6 described an oxide semiconductor memory that can be fabricated at low temperature, was developed to offer an infinite number of read/write operations combined with an ultra-low-power consumption. These research results contribute to both the creation of new functionality and a reduction in power consumption for mobile applications.

Power consumption is becoming the limiting factor for the functionality of ultra-low-power mobile devices and reducing standby power is the key challenge for further power reduction. The oxide semiconductor memory has the potential to eliminate the static and dynamic power consumption caused by the inherent leakage of silicon MOSFET, because the band-gap of oxide semiconductor films is as large as the effective barrier height for electrons at the Si/SiO_2 interface. The quasi-nonvolatile system combining an oxide semiconductor memory with power gating techniques in a fine manner is the most promising approach to realizing next-generation ultra-low-power wearable devices that integrates all circuit elements on a common plastic substrate.

Lists of Publications, Conferences and Patents

I. PUBLICATIONS

[1] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "Source-side Injection Single-Polysilicon Split-gate Flash Memory," to be published in *Japanese Journal of Applied Physics*, vol. 43, Mar. 2014.

[2] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "Scalable Virtual-Ground Multi-Level-Cell Floating-Gate Flash Memory," *IEEE Transactions on Electron Devices*, vol. 60, no. 8, pp. 2518 − 2524, Aug. 2013.

[3] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "Study of Novel Floating-Gate Oxide Semiconductor Memory Using Indium-Gallium-Zinc Oxide for Low-Power System-on-panel Applications," *Japanese Journal of Applied Physics*, vol. 52, pp. 094101-1 – 094101-7, Aug. 2013.

[4] Y. Yamauchi, N. Ueda, Y. Ogawa, K. Tanaka, and K. Yamamoto, "A Novel Pixel Memory using Integrated Voltage-Loss-Compensation (VLC) Circuit for Ultra-Low-Power TFT-LCDs," *Journal of the SID*, vol. 19, pp. 57 – 62, Jan. 2011.

[5] Y. Yamauchi, K. Tanaka, K. Sakiyama, and R. Miyake, "A Versatile Stacked Storage Capacitor on FLOTOX Cell for Megabit NVRAMs," *IEEE Transactions Electron Devices*, vol. 39, no. 12, pp. 2791 – 2796, Dec. 1992.

II. INTERNATIONAL CONFERENCES

[1] Y. Yamauchi, Y. Kamakura, and T. Matsuoka, "A Source-side Injection Single-poly Split-Gate Cell Technology for Embedded Flash Memory," *International Conference Solid State Devices and Materials*, 2013, pp. 110 − 111.

[2] Y. Isagi, Y. Yamauchi, and Y. Kamakura, "Monte Carlo Simulation of Program Disturb in Contact-Less Virtual Ground NOR Flash Memory," *International*

Conference on Simulation of Semiconductor Processes and Devices, 2012, pp. 7 − 10.

[3] Y. Yamauchi and N. Ueda, "(Invited) Low Power Driving Technologies for High Resolution AM-LCD," *Active-Matrix Flat-panel Display and Device*, 2010, pp. 85 − 88.

[4] N. Ueda, Y. Ogawa, K. Tanaka, K. Yamamoto, and Y. Yamauchi, "A Novel Multi-level in Pixel Technology for Ultra low Power LTPS TFT LCD," *SID International Symposium Digest of Technical Papers*, 2010, pp. 615 − 618.

[5] N. Ito, Y. Yamauchi, N. Ueda, K. Yamamoto, Y. Sugita, T. Mineyama, A. Ishihama, and K. Moritani, "A Novel Program and Read Architecture for Contact-Less Virtual Ground NOR Flash Memory for High Density Application," *Symposium on VLSI Circuits Digest of Technical Papers*, 2006, pp. 204 − 205.

[6] N. Ueda, Y. Yamauchi, and T. Ohm, "A New Two-Step Round Oxidation STI Technology for Highly Reliable Flash Memory," *International Conference Solid State Devices and Materials*, 2001, pp. 538 − 539.

[7] N. Ueda, Y. Saito, M. Hirayama, Y. Yamauchi, S. Sugawara and T. Ohmi, "Highly Reliable MOS Trench Gate FET by Oxygen Radical Oxidation," *International Conference Solid State Devices and Materials*, 2001, pp. 164 - 165.

[8] Y. Yamauchi and K. Sakiyama, "A Novel NOR Virtual-Ground Array Architecture for High Density Flash," *International Conference Solid State Devices and Materials*, 1996, pp. 269 − 271.

[9] Y. Yamauchi, M. Yoshimi, S. Sato, H. Tabuchi, N. Takenaka, and K. Sakiyam, "A New Cell Structure for Sub-quarter Micron High Density Flash Memory," *International Electron Devices Meeting Technical Digest*, 1995, pp. 267 − 270.

[10] Y. Yamauchi, M. Yoshimi, K. Omori, S. Nitta, N. Mizukoshi, and K. Suzuki, "A Novel Source-side Injection Split Gate Flash Cell Technology for High Density Low Voltage Applications," *International Conference Solid State Devices and Materials*, 1994, pp. 910 − 912.

[11] Y. Ma, C.S. Pang, J. Pathak, S.C. Tsao, C.F. Chang, M. Yoshimi, and Y. Yamauchi, "A Novel High Density Contactless Flash Memory Array Using Split-Gate Source-Side-Injection Cell for 5V-only Applications," *Symposium on VLSI Technology Digest of Technical Papers*, 1994, pp. 49 − 50.

[12] Y. Yamauchi, K. Tanaka, H. Shibayama, and R. Miyake, "A 5V-Only Virtual Ground Flash Cell with An Auxiliary Gate for High Density and High Speed Application," *International Electron Devices Meeting Technical Digest*, 1991, pp. 319 − 322.

[13] Y. Yamauchi, N. Mizukoshi, K. Ohmori, and J. Taketani, "A Highly Reliable Stacked storage Capacitor on FLOTOX Cell Technology for Megabit NVRAMs," *IEEE Non-Volatile Semiconductor Memory Workshop*, 1992.

[14] K. Fukumoto, M. D. Eby, M. J. Griffus, G. N. Pham, Y. Yamauchi, K. Tanaka, K. Sakiyama, Y. Torimaru, and R. Miyake, "A 256K-bit Non-Volatile PSRAM with Page Recall and Chip Store," *Symposium on VLSI Circuits Digest of Technical Papers*, 1991, pp. 91 − 92 .

[15] Y. Yamauchi, K. Tanaka, and K. Sakiyama, "A Highly Reliable Stacked Storage Capacitor on FLOTOX Cell Technology for Megabit NVRAMs," *IEEE Non-Volatile Semiconductor Memory Workshop*, 1991.

[16] Y. Yamauchi, H. Tateoka, K. Tanaka, K. Sakiyama, and R. Miyake, "A 4M Bit NVRAM Technology using a Novel Stacked Capacitor on Selectivity Self-Aligned FLOTOX Cell Structure," *International Electron Devices Meeting Technical Digest*, 1990, pp. 931 − 933.

[17] Y. Yamauchi, H. Ishihara, K. Tanaka, K. Sakiyama, and R. Miyake, "A Versatile Stacked Storage Capacitor on FLOTOX Cell for Megabit NVRAMs," *International Electron Devices Meeting Technical Digest*, 1989, pp. 595 − 598.

[18] Y. Yamauchi, K. Tanaka, and K. Sakiyama, "A New Cell Technologies for Implementation in High Density NVRAMs," *IEEE Non-Volatile Semiconductor Memory Workshop*, 1989.

[19] Y. Yamauchi, K. Tanaka, K. Sakiyama and R. Miyake, "A novel NVRAM cell technology for high density applications," *International Electron Devices Meeting Technical Digest*, 1988, pp. 416 − 419.

III. UNITED STATES PATENTS (REGISTERED)

[1] N. Ueda and Y. Yamauchi, "Semiconductor Memory and Process for Manufacturing the Same," United States Patent 8610197, 2013.

[2] Y. Yamauchi, "Nonvolatile Semiconductor Memory Device," United States Patent 8390052, 2013.

[3] Y. Yamauchi, "Pixel Circuit and Display Device," United States Patent 8384835, 2013.

[4] Y. Yamauchi, "Display Device", United States Patent 8339531, 2012.

[5] Y. Yamauchi, N. Ueda, and F. Nakano, "Pixel Circuit and Display Apparatus," United States Patent 8310638, 2012.

[6] N. Ueda and Y. Yamauchi, “Nonvolatile Semiconductor Memory Device, Manufacturing Method Thereof and Method of Programming Information into the Memory Device,” United States Patent 7728378, 2010.

[7] K. Yamamoto, N. Ito, N. Ueda, and Y. Yamauchi, “Semiconductor Memory Device,” United States Patent 7630243, 2009.

[8] N. Ueda and Y. Yamauchi, “Memory Cell Having First and Second Capacitors with Electrodes Acting as Control Gates for Nonvolatile Memory Transistors,” United States Patent 7612397, 2009.

[9] Y. Yamauchi, “Nonvolatile Semiconductor Memory and Method of Operating the Same,” United States Patent 7326991, 2008.

[10] N. Ueda, N. Ito, and Y. Yamauchi, “Semiconductor Memory Device,” United States Patent 7283391, 2007.

[11] Y. Yamauchi, “Semiconductor Memory Device Having Insulating Film of Varying Thickness over Bit Lines,” United States Patent 7276761, 2007.

[12] K. Yamamoto, N. Ito, and Y. Yamauchi, “Virtual Ground Line Type Memory Device with High Speed Readout Circuit,” United States Patent 7224611, 2007.

[13] Y. Sugita and Y. Yamauchi, “Nonvolatile Semiconductor Memory Device with Floating Gate and Two Control Gates,” United States Patent 7187029, 2007.

[14] Y. Yamauchi, “Method of Writing Data to a Non-volatile Semiconductor Memory,” United States Patent 7151693, 2006.

[15] N. Ito, K. Yamamoto, and Y. Yamauchi, “Semiconductor Readout Circuit,” United States Patent 7057944, 2006.

[16] N. Ueda, Y. Sugita, and Y. Yamauchi, “Non-volatile Semiconductor Memory and Process of Fabricating the Same,” United States Patent 7002204, 2006.

[17] Y. Yamauchi, “Nonvolatile Semiconductor Memory and Method of Operating the Same,” United States Patent 6952031, 2005.

[18] Y. Yamauchi, “Non-volatile Semiconductor Memory and Method of Operating the Same,” United States Patent 6894913, 2005.

[19] Y. Sugita and Y. Yamauchi, Nonvolatile Semiconductor Memory Device and Process for Producing the Same, United States Patent 6873007, 2005.

[20] Y. Yamauchi and N. Ito, “Nonvolatile Semiconductor Storage Device and Test Method therefor,” United States Patent 6643175, 2003.

[21] Y. Sugita and Y. Yamauchi, “Nonvolatile Semiconductor Memory Device, Process of Manufacturing the Same and Method of Operating the Same,” United States Patent 6531732, 2003.

[22] Y. Yamauchi and N. Ito, “Nonvolatile Semiconductor Storage Device and Test Method therefor,” United States Patent 6512692, 2003.

[23] Y. Yamauchi, “Nonvolatile Semiconductor Memory, Method of Reading from and Writing to the Same and Method of Manufacturing the Same,” United States Patent 6493264, 2002.

[24] K. Yamamoto, N. Itoh, and Y. Yamauchi, “Nonvolatile Semiconductor Storage Device,” United States Patent 6438035, 2002.

[25] Y. Yamauchi, “Cell Array, Operating Method of the Same and Manufacturing Method of the Same,” United States Patent 6392927, 2002.

[26] Y. Yamauchi, “Method of Making Nonvolatile Semiconductor Memory,” United States Patent RE37199, 2001.

[27] Y. Yamauchi, “Nonvolatile Semiconductor Memory and Driving Method and Fabrication Method of the Same,” United States Patent 6101128, 2000.

[28] Y. Yamauchi, “Non-volatile Memory and Method for Operating the Same,” United States Patent 6009011, 1999.

[29] Y. Yamauchi, M. Yoshimi, S. Sato, and K. Sakiyama, “Nonvolatile Semiconductor Memory with a Floating Gate that has a Bottom Surface that is Smaller than the Upper Surface,” United States Patent 5962889, 1999.

[30] Y. Yamauchi, “Method of Making Nonvolatile Semiconductor Memory,” United States Patent 5877054, 1999.

[31] N. Ueda and Y. Yamauchi, “DRAM and MROM cells with Similar Structure,” United States Patent 5606193, 1997.

[32] Y. Yamauchi, “Method for Manufacturing an Asymmetric Non-volatile Memory,” United States Patent 5510284, 1996.

[33] Y. Yamauchi, “Method for Manufacturing a Non-Volatile Memory Having a Floating Gate Oxide Type FET Provided with a Tunnel Oxide Film,” United States Patent 5424232, 1995.

[34] Y. Yamauchi, “Method for Forming a Semiconductor Device,” United States Patent 5422297, 1995.

[35] Y. Yamauchi, “Nonvolatile Memory, Method of Fabricating the Same, and Method of Reading Information from the Same,” United States Patent 5414286, 1995.

[36] Y. Yamauchi, K. Tanaka, K. Sakiyama, A. Ayukawa, “Process for fabricating Nonvolatile Random Access Memory Having a Tunnel Oxide Film,” United States Patent 5411904, 1995.

[37] Y. Yamauchi, K. Tanaka, and K. Sakiyama, “Non-Volatile Memory,” United States Patent 5401993, 1995.

[38] Y. Yamauchi, “Nonvolatile Memory and a Method of Writing Data thereto,” United States Patent 5400280, 1995.

[39] Y. Yamauchi, “Non-volatile Memory,” United States Patent 5,338,952, 1994.

[40] K. Tanaka, Y. Yamauchi, and K. Sakiyama, “Non-Volatile Semiconductor memory,” United States Patent 5331181, 1994.

[41] Y. Yamauchi, “Semiconductor Memory and Process for Manufacturing the Same,” United States Patent 5303186, 1994.

[42] M. Yoshimi, Y. Yamauchi, and K. Omori, “Method of Making a Stacked Semiconductor Nonvolatile Memory Device,” United States Patent 5290721, 1994.

[43] N. Ueda, Y. Yamauchi, and K. Tanaka, “Semiconductor Memory with Increased Capacitive Storage Capabilities and Reduced Size,” United States Patent Patent 5282159, 1994.

[44] Y. Yamauchi, “Non-Volatile Memory and Method of Manufacturing the Same,” United States Patent 5268585, 1993.

[45] Y. Yamauchi, “Semiconductor Memory Device with Volatile Memory and Non-Volatile Memory in Latched Arrangement,” United States Patent 5262986, 1993.

[46] Y. Yamauchi, “Method of Operating a Semiconductor Memory Device,” United States Patent 5251171, 1993.

[47] Y. Yamauchi, K. Tanaka, K. Sakiyama, and K. Fukumoto, “Semiconductor memory device,” United States Patent 5181188, 1993.

[48] Y. Yamauchi, K. Tanaka, K. Sakiyama, “Compact Nonvolatile Semiconductor Memory Device using Stacked Active and Passive elements,” United States Patent 5172199, 1992.

[49] Y. Yamauchi, K. Tanaka, K. Sakiyama, “Semiconductor Memory Device Having a Volatile Memory Device and a Non-Volatile Memory Device,” United States Patent 5140552, 1992.

[50] S. Ohnishi, K. Tanaka, Y. Yamauchi, and K. Sakiyama, “Method for Forming Isolation Region of Semiconductor Device,” United States Patent 5139964, 1992.

[51] M. Tanigawa H. Tateoka, K. Sakiyama, S. Ohnishi, Y. Yamauchi, and K. Tanaka, “Process for Producing Semiconductor Devices,” United States Patent 5108783, 1992.

[52] K. Sakiyama, K. Tanaka, and Y. Yamauchi, “Semiconductor Device Having A Volatile Memory Device and A Non-Volatile Memory Device,” United States Patent 5075888, 1991.

[53] Y. Yamauchi and K. Sakiyama, “Semiconductor Memory Device with Capacitor over Non-Volatile Memory Cell Gate Structure,” United States Patent 5063425, 1991.

[54] Y. Yamauchi, “Semiconductor Memory Device,” United States Patent 5065201, 1991.

[55] Y. Yamauchi, K. Tanaka, and K. “Sakiyama, Semiconductor Memory Device,” United States Patent 5043946, 1991.